UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

DOCTORAL THESIS

Contributing factors of teachers' acceptance intention to gamified learning tools in secondary schools

A scale development study

Zhanni Luo


UC
UNIVERSITY OF
CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

CONTRIBUTING FACTORS OF TEACHERS' ACCEPTANCE INTENTION TO GAMIFIED LEARNING TOOLS IN SECONDARY SCHOOLS

A scale development study

STUDENT NAME: Zhanni Luo

STUDENT NUMBER: 98799800

COURSE CODE: EDU790

DEPARTMENT: School of Educational Studies and Leadership

SUPERVISORs: Dr. Billy O'Steen

Dr. Cheryl Brown

DOCTORAL THESIS

Contributing factors of teachers' acceptance intention to gamified learning tools in secondary schools

A scale development study

Zhanni Luo

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

ABSTRACT

In the past decades, educationalists were fighting with video games since students naturally prefer to play video games rather than spending time on "more meaningful" activities such as book reading. After the long-term fruitless struggle, educationalists began to consider using the engaging characteristics of video games on top of formal education for enhanced teaching results (Rueckert, Pico, Kim, & Calero Sánchez, 2020). Under this guideline, the concept *gamification* gained popularity in the field of education for both practitioners and researchers (Baldauf, Brandner, & Wimmer, 2017; Cruaud, 2018; Honório, Moura, Brito, Menezes, & de Barros, 2018; Werbach, 2014), which refers to the use of game design elements in non-game contexts (Deterding, Sicart, et al., 2011). Accordingly, *gamified learning*, also named educational refers to the use of gamification for educational purposes. Though gamification does not necessarily connect with digital devices, *gamified learning tools* in the current thesis refer to the education websites, systems, software, or mobile applications that use gamification techniques.

Despite gathering academic interests, the actual implementation of gamification in school contexts is not prevalent (Luo, Brown, & O'Steen, 2021). It is important to understand and improve teachers' willingness of accepting a new approach or a new technology since teachers' attitudes affect the selection, implementation, and evaluation of the approach/technology. The effectiveness of a new approach/technology is also "largely depend on" teachers' acceptance (Bourgonjon et al., 2013, p. 22). However, a large number of previous studies focused on the technology and its impact, with the role of teachers neglected (Bourgonjon et al., 2013).

Bearing the aforementioned gap in mind, this author aims to explore the factors influencing secondary school teachers' acceptance intention to gamified learning tools in formal education. Two approaches were conducted: the open-ended survey that involved 347 teachers for general perspectives, and the face-to-face interview that involved 14 teachers for detailed explanations. This author obtained abundant data that reveal in-depth explanations, which were summarised into 16 themes. The specification of the 16 themes enables researchers to better understand what motivates teachers to use this new technology, what hinders teachers from accepting it, how to enhance the motives, and how to address the hindrances. The interview responses were furtherly being shaped into survey items, which were drafted and examined following the typical scale development and evaluation guidelines by DeVellis (2003) and Robertson (2017). The survey study involved 512 valid participants. This study aims to answer three main research questions (RQs) as follows:

RQ 1: What are the factors contributing to teachers' acceptance intention of gamified learning tools in secondary schools?

RQ 2: What items should be contained in the scale measuring teachers' acceptance intention of gamified learning tools in secondary schools (TAI-GLT)?

RQ 3: How reliable and valid is the scale measuring teachers' acceptance intention of gamified learning tools in secondary schools (TAI-GLT)?

It is noteworthy that the scale evaluation revealed surprising results. Based on the interview responses, this author conceptualised a framework containing six main factors that contribute to teachers' acceptance intention to educational gamification, including perceived enjoyment, perceived usefulness (PU), perceived ease of use (PEoU), perceived risks, facilitating conditions, and control variables. Surprisingly, the results indicated that only two factors can be retained in the proposed framework (PU and PEoU). However, PU and PEoU in the current research contexts brought new connotations, which is of significance for future studies. For example, while most of the previous studies

use PEoU to indicate the PEOU of *the technological or the devices*, findings of this study revealed that the concept of PEoU also indicates the PEoU of the implementation process. In other words, the perceived ease of use of gamified learning tools in secondary-school contexts is connected with not only the cumbersomeness in using digital devices but also the effort-demandingness of dealing with correspondent issues such as classroom discipline management.

To furtherly understand the topic, this author added an interview and a survey study to explore the determinants of PU and PEoU of gamified learning tools, which involved 263 and 239 participants respectively. Under the suggestion of field experts, this author selected foreign language learning as the specific subject. The newly added study aims to answer the following three research questions (RQs):

RQ 4: What are the determinants of teachers' perceived usefulness (PU) of gamified learning tools (GLT) for foreign language learning in formal education? What items should be contained in the correspondent measuring scale PU-GLT?

RQ 5: What are the determinants of teachers' perceived ease of use (PEoU) of gamified learning tools (GLT) for foreign language learning in formal education? What items should be contained in the correspondent measuring scale PEoU-GLT?

RQ 6: How reliable and valid are the scales PU-GLT and PEoU-GLT?

The current study presents five main outputs: a redefinition of gamified learning and gamified learning tools; a summary of factors influencing whether secondary teachers accept or resist using gamified learning tools; a theoretical framework that illustrates the contributing factors; a validated survey that measures teachers' acceptance intention to gamified learning tools; and two validated survey that measure teachers perceived usefulness (PU) and perceived ease of use (PEoU) of gamified learning tools respectively. The two scales were named PU-GLT and PEoU-GLT.

Keywords: Gamification, gamified learning tools, technology acceptance, perceived usefulness (PU), perceived ease of use (PEoU), scale development

KEY IDEA

Gamification & technology acceptance intention

Factors influencing teachers' acceptance intention

Scales: construction & evaluation

TABLE OF CONTENTS

Tables

Figures

ATTESTATION OF AUTHORSHIP

This author hereby declares that this thesis is her own work that contains no materials published before or owned by other people.

The thesis work was conducted from September 2018 to December 2021 under the supervision of Associate Professor Billy O'Steen and Associate Professor Cheryl Brown at the University of Canterbury, Christchurch, New Zealand.

RELATED PUBLICATION

Luo, Z., Brown, C., & O'Steen, B. (2021). Factors contributing to teachers' acceptance intention of gamified learning tools in secondary schools: An exploratory study. *Education and Information Technologies, 26*(5), 6337-6363. https://doi.org/10.1007/s10639-021-10622-z

ACKNOWLEDGEMENT

This author, Zhanni Luo, would like to thank her supervisors Billy O'Steen and Cheryl Brown for their kind support during the doctoral study.

CHAPTER 1

Introduction

01

Introduction

This chapter as the inaugural part of the thesis is to give an overall introduction to the background, concept, and benefits of gamification, research gaps presented in previous studies, research aims, research significance, research questions (RQs), and thesis structure arrangement.

CHAPTER ONE: INTRODUCTION

1.1. Background overview

In the last three decades, games, especially video games, are gaining attention in the field of education. Video games have been viewed as a major distraction from more 'worthy' activities such as doing homework or reading books; however, educationalists began to ask why video games are so engaging and how to make good use of the lure of video games to support teaching and learning (Kirriemuir & McFarlane, 2004).

It has been argued that games are intrinsically educational in nature (Carr, 2011), and game-like experiences have the power to motivate learners to take part in learning activities (de Freitas & Liarokapis, 2011). It is desirable to use the motivational power of digital games to "making learning fun", or use video games as simulations to achieve "learning through doing" (Kirriemuir & McFarlane, 2004, p. 21).

Afterward, the concept of gamification gained widespread development, which does not mean using specific games to teach. Deterding, Sicart, et al. (2011) specifically highlighted the gamification definition as "the use of design (rather than game-based technology or other game-related practices) elements (rather than fully developed games) characteristic for games (rather than play or playfulness) in non-game contexts (regardless of specific usage intentions, contexts, or implementation media)" (p. 5).

The benefits of gamification have been explored in a number of studies. Generally, gamification has positive impacts on engagement and motivation (Deterding, Sicart, et al., 2011; Lopez & Tucker, 2019; McGonigal, 2011), participation (Lister, 2015), achievement (Hamari, 2017; Hattie & Yates, 2013), satisfaction (Sailer, Hense, Mayr, & Mandl, 2017), enjoyment (Clements, Ahmed, & Henderson, 2017; Peng, Song, Kim, & Day, 2016), performance (Lopez & Tucker, 2019; Pedersen, Rasmussen, Sherson, & Basaiawmoit, 2017) and problem-solving skills (McGonigal, 2011).

According to McGonigal (2011), gamification is of great potential in increasing user satisfaction, conveying optimism (by providing a sense of achievement or the hope of experiencing success), facilitating social interactions and providing meanings (by assisting users to go beyond their personal possibilities). Gamification is also evidenced important in supporting behaviour changes and learning processes (Blohm & Leimeister, 2013).

DEFINITION

Gamification: "the use of design (rather than game-based technology or other game-related practices) elements (rather than fully developed games) characteristic for games (rather than play or playfulness) in non-game contexts (regardless of specific usage intentions, contexts, or implementation media)" (Deterding, Sicart, Nacke, O'Hara, & Dixon, 2011, p. 5).

Gamification is also widely implemented in educational contexts (Buckley & Doyle, 2017; DomíNguez et al., 2013). Even though the definition of gamification refers to the use of game elements in non-game contexts, which involves digital-free activities as well, the majority of gamified learning services or tools are still based on digital equipment such as smartphones and computers. Generally speaking, educational gamification is "a clear trend" of recent years (Purgina, Mozgovoy, & Blake, 2020, p. 126) for its great potential in engaging learners (Dehghanzadeh, Fardanesh, Hatami, Talaee, & Noroozi, 2019; Y.-L. Wu & Wang, 2014), especially for the subjects requiring long-time commitment (e.g. foreign language) (Purgina et al., 2020).

1.2. Research aims and methods

This study aims to investigate the contributing factors to teachers' acceptance intention of gamified learning tools in China. Three main academic outputs were expected:

1) A summary of teacher perceptions of using gamified learning tools in secondary schools.
2) A theoretical framework illustrating the factors contributing to teachers' acceptance intention to gamified learning tools in secondary schools.
3) A validated scale measuring teachers' acceptance intention to gamified leanring tools in secondary schools.

Based on literatures, the contributing factors were synthesised into a theoretical framework illustrating the potential factors contributing to teachers' acceptance intention to gamified learning tools. Based on interview responses collected from the target participants, the framework was extended into a scale measuring teachers' acceptance intention of gamified learning tools. After confirming with two experts in the field, the scale was administered to the target participant, followed by a thorough scale evaluation process proposed by DeVellis (2003).

Generally speaking, the current study used the exploratory sequential mixed method. This author collected qualitative data via interviews and quantitative data via survey. The interview responses provided inspirations and evidences for survey item generation and survey item analysis, which significantly improves the reliability and validity of the constructed survey items. This author also consulted experts in the field at each step for enhanced reliability and validity.

1.3. Research rationale

1.3.1. Why focuses on gamified learning rather than game-based learning (GBL)?

As will be detailed in the Literature Review section, while the game-based learning tends to involve video games (Luo, 2021), the use of video games for educational purposes brings new challenges such as poor integration of game features with pedagogical content (Kim & Lee, 2015; B. Morris, Croker, Zimmerman, Gill, & Romig, 2013; NFER, 2009; Van Rosmalen, Wilson, & Hummel, 2015).

1.3.2. Why focuses on teachers?

Quazi and Talukder (2011) pointed out that the potential users of an innovation are the "focal point in making technological innovations a success" (p. 34), so gamification studies should not only about the design of the gamified tools but also about the perceptions of potential users. In the current context, the main potential users of gamified learning tools are teachers and students.

Compared with students, teachers are "neglected in gamification research with few research focused on them" (José Martí-Parreño, Seguí-Mas, & Seguí-Mas, 2016, p. 682), even though they are of great importance (Hu, Clark, & Ma, 2003). Shaban and Egbert (2018) even stated that "the probability that technology will be adopted and used efficiently and effectively by teachers is often low" (p. 234).

Teachers play a critical role in introducing pedagogical innovations in the classroom, especially when the innovations are related to new technologies (José Martí-Parreño et al., 2016). To be more specific, the effectiveness of game-based learning or gamified learning "largely depend on" teachers' acceptance (Bourgonjon et al., 2013, p. 22) since the new form of technology-driven innovations can challenge teachers' attitudes and provoke anxiety of uncertain outcomes (Sánchez-Mena & Martí-Parreño, 2017), which therefore lead to negative attitudes toward gamified learning, and eventually impede the realization of gamification' full potentials in classrooms (Y.-H. Lee, Hsieh, & Hsu, 2011; Sugar, Crawley, & Fine, 2004; Tuparova, Tuparov, Veleva, & Nikolova, 2018).

In other words, teachers play a crucial role in technology acceptance: if teachers do not see the value of the technology, the tool selection and implementation would be significantly restrained (Ale & Chib, 2011; Huizenga, Ten Dam, Voogt, & Admiraal, 2017; José Martí-Parreño et al., 2016)

Besides, based on this author's observation, theories and frameworks in gamification studies are comparatively complicated, so users are mainly researchers rather than teachers. As a consequence, there is a gap between the gamification being studied in research contexts and the gamification being implemented in real-life practices. Cuban (1986) has criticised that many attempts to introduce innovative technologies to education are top-down activities, which have failed to achieve long-term effects. To address this issue, researchers should consider more about teachers' perspectives of gamification and the factors affecting their acceptance intention.

1.3.3. Why focuses on secondary school teachers in China?

This study can be duplicated in other countries on any group of teachers who implement gamification. This thesis focuses on secondary school teachers in China due to this author's accessibility to potential participants.

1.3.4. Why develops a scale?

Small sample size brings limited generalisability and "underpowered" analysis (Sailer & Homner, 2020, p. 79), so there is a need of survey studies to obtain large sample outcomes for gamification (Baydas & Cicek, 2019). However, the availability of validated gamification-related scales is an issue: the sample size in studies assessing motivation can be 10,000 in one single study (You & Dörnyei, 2016), while that in gamification studies mainly is mainly below 300.

1.3.5. Why interviews teachers before developing the scale?

One highlight of the current scale-development study is the use of exploratory sequential mixed method, which combines both qualitative and quantitative studies (Creswell & Clark, 2017). The qualitative data were interview responses collected from secondary school teachers in China, which function as the basis constructing the survey items for the quantitative survey study.

The use of exploratory sequential mixed method significantly improves the reliability and validity of the survey items (Mihas, 2019; Moral-Bofill, López-delallave, Pérez-Llantada, & Holgado-Tello, 2020).

Besides, there is a shortage of qualitative studies on how users perceive gamified learning tools. According to Vandercruysse, Vandewaetere, Cornillie, and Clarebout (2013), the perceived usefulness of gamification, one main

factor contributing to the final acceptance intention, depends on the perceptions of users. However, there is a lack of studies on how users perceive the use of gamification in classroom settings, especially teachers.

1.3.6. Why shifts from effectiveness study to technology acceptance study?

> Note: Chapter five is named pilot study, which mainly investigates the effectiveness of gamification in educational practices.
>
> Though seemed far-fetched with the theme "technology acceptance intention", the results were included in the current thesis since they provide insightful information for future studies.

The effectiveness of gamified learning has been debated for a long time, which is an unignorable issue in academic research (C.-M. Chen, Li, & Chen, 2020). As claimed by Baydas and Cicek (2019), previous studies claimed that tend to focus on the effects of educational gamification. Therefore, when first cut into gamification studies, this author planned to focus on the assessment of its effectiveness, too. The literature reading and pilot study lead to a disappointing conclusion: effectiveness assessment might not fit into the current research scope since it tends to involve controlled experiments or longitude studies.

As a compromise, this author shifted the research focus from effectiveness to a related concept "usefulness". In reading literature introducing the usefulness of new technologies, this author noticed that usefulness is an important concept in technology acceptance studies. Therefore, the research focus shifted from effectiveness assessment to technology acceptance intention.

1.3.7. Why conducts a second round of scale development?

This author planned to develop one scale only, namely the TAI-GLT scale measuring teachers' acceptance intention of gamified learning tools in secondary schools.

As briefed above, the evaluation of the TAI-GLT scale revealed unexpected results: though this author proposed six contributing factors to teacher's acceptance intention, the factors confirmed by the statistical analysis were two only: perceived usefulness (PU) and perceived ease of use (PEoU). However, both PU and PEoU contain new connotations in the current context.

To better understand the two concepts, a second round of scale development was conducted, which aimed to explore the determinants of PU and PEoU of teachers in using gamification for education. Since the interview responses indicated that the connotations may vary in different contexts (Nacke & Deterding, 2017), this author conducted the survey study in foreign language learning (FLL) contexts.

Using the same procedure of developing and evaluating the TAI-GLT scale, this author established two scales measuring secondary school teachers' perceived usefulness (PU) and perceived ease of use (PEoU) of gamified learning tools (GLTs). The two scales were named PU-GLT and PEoU-GLT respectively.

1.4. Research significance

The current study is of great significance on the following aspects:

First, it helps clarify the relationship between gamification and education. Though gamification has been clearly defined as the use of *game design elements* in non-game contexts where fully-developed games are not usually present (Deterding, Dixon, Khaled, & Nacke, 2011; DomíNguez et al., 2013; Kapp, 2012), readers or researchers tend to make automatic connections between gamification with digital video games. The use of digital video games for educational practices has been criticised for various reasons, such as homogenous content, poor integration of game features and pedagogical content, questionable justification for some subjects, and side effects (e.g. violence and addiction) (Kirriemuir & McFarlane, 2004). The current study specifically distinguishes the concept of gamification from game-based learning. This change furtherly helps educationalists cut into gamification studies, which promotes the development of gamification in educational areas.

Second, it examines whether the Technology Acceptance Model (TAM) proposed by Davis (1989) applies to educational gamification. The TAM model is a popular framework that describes and predicts users' acceptance intention to innovative technologies.

Third, it enhances the understanding of factors influencing teachers' decisions in deciding whether to use educational technologies, which therefore promotes the eventual adoption of gamified learning tools in secondary schools in China.

Fourth, it helps school principals or educationalists support the adoption and implementation of new educational technologies, via leading them to focus on the alleviation of hindering factors of teachers' acceptance intention to gamified learning tools. By better understanding the contributing factors and making correspondent copying strategies, the degree of technology adoption can be improved (van Braak & Tearle, 2007).

Fifth, it constructs three scales, which are designed to measure teachers' acceptance intention to gamified learning tools, the perceived usefulness of gamified learning tools, and the perceived ease of use of gamified learning tools. The three scales can be directly used in future studies.

1.5. Research questions (RQs)

This author attempted to answer three research questions through a scale-development study. Naming the scale as TAI-GLT (teachers' acceptance intention of gamified learning tools), the three research questions are:

> RQ 1: What are the factors contributing to teachers' acceptance intention of gamified learning tools in secondary schools?
>
> RQ 2: What items should be contained in the scale measuring teachers' acceptance intention of gamified learning tools in secondary schools (TAI-GLT)?
>
> RQ 3: How reliable and valid is the scale measuring teachers' acceptance intention of gamified learning tools in secondary schools (TAI-GLT)?

It is notable that the evaluation of the TAI-GLT scale revealed unexpected results, so further investigations are essential. In evaluating the TAI-GLT scale, it is assumed that six factors contribute to users' acceptance intention to gamification; however, the scale evaluation results indicated that only two factors should be included in the TAI-GLT scale (PU and PEoU).

Even more surprisingly, each of the remaining two factors contains varied connotations in the current research context. To be specific, the perceived usefulness of gamified learning tools in secondary-school contexts is not only about the usefulness in improving academic achievement but also in enhancing the learning experience (e.g., enjoyment). Similarly, the perceived ease of use of gamified learning tools in secondary-school contexts is connected with not only the cumbersomeness in using digital devices but also the effort-demandingness of dealing with correspondent issues such as classroom discipline management.

To better understand the connotations of teachers' PU and PEoU in secondary-school contexts, this author conducted a second-round of scale development, which aims to answer the following three questions:

> RQ 4: What are the determinants of teachers' perceived usefulness (PU) of gamified learning tools (GLT) in secondary schools? What items should be contained in the correspondent scale (PU-GLT)?
>
> RQ 5: What are the determinants of teachers' perceived ease of use (PEoU) of gamified learning tools (GLT) in secondary schools? What items should be contained in the correspondent scale (PEoU-GLT)?
>
> RQ 6: How reliable and valid are the scales PU-GLT and PEoU-GLT?

Note

TAI-GLT: the scale measuring teachers' acceptance intention of gamified learning tools (GLT).
PU-GLT: the scale measuring teachers' perceived usefulness (PU) of gamified learning tools (GLT).
PEoU-GLT: the scale measuring teachers' perceived ease of use (PEoU) of gamified learning tools (GLT).

1.6. Thesis structure

The current thesis contains nine chapters, as detailed in Table 1.

Chapter 1 is the introduction of research backgrounds, research aims, methodology, research rationale, research significance, research questions, and thesis structure arrangement.

Both Chapter 2 and Chapter 3 are Literature Review sections, with the former one focusing on gamification and technology acceptance intention and the latter one focusing on the effectiveness of gamification.

Chapter 2 mainly reviews literatures about gamification and technology acceptance intention. First and foremost, the author introduced and analysed the definitions of involved terms, including game, educational game, game-based learning (GBL), serious game, pervasive game, simulation, edutainment, etc. Since the definitions of gamification are of great variance, the author furtherly redefined the terms gamification and game elements (Section 2.1). Gamification-related frameworks and theories were introduced, such as flow theory, the self-determination theory (SDT), and the mechanics-dynamics-aesthetics (MDA) framework (Section 2.2).

Section 2.3, 2.4, and 2.5 are the main parts being connected with findings and discussion. After discussing the importance of investigating teacher perceptions, this author synthesised teacher perceptions of the advantages of game-based learning and educational gamification, followed by the analysis of factors contributing to teachers' acceptance intention (Section 2.3). Focusing on the Technology Acceptance Model (TAM), this author listed the TAM-expanding frameworks in previous studies to better understand the topic (Section 2.4). The determinants of perceived usefulness (PU) and perceived ease of use (PEoU), the two main constructs of the TAM model, were also investigated (Section 2.5).

Generally speaking, section 2.1 is to define related terms, while section 2.2 is to introduce related theories. The remaining three sections synthesise previous findings of teacher perceptions to educational technologies, discuss the theoretical framework, and list determinants of the two main factors influencing users' technology acceptance (Section 2.3, 2.4, and 2.5).

Chapter 3) mainly investigated the effectiveness of gamification used in educational practices, including what results have been reported in empirical studies (positive, negative, mixed, neutral, etc.) (Section 3.1), how to measure the effectiveness of gamification (Section 3.2), and what are the factors influencing the effectiveness results in gamification studies (Section 3.3).

As will be detailed in Chapter 3, effectiveness assessment was the first attempt this author made when cutting into gamification studies, but this author failed to hatch a meaningful and practical research plan on the topic. Though shifted to technology acceptance intention, the literature review about the effectiveness of gamification was kept due to its value in providing insightful information.

Chapter 4, the Methodology, introduces the method selection, participants, research instrument, data collection, data analysis, and reliability and validity. Since this study is a scale development study, the construction and validation of new scales are important; therefore, Chapter 5 as a sub-section of methodology was separated as an individual chapter, which details the procedure and techniques for scale development and evaluation.

Chapter 6 to 8 are findings and discussions, followed by the conclusion of the whole thesis (Chapter 9).

Table 1. Thesis structure

#	Title	Content
Chapter 1	Introduction	Research background.
		Research aims and objectives.
		Research rationale.
		Research significance.
		Research questions.
		Thesis structure.
Chapter 2	Literature review-1	Concepts and definitions.
	(Gamification and technology acceptance intention)	Gamification-related frameworks and theories.
		***Teacher perceptions to gamified learning tools.**
		***Technology acceptance intention.**
		***Determinants of PU and PEoU.**
Chapter 3	Literature review-2	The effectiveness of gamified learning.
	(Effectiveness of gamification)	Effectiveness assessment of gamified learning.
		Factors influencing the effectiveness results of gamification studies.
Chapter 4	Methodology	Overall method selection: The exploratory sequential mixed method.
		Qualitative method selection: Grounded theory.
		Quantitative method selection: Scale development.
		Participants.
		Research instrument.
		Data collection.
		Data analysis.
		Reliability and validity.
Chapter 5	Scale development techniques	Procedure and techniques for scale development.
		Procedure and techniques for scale evaluation.
Chapter 6	Findings and discussion (1)	Findings and discussion of the interview study.
Chapter 7	Findings and discussion (2)	Findings and discussion of the scale measuring teachers' acceptance intention to gamified learning tools (TAI-GLT).
Chapter 8	Findings and discussion (3)	Findings and discussion of the scales measuring teachers' PU and PEoU of gamified learning tools (PU-GLT and PEoU-GLT).
Chapter 9	Conclusion	Thesis summary.
		Research limitations.
		Implications.
		Suggestion for future studies.

Note: The sentences in bold and with the "*" mark are literature review that closely related with findings and discussion.

CHAPTER 2

Literature Review

02

Literature Review

To make the outcomes of these disjointed literature studies easy to read, this author divided them into five separate sections:

Section 2.1 is the brief introduction and analysis of related concepts. It helps redefine the term gamification and game elements.

Section 2.2 introduces gamification-related theories and frameworks.

Section 2.3 introduces previous findings on teachers' perceptions to educational gamification.

Section 2.4 summarises framework and theories focusing on users' technology acceptance intention.

Section 2.5 introduces the determinants of perceived usefulness (PU) and perceived ease of use (PEoU) in deciding whether to adopt new technologies.

CHAPTER TWO: LITERATURE REVIEW-1

2.1. Concepts and definitions

2.1.2. Definitions of the related terms

There are various terms that relate with gamification, including game, educational game, game-based learning (GBL), serious game, pervasive game, simulation, edutainment, etc. The specific definitions are as follows.

Game. There are various types of games, such as board games and video games. In studies related to gamification, games in most cases refer to video games, digital games, computer games, commercial video games, or mainstream games (Kirriemuir & McFarlane, 2004). Since gamification is defined as the use of game design elements only (namely a part of games rather than the full ones), games in gamification studies are also specified as complete games, full-developed games, or full-fledged games (Kirriemuir & McFarlane, 2004).

There are differences among these terms; for example, according to Kirriemuir and McFarlane (2004), computer games are PC-based while video games are console-based. However, in gamification studies, these terms are used "interchangeably" (Kirriemuir & McFarlane, 2004, p. 8). In the current study, the term digital games will be used as the unified term.

Educational games. Educational games refers the game software that is either developed explicitly as being educational or the game software that can be used for educational purposes (Berg Marklund, 2015).

Game-based learning (GBL). Game-based learning is used to describe the act of "conducting pedagogical work with a game as the core educational tool", which can be followed by words including learning, curriculum, and activities (Berg Marklund, 2015, p. 16). Berg Marklund (2015) suggested that game-based learning is the activity in which used educational games.

Serious game. A serious game is a game in which "the primary goal is to educate, rather than just entertain players" (Fitz-Walter, 2015, p. 26). It is designed for any purpose rather than entertainment (Zyda, 2005). Groh (2012) indicated that serious games are a branch of video games, as they are "full-fledged games for non-entertainment purposes" (p.40). Berg Marklund (2015) concluded that the term serious games is used to describe the "entire range of games with a purpose beyond just providing an engaging experience" (p. 16). Serious games are used "in conjunction to acquiring knowledge and skills applied to a particular discipline, module or educational topic" (Lameras et al., 2017, p. 979).

Researchers suggested that the line between serious games and gamification is also blurry (Fitz-Walter, 2015), though serious games are defined as the use of complete games to educate learners while gamification is to employ the game elements (namely part of a complete game) for educational purposes (Fitz-Walter, 2015; Groh, 2012). Commercial video games can be used for non-entertainment purposes, which can be regarded the same with what serious games aim to achieve (Berg Marklund, 2015). For example, the game "America's Army" is a full-fledged video game

providing virtual military trainings. The players reported that after playing the game, their willingness to join the US Army increased: they are "twice as likely to consider a career in the US Army" (Zyda, 2005, p. 27). From the solider recruitment perspective, the game "America's Army" achieved educational functions, so it can be regarded as a serious game.

Marczewski (2015) attempted to distinguish gamification from serious games, games, or playful design. According to Marczewski (2015), gamification is a purposeful activity without gameplay, while serious games are purposeful games with gameplay. However, how to define "gameplay" remains an issue (see Figure 1). Therefore, the distinction between serious games and gamification may be personal and subjective that depends on users' "perceptions and enactments" (Groh, 2012, p. 40).

Simulation. Also named simulation game. Simulations is a branch of serious gaming, which refers to the representation of reality (Deshpande & Huang, 2011). "A simulation game is a game, which has elements like score, performance rating, conflict, and payoff, and simulates a real world situation for decision-making or alternative evaluation" (Deshpande & Huang, 2011, p. 400).

Edutainment. Edutainment is the combination of educational software "lightly sprinkled with game-like interfaces and cute dialog" (Zyda, 2005, p. 29).

Playful design. Researchers debated how to distinguish the concept gamification from playful design. Fitz-Walter (2015) proposed that gamification is close to "structured activities with explicit rules" (p.24) while playful design is related to playfulness, which contains unstructured and spontaneous activities. Generally speaking, playful design aims to elicit playful experience when designing the product, without specific rules. However, gamification has also been named playful design, so the two terms can be overlapping (Deterding, Sicart, et al., 2011).

Gamification. In the extant literature, the most widely-recognised definition of gamification is the use of *game design elements* in non-game contexts to increase user experience and engagement where fully-developed games are not usually present (Deterding, Dixon, et al., 2011; DomíNguez et al., 2013; Kapp, 2012).

The concept of gamification gained widespread attention, which does not mean using specific games to teach; instead, it refers to using *game design elements* in non-game contexts to increase user experience and engagement where fully-developed games are not usually present (Deterding, Dixon, et al., 2011; Deterding, Sicart, et al., 2011; DomíNguez et al., 2013; Kapp, 2012). To be specific, gamification is defined as "the use of design (rather than game-based technology or other game-related practices) elements (rather than fully developed games) characteristic for games (rather than play or playfulness) in non-game contexts (regardless of specific usage intentions, contexts, or implementation media)" (Deterding, Sicart, et al., 2011, p. 5).

Gamification in the context of learning "can be referred to as gamified learning" (Sailer & Homner, 2020, p. 78).

Figure 1. Serious games vs. gamification quadrant (Marczewski, 2015)

2.1.2. Necessity of using gamification rather than GBL

Both educational gamification and game-based learning (GBL) are of great potential in engaging learners; however, the two concepts are not interchangeable in educational practices. To be more specific, while GBL in most cases involves video games, the use of GBL for educational purposes have varied disadvantages (see Table 2).

The NFER survey (NFER, 2009) investigated the perceptions of 1632 primary teachers on the use of computer games in teaching in United Kingdom. The results reported that a majority of teachers (85%) were positive about the impact of video games on learning, however, more than half of teachers (64%) never used games in classrooms for educational purposes. There are several barriers hindering teachers from using GBL, including the expense of video games (74%), difficulties in obtaining licenses (69%), teachers' lack of knowledge (56%), inappropriate nature of game content (51%), poor integration of game content and learning objectives (50%) and so on. Van Rosmalen et al. (2015) concluded that the main barriers hindering the development of GBL in classroom settings are teachers' lack of expertise, aspects of the school system, financial barriers, and technical barriers. Van Rosmalen et al. (2015) also mentioned that the integration of games with educational content is a problem: even though teachers and students are the main characteristics in GBL with digital games, the games are almost exclusively designed by professionals. Therefore, the integration of games with curriculum is very difficult to be managed (Van Rosmalen et al., 2015).

Other drawbacks in the implementation of GBL in classrooms were also identified by researchers, such as the unconvincing effectiveness of GBL (Carr, 2011), the time-demanding nature in class preparation, the dangers of getting addicted, the distractions of repetitive or irrelevant tasks, the negative impact of violence, the irrelevance with educational contexts, the harm to intrinsic motivation, etc. (Kirriemuir & McFarlane, 2004).

Table 2. Disadvantages of using video games in educational contexts

Perspective	Disadvantages
Game design	Simple, homogenous, or irrelevant content.
	Repetitive or poorly-designed tasks.
	Side effects (e.g. violence, addiction).
Curriculum	Poor integration of game features and curriculum.
	Questionable justification in some subjects.

	Untested effectiveness compared with traditional face-to-face teaching.
Teachers	Lack of time to identify a relevant game and get familiar with it. Lack of incentives or support. Lack of expertise in the use of educational games. Difficult to convince shareholders or school managers.
School	Financial barriers in purchasing, upgrading, and maintaining the software/hardware.

2.1.3. Redefining gamification

Though gamification has been defined specifically as the use of *game design elements* in non-game contexts (Deterding, Dixon, et al., 2011; Deterding, Sicart, et al., 2011; DomíNguez et al., 2013; Kapp, 2012), the definition has been criticised for being blurred (Werbach, 2014).

Werbach (2014) suggested that though Deterding, Dixon, et al. (2011) separate gamification (which involves only parts of games) from serious games (which involve full-developed games), the dividing line is "often difficult to see" (Werbach, 2014, p. 267). The author took Duolingo as an example, saying that Duolingo as a language learning tool is game-like but not immersive simulations like a typical serious game, so it can be seen as a gamification example without struggling over whether it involves non-game contexts. Werbach (2014) furtherly stressed that "there is no need to insist that games cannot be gamified" (p. 267). The problem is, if so, the definition of gamification becomes blurred.

Werbach (2014) also questioned the concept of "game elements" in the definition. Werbach (2014) pointed out that though gamification was defined as the use of game design elements in non-game context, "there is no universal list of game elements" (Werbach, 2014, p. 267). Therefore, different studies employed different game elements with different interpretation for different purposes, which makes the research conclusions of high variance. Some definitions, according to Werbach (2014), even excluded typical features of game or gamification such as points and reward structures. The involvement of varied game design elements also makes it confusing about how gamification works, such as which element is engaging, to what extend it is engaging, and a specific game design element will be engaging in what specific contexts.

What Werbach (2014) proposed reflects three main concepts: the definition of gamification needs further investigation, the existing definitions of gamification did not highlight gamefulness, and the use of varied game elements brings challenges to the comparison of the effectiveness of gamification practices.

Werbach (2014) furtherly emphasised the importance of redefining the gamification concept. Werbach (2014) stressed that gamification is still a young field, so the definition affects the coherence of research efforts (Werbach, 2014). The key for redefining gamification is not whether it is "right" in an abstract sense, but "whether the distinctions and boundaries it creates are useful" (Werbach, 2014, p. 269).

As detailed above, it is essential to redefine gamification in new contexts, which involves two main concerns: whether the gamification is the meaningful gamification which promotes gamefulness, and how to categorise various game elements into certain groups for enhanced comparison feasibility (Werbach, 2014).

The explicit-and-implicit gamification concept provided by Chou (2015) addresses these two concerns. According to Chou (2015), explicit gamification involves strategies utilising applications that are obviously game-like, while implicit gamification refers to human-focused design that utilises game elements in non-game contexts. Chou (2015) highlighted the importance of implicit gamification, saying that implicit gamification focuses on individuals' natural desires, provides motivational drives, and eventually results in human behaviours.

However, though Chou (2015) clearly concluded that explicit gamification involves "obviously game-like strategies" while implicit gamification involves "human-focused design that involves game elements", the specific indication of "strategies" or "game elements" remain obscure. This author extended the explicit-and-implicit gamification concept to specific definitions, as shown in the following table (Table 3).

Table 3. Redefining gamification

Term	Definition
Explicit gamification	The strategies that utilise obviously game-like applications. (Chou, 2015)
Implicit gamification	The human-focused designs that utilises game elements in non-game contexts (Chou, 2015).
Gamified learning	The use of game-like elements or human-focused designs to promote human innate psychological needs, which eventually enhances learner engagement or performances.

2.1.4. Redefining game elements and gamification mechanisms

> There is "no universal list of game elements" (Werbach, 2014, p.267).

This section aims to separate the concept "gamification mechanisms" from "game elements".

Werbach (2014) pointed out that though gamification was defined as the use of "game design elements" in non-game contexts (Deterding, Dixon, et al., 2011, p. 10), there was a lack of uniformed understanding of available elements: there is "no universal list of game elements" (Werbach, 2014, p. 267).

The lack of universal list of game elements bought high variance of the effectiveness results presented in previous studies: different studies may use *different* game elements with *different* definitions in *different* educational contexts on *different* target group of people, so the effectiveness results are not suitable to be compared; or, the process will be like comparing bananas with apples. The use of different game elements explains why previous studies reported significantly different conclusion regarding to the effectiveness of gamification.

Game elements listed in previous studies

Researchers attempted to secure a list of game elements for a specific context, either from literature or from consulting participants.

Reeves and Read (2009) suggested that there are ten ingredients of great games via cooperating with gamers and game designers: 1) self-representation with avatars, 2) three-dimensional environments, 3) narrative context, 4) feedback, 5) reputations, ranks, and levels, 6) marketplaces and economies, 7) rules that are explicit and enforced, 8) teams, 9) communication systems that can be reconfigured by participants, and 10) time pressure (pp. 231-236). Toda et al. (2019) criticised that previous game-element frameworks, the one by Reeves and Read (2009) included, did not consider specifications from educational domains, or tied to particular contexts.

Accordingly, Toda et al. (2019) stressed the need to list the game elements apply in the educational domain. By reviewing literature and consulting experts, the authors nominated 21 game elements, which include acknowledgement, chance, competition, cooperation, economy, imposed choice, level, narrative, novelty, objectives, point, progression, puzzles, rarity, renovation, reputation, sensation, social pressure, stats, storytelling, and time pressure. The expert investigation results suggested that some elements were considered crucial, such as objective, level, and progression, and some others are not (e.g. narrative and storytelling).

However, the 21 game elements were not divided into sub-categories, which should be explored in future studies. Most importantly, due to methodological limitations, such as the variance of involved articles and the small number of experts, the nominated 21 game elements "cannot (be) generalise(d)… to all contexts" (Toda et al., 2019, p. 88). More participants are expected to be involved for enhanced reliability, which means the nominated game elements might also change.

The nomination of game elements by other researchers also presented a tendency of being casual. Chapman and Rich (2018) attempted to investigate which game element work best on improving students' motivation by surveying students' perceptions. The nominated game elements include non-typical ones such as "doing peer reviews", "doing assignment" and "starting with zero points" (p. 317). It is true that game elements is "a vast, seemingly never-ending list of concepts" (Cavanagh, 2019, p. 22), so the non-typical ones can also be considered. However, the nomination

of game elements with high variance brings challenges to the comparison of research findings, which consequently weakens the theory generalisability and research value.

This author proposes that reviewing or listing game elements brings limited academic value. However, it is essential to categorise the game elements into groups. The categorisation enables researchers to end the effort-demanding process of analysing the never-ending list of concepts, as well as allowing researchers to freely add new game elements to the list based on the category characteristics.

Previous studies also attempted to divide the never-ending game-element concepts into groups.

Seaborn and Fels (2015) suggested to put the game elements into three groups based on the three-factor self-determination theory (autonomy, competence and relation). For example, the authors suggested that several game elements contribute to the establishment of learning autonomy, such as avatars, privacy control, notification control and alternative activates.

Groh (2012) proposed that "elements of games can be found outside games" (p.40), which means in gamification the "game design elements" proposed by Deterding, Sicart, et al. (2011) should not only be the game elements defined from technical aspects. Groh (2012) categorised the game elements into five categories: game interface design patterns, game design patterns and mechanics, game design principles and heuristics, game models, and game design methods, as shown in Table 4.

Table 4. Levels of game design elements (Groh, 2012, p. 40)

Category	Description	Example of the element
Game interface design patterns	Common, successful interaction design components and design solutions for a known problem in a context, including prototypical implementations	Badge, leaderboard, level
Game design patterns and mechanics	Commonly reoccurring parts of the design of a game that concern gameplay	Time constraint, limited resources, turns
Game design principles and heuristics	Evaluative guidelines to approach a design problem or analyse a given design solution	Enduring play, clear goals, variety of game styles
Game models	Conceptual models of the components of games or game experience	challenge, fantasy, curiosity; game design atoms
Game design methods	Game design-specific practices and process	Playtesting, value conscious game design

Bunchball (2010) attempted to divide game elements into game mechanics and game dynamics, which is adopted from the Mechanics-Dynamics-Aesthetics (MDA) framework (Hunicke, LeBlanc, & Zubek, 2004).

Game mechanics refer to the "various actions, behaviours, and control mechanisms that are used to 'gamify' an activity" to create a compelling and engaging user experience (Bunchball, 2010, p. 2) The desires and motivations results in the compelling and motivational nature of the user experience are called game dynamics (Bunchball, 2010). As shown in Table 5, Bunchball (2010) proposed that game mechanics include points, levels, challenges, virtual goods and spaces, leaderboards, and gifts/charity.

Game dynamics refer to the desires and motivations results in the compelling and motivational nature of the user experience are called game dynamics. Bunchball (2010) proposed that six main game elements are constitutes game dynamics, including reward, status, achievement, self-express, completion, and altruism.

Table 5. Game mechanics and game dynamics (Bunchball, 2010)

Term	Definition	Elements
Game mechanics	"Various actions, behaviours, and control mechanisms that are used to 'gamify' an activity" to create a compelling and engaging user experience (Bunchball, 2010, p. 2).	Points Levels Challenges Virtual goods and spaces Leaderboards Gifts and charity
Game dynamics	The desires and motivations results in the compelling and motivational nature of the user experience are called game dynamics.	Reward Status Achievement Self-express Competition Altruism

Blohm and Leimeister (2013) made the same categorisation though with different game element list. According to Blohm and Leimeister (2013), the elements constitude game mechanics are: documentation of behaviour, scoring systems/badges/trophies, rankings, ranks/levels/reputation points, group tasks, time pressure/tasks/quests, and avatars/virtual words/virtual trade. Regarding game dynamics, the related game elements include exploration, collection, competition, acquisition of status, collaboration, challenge, and development/organization.

Blohm and Leimeister (2013) added one more category: motives. According to Blohm and Leimeister (2013), gamification is an incentive mechanism that motivates users for the expected behaviours. In another word, users are engaged by gamification only merely because of the game elements, but because their motives are met with the use of game elements. The motives include intellectual curiosity, achievement, social recognition, social exchange, cognitive stimulation, and self-determination (see Table 6).

Notes

This author proposes that reviewing or listing game elements brings limited academic value. However, it is essential to categorise the game elements into groups.

Benefits:

1. The categorisation enables researchers to end the effort-demanding process of analysing the never-ending list of concepts;

2. The categorisation enables researchers to freely add new game elements to the list based on the category characteristics.

Table 6. Game-design elements and motives (Blohm & Leimeister, 2013, p. 276)

Category	Element
Game design element: mechanics	Documentation of behaviour, scoring systems/badges/trophies, rankings, ranks/levels/reputation points, group tasks, time pressure/tasks/quests, and avatars/virtual words/virtual trade
Game design element: dynamics	Exploration, collection. Competition, acquisition of status, collaboration, challenge, and development/organization
Motives	Intellectual curiosity, achievement, social recognition, social exchange, cognitive stimulation, and self-determination

Some researchers did not consider categorising the game elements. Buckley and Doyle (2017) specified the elements as ”elements of gamification” (p.45), even though some of those elements are with obvious-game-like features. The listed elements of gamification include achievements, avatars, badges, boss fights, collections, combat, content unlocking, gifting, leaderboards, levels, points, quests, social graphs, teams, and virtual goods. When describing the game elements in the learning management system Moodle, Vanduhe, Nat, and Hasan (2020) identified eight game elements: leaderboard, badge, progress bar, points, level, chat, unlocking content, and completion tracking, among which each one is connected with one correspondent game dynamic. The eight game dynamics include completion, achievement, ranking, rewards, status, social, trigger, and measurement or monitoring.

Game elements categorised by this author

Based on the implicit-and-explicit gamification theory, this author proposes to use the term “game elements” to refer to the obviously game-like strategies that construct explicit gamification, and the term “gamification mechanisms” to refer to the human-focused designs that help realise implicit gamification. To be more specific, this author proposes to distinguish the concept game mechanisms from game elements: game elements are the obvious game-like elements that are frequently used in typical digital or board games, which are normally concrete nouns (e.g., points, badges, avatar, virtual goods), while game mechanisms are the underlying guidelines that make an activity engaging, which are abstract nouns that relate to innate psychological needs (e.g., challenge, feedback, achievement, competition, fantasy) (see Table 7).

Correspondingly, this author defined gamified learning or gamification in education as the use of game-like elements and game mechanisms to promote human innate psychological needs, which eventually enhances learner engagement or performances (see Table 7).

The term gamification mechanisms appeared in previous studies; however, the explanations were highly inconsistent. For example, Schacht and Schacht (2012) described gamification mechanisms as the summarised name of a cluster including “achievements, bonuses, levelling up, progression, leaderboards, cascading information and virality” and “additional” ones when needed (p. 195). From this perspective, gamification mechanisms are similar to game elements. Arakawa and Matsuda (2016) described gamification mechanism as the providence of tasks, rewards and communication which promotes value, ideology and fun. According to Arakawa and Matsuda (2016), gamification mechanism is a process rather than a concretised item. What’s worse, to my knowledge, no researchers secured the definition of gamification mechanism yet.

To minimise confusion, this author redefined the term gamification mechanisms after analysing the selected articles (see Table 7). Terminology differences are illustrated in Figure 2, while examples of game elements and gamification mechanisms are detailed in Table 8.

2.1.5. Redefinition of gamified learning tools

Besides gamification, Findings revealed that *there is a need to define the concept "gamified learning tools".* previous studies revealed that numerous tools can be added with gamified technique, such as classroom response system, interactive whiteboard, robots, MOOC (Massive Open Online Courses) learning system, virtual reality (VR), augmented reality (AR), or even non-technology activity designs. Undoubtedly, the involvement of different tools is accompanied by varied entry barriers, design principles, implementation difficulties, educational implications, and research conclusions. Therefore, comparing the effectiveness of studies using different gamified learning tools is like comparing bananas with apples. To address this issue, there should be a specific definition of "gamified learning tools".

Varied researchers contributed to the redefinition of gamified learning tools. At the very first beginning, researchers attempted to use full-fledged commercial games for educational purposes, which have been criticised for various reasons, such as containing unnecessary content and bringing distractive information (Kirriemuir & McFarlane, 2004). In this context, the learning tools were mainstream video games or commercial video games. Then, Kirriemuir and McFarlane (2004) proposed to use the "lite" version of mainstream video games, which exclude unnecessary or irrelevant content as well as including curriculum-relevant tasks (p. 26). However, the "lite" mainstream video games need researchers to request the code of the games from the game developers and make revisions, which sets a high barrier for researchers or educational practitioners; other researchers attempted to design customised tools for gamification activities. However, researchers reflected that the customised tools might not be engaging enough due to design limitations.

Shang, Li, and Li (2005) reshaped the lite game concept from pedagogical perspectives. According to Shang et al. (2005), a "lite game" for educational purposes should be an educational website or software that used the motivating components of video games. Focusing on educational website or software, the lite game concept proposed by Shang et al. (2005) distinguished the gamification designs (e.g., fantasy, challenge, competition, curiosity) from obvious game-like elements (e.g. boss fight, avatar, sound effects). Though the lite game concept defined by Shang et al. (2005) was much earlier than the well-acknowledged gamification definition by Deterding, Dixon, et al. (2011), the core ideas were overlapping. Therefore, this author supports defining the gamified learning tool as "educational website or software that use the motivating components of video games to engage learners or improve performance", as initialled by Shang et al. (2005).

It was highlighted that gamification is still a young field, so the definition affects the coherence of research efforts (Werbach, 2014). The key for redefining gamification is not whether it is "right" in an abstract sense, but "whether the distinctions and boundaries it creates are useful" (Werbach, 2014, p. 269). The current definition explicitly

It is notable that gamified learning does not necessarily involve digital devices even if people tend to make involuntary associations. Non-digital gamification is also of great potential in engaging students for enhanced performances; however, since a successful non-digital gamified learning practice is comparatively more difficult to design, record, spread and reproduce, together with the fact that the front-line teachers implementing non-digital gamified learning are not necessarily researchers writing academic papers, non-digital gamified learning has been rarely discussed in academic publications. In most cases, gamified learning tends to refer to digital gamified learning; therefore, gamified learning tool in this paper refers to a website, a system or mobile software that employs gamification techniques, as shown in Table 7. The specific examples are shown in Table 8.

Table 7. Redefining game elements and gamification mechanisms

Feature	Term	Definition
Game-like	Game-based learning (GBL)	The use of fully-developed video games for educational purposes.
	Explicit gamification	The strategies that utilise obviously game-like applications. (Chou, 2015)
	Game elements	The obvious game-like elements that are frequently used in digital games or gamification, which are concrete nouns.
Human-desire based	Gamified learning (Gamification in education)	The use of game-like elements or gamification mechanisms to promote human innate psychological needs, which eventually enhances learner engagement or performances.
	Implicit gamification	The human-focused design that utilises game elements in non-game contexts (Chou, 2015).
	Gamification mechanisms	The underlying guidelines that make gamification activities engaging, which are abstract nouns that relate to humans' innate psychological needs.
Other	Gamified learning tool	The educational website, software or mobile app that uses game-like elements and gamification mechanisms to promote human innate psychological needs, which aims to enhance learner engagement or performances.

Table 8. Game elements and gamification mechanisms: definition, feature, and examples

Term	Definition	Feature	Examples
Game elements	The obvious game-like elements that are frequently used in digital games or gamification.	Concrete nouns that are obvious game-like	Points, badges, leaderboard, trophies, levels, tasks, scoring systems, avatar, boss fight, virtual goods, virtual trade, alternative activities, group, etc.
Gamification mechanisms	The underlying guidelines that make gamification activities engaging, which originate from human needs and desires.	Abstract nouns that relate to human needs or desires	Feedback, achievement (accomplishment), meaning, self-expression, competition, curiosity, exploration, collaboration, fantasy, challenge, self-determination, social influence (social recognition), unpredictability (uncertainty), fun (fun failure), avoidance, scarcity, ownership, engagement, enjoyment, interaction, user control, etc.

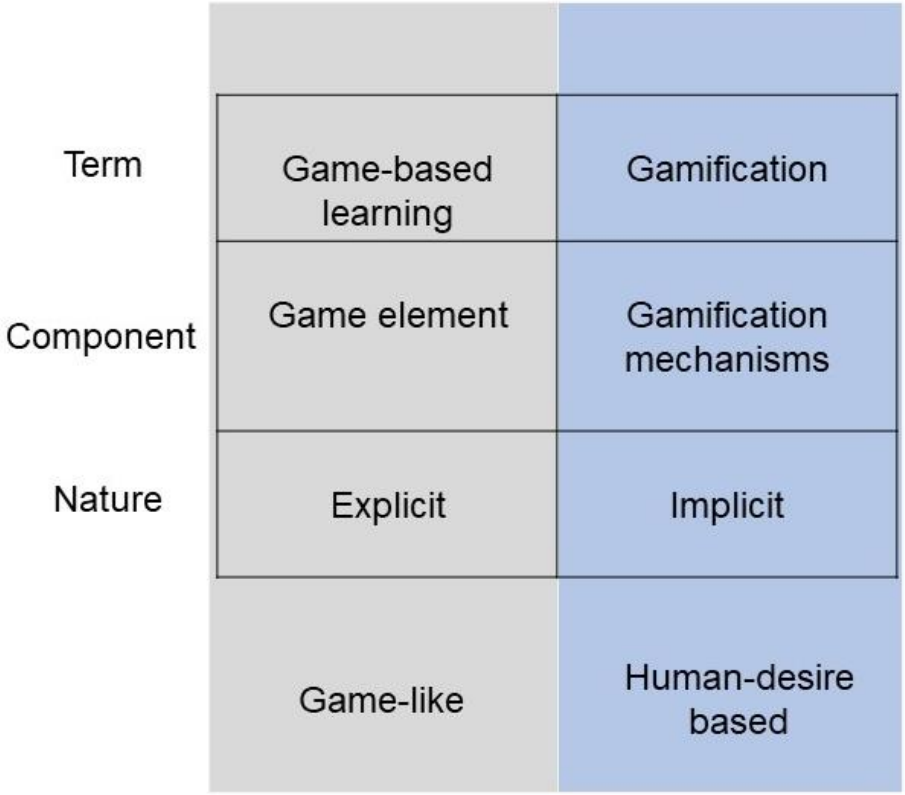

Figure 2. Terminology differences

Notes

Definition of gamified learning tools:

The educational website, software or mobile app that uses game-like elements and gamification mechanisms to promote human innate psychological needs, which aims to enhance learner engagement or performances.

2.2. Gamification-related frameworks and theories

Gamification is an inter-disciplinary field that involves game design principles. Considering the research scope, the frameworks and theories for game design were excluded, though some of them contribute to the design of gamification tools. Frequently-mentioned frameworks or theories in gamification studies include flow theory, the self-determination theory (SDT), and the mechanics-dynamics-aesthetics (MDA) framework.

2.2.1. Flow theory

Flow theory, proposed by Csikszentmihalyi (1990), is a popular theory in gamification studies, which originated from the field of Psychology. Flow is summarised by several researchers as "the state in which we are so involved in something that nothing else matters", which "has clear relevance to research into games (and play)" (Kirriemuir & McFarlane, 2004, p. 9). According to Csikszentmihalyi (1990) to create the optimal experience, either for gameplay or for other activities, there should be a good balance between the perceived challenge of the task and the users' skills. As illustrated in Figure 3, when the task is too challenging, anxiety would be raised; if the task is too simple, the users would experience boredom; only when the challenge is balanced with user skill, the users would reach the flow status. At the same time, the activities should provide explicit and reachable task goals while providing instant feedback to the users.

There are also other features of the flow status. As summarised by Hamari and Koivisto (2014), Csikszentmihalyi (1990) defined nine characteristics of the flow status: "(1) a balance between the challenge of the task and skills of the individual, (2) a merging of action and awareness, i.e. one performs the activity almost "automatically", (3) clear perceived goals, (4) unambiguous feedback, (5) focusing on the task at hand, (6) a sense of control of the activity, (7) a loss of self-consciousness or a reduced awareness of self, (8) time transformation, i.e. sense of time becomes distorted, and (9) an autotelic, intrinsically rewarding experience, implying that the activity in itself is a reason for performing it instead of any external objectives" (p. 134). One of the most eye-catching features of games for educational purposes is their potential in promoting the flow status.

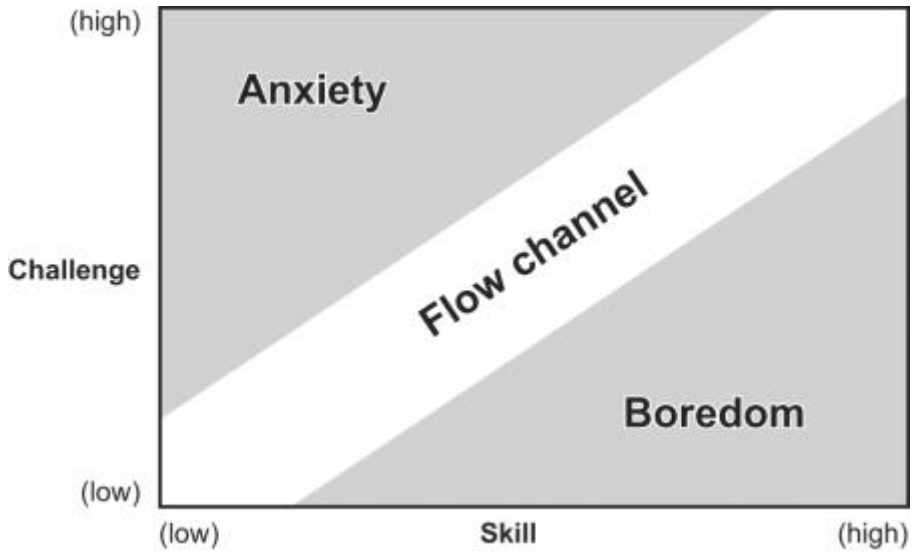

Figure 3. Flow channel (Csikszentmihalyi, 1990, p. 74)

2.2.2. Self-determination theory (SDT)

Seaborn and Fels (2015) proposed that there are several key theoretical foundations used in gamification frameworks, among which one main is the self-determination theory (SDT) (Ryan & Deci, 2000). SDT emphasises the importance of "human inner competence" on personality development and self-control of behaviours (Liu, Wang, Huang, & Tang, 2019, p. 270), which also explains human psychological needs as the foundation of self-motivation.

According to SDT, there are three main factors that drive people to engage in tasks, which are competence, autonomy, and relatedness (Ryan & Deci, 2000). Competence refers to people's subjective confidence about their capability to overcome tasks and challenges; autonomy represents the awareness of using actions to demonstrate personal interests and integrated values, and relatedness involves meaningful connections and relationships with others (Ryan & Deci, 2000).

The self-determination theory also applies to gamification contexts. Roy and Zaman (2017) specified how SDT and gamification relate by listing gamification designs that achieve autonomy, competence, and relatedness respectively, as summarised in Table 9.

Gamification designs and autonomy. Roy and Zaman (2017) proposed that when a gamified system provides a variety of meaningful challenges, the system is likely to trigger autonomy. The challenges cannot be obligatory, to avoid giving users the feeling of being control*led*. Besides, previous studies have confirmed that a moderate amount of choices is "likely to incite the perception of being autonomously motivated" (Roy & Zaman, 2017, p. 497), so it is suggested to find the "sweet spot" between providing learners with at least one meaningful option and placing students in a dilemma by offering too many options.

Gamification designs and competence. The need for competence refers to the desire of feeling the possibility of successfully achieve a goal (Deci & Ryan, 2000). Roy and Zaman (2017) suggested that in educational contexts, learners who are competent tend to be more persist than those who are incompetent. Nevertheless, enabling learners to be competent does not mean to make the educational activities as simple as possible; align with the flow theory, it is suggested to create tasks that post a significant challenge while remaining perceived as feasible to fulfil (Roy & Zaman, 2017).

According to Roy and Zaman (2017), another way of fostering feelings of competence is providing constructive and meaningful feedback. Gamified applications (or gamified learning tools) provide feedback by presenting game design elements such as badges; however, not all types of feedback would positively work on learning motivation: if the feedback is given according to behavioural performance rather than learner competence, the feedback is likely to be perceived as controlling, which therefore undermines autonomous motivation. Accordingly, the authors stressed that the provided feedback should be positive and competence-related.

Table 9. SDT and gamification designs

#	Components in SDT	Gamification design
1	Autonomy	Avoid obligatory uses.
2	Autonomy	Provide a moderate amount of meaningful options.
3	Competence	Set challenging but manageable goals.
4	Competence	Provide positive, competence-related feedback.
5	Relatedness	Facilitate social interaction

Gamification designs and relatedness. Roy and Zaman (2017) suggested that relatedness in gamification designs should be connected with social interactions. Roy and Zaman (2017) believed that in educational contexts, learners who work together with others have stronger satisfaction and autonomous motivation. Consequently, Roy and Zaman (2017) suggested minimising factors that hinder social interactions between users, and facilitate them to interact and support their feelings of relatedness instead.

2.2.3. The mechanics-dynamics-aesthetics (MDA) framework

The mechanics-dynamics-aesthetics (MDA) framework proposed by Hunicke et al. (2004) is based on game design principles, which aims to build connections between the base mechanisms of game systems, the overarching design goals, and the desired experiential results of gameplay. The MDA framework breaks down a player's consumption process of games into three parts: *rules, system,* and *fun*, followed by three design counterparts *mechanics, dynamics,* and *aesthetics*.

Mechanics are the "various actions, behaviours and control mechanisms afforded to the player" within a game context (Hunicke et al., 2004, p. 3), which describe "the particular components of the game, at the level of data representation and algorithms (p. 2).

Dynamics "describes the run-time behavior of the mechanics acting on player inputs and each other's outputs over time", and aesthetics describes "the desirable emotional responses evoked in the player, when she interacts with the game system" (Hunicke et al., 2004, p. 2).

Aesthetics "describes the desirable emotional responses evoked in the player, when she interacts with the game system" (Hunicke et al., 2004, p. 2). There are many types of aesthetics, including but not limited to: sensation, fantasy, narrative, challenge, fellowship, discovery, expression, and submission (Hunicke et al., 2004, p. 2).

The MDA framework is a popular one in game design. However, the involved concepts are not familiar for researchers in fields other than game design.

2.2.4. A Taxonomy of Motivational Affordances

Though the concepts in MDA framework were widely used in game-design-related studies, this author found the MDA framework complex to understand. For example, in the study of Weiser, Bucher, Cellina, and Luca (2015), game mechanics include feedback, education, challenges, rewards, competition, and cooperation. The same with what was suggested by Werbach (2014) about game elements, the list of game mechanics is also never-ending, while the concept of game mechanics remains mixed.

The taxonomy of motivational affordances proposed by Weiser et al. (2015) helps clarify the term mechanics. Motivational affordances refer to the properties that "afford user motivation" (Weiser et al., 2015, p. 274). According to Weiser et al. (2015), there are three aspects of considerations in designing a gamified tool: game elements, game mechanics, and general design principles (see Figure 4).

Game elements refer to the building blocks to implement mechanics, which tend to be concrete objects such as assignments, points, levels, badges, friends, and reminders; mechanics are the "possible means of interaction between user and system", such as feedback, rewards, competition, challenges, and cooperation (Weiser et al., 2015, p. 274); general design principles are the abstract guidelines for process design, such as personalise the experience and offer meaningful suggestions.

Weiser et al. (2015) suggested that the game elements help implement mechanics, and the combination of game elements and game mechanics helps achieve the general design principles. Namely, "mechanics can be implemented using one or several concrete elements" (p.274), and the use of mechanics helps what the general design principles target to achieve. Gamification designers can choose components depending on the objectives and needs of the users.

The three components of the taxonomy of motivational affordances provided a valuable cut point to study gamification: dividing concrete nouns and abstract nouns of the "game elements". In the framework, general design principles are words in the form of abstract nouns (e.g., fantasy, challenge) while the elements are words in the form of concrete nouns (e.g., points, badges, leaderboard). The taxonomy enables programmers to quickly make connections between the handy game elements with the to-achieve expectations. The re-definition of game elements and the correspondent components are two main contributions of the current thesis, which are detailed in Table 7 and Table 8.

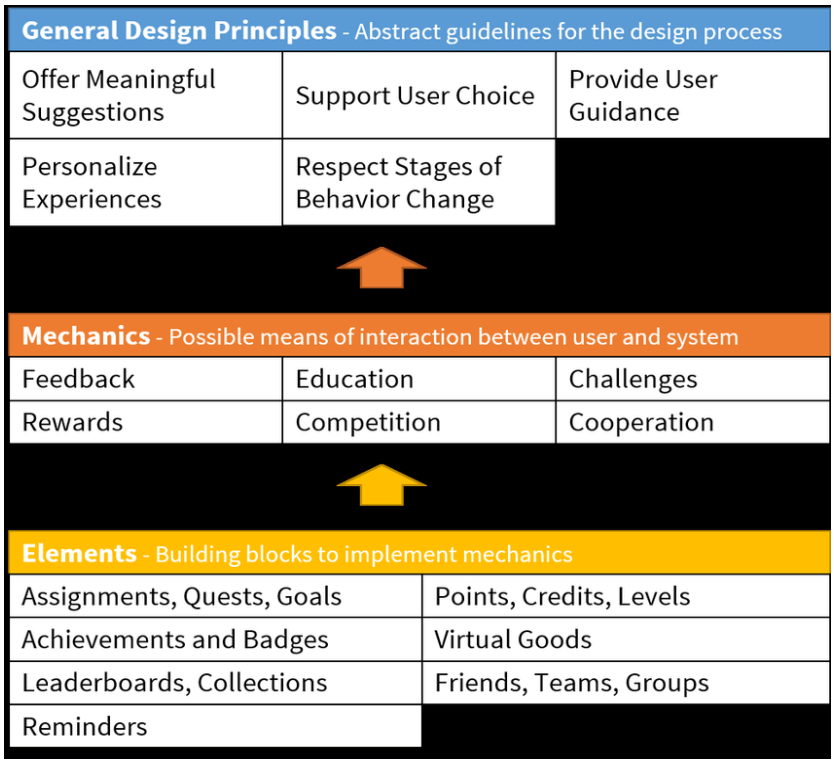

Figure 4. Taxonomy of motivational affordances (Weiser et al., 2015, p. 274)

2.3. Teacher perception

2.3.1. The importance of investigating teacher perception

How and why individuals adopt a particular technology has motivated a great deal of research, especially for gamification (Straub, 2009).

Cuban (1986) proposed that many top-down attempts to introduce innovative technologies in education have failed to achieve long-term effects. Reasons could be the lack of awareness of teacher attitude towards the technology, which plays a crucial role in influencing teachers' selection, implementation and evaluation of pedagogical interventions in the classroom (Ale & Chib, 2011; Huizenga et al., 2017; José Martí-Parreño et al., 2016). If teachers do not see the value of a technology or approach, their technology acceptance intention would be negatively influenced (Huizenga et al., 2017).

Despite the increasing popularity of gamified learning, teachers as the key role in implementing it in schools have been "neglected in gamification research with few research focused on them" (José Martí-Parreño et al., 2016, p. 682). Teachers play a critical role in introducing pedagogical innovations in the classroom, especially when the innovations are related to new technologies (José Martí-Parreño et al., 2016). To be more specific, the effectiveness of game-based learning or gamified learning "largely depend on" teachers' acceptance (Bourgonjon et al., 2013, p. 22) since the new form of technology-driven innovations can challenge teachers' attitudes and provoke anxiety of uncertain outcomes(Sánchez-Mena & Martí-Parreño, 2017), which therefore lead to negative attitudes toward gamified learning, and eventually impede the realization of gamification' full potentials in classrooms (Sugar et al., 2004; Tuparova et al., 2018).

Albirini (2006) pointed out the reason why teachers' perceptions have been neglected in gamification studies. According to Albirini (2006), a large number of new technologies focus too much on the technical aspects and the effect on students' academic achievement, with the role of teachers neglected. However, "teachers are in many areas the true change agents of schools in terms of modes of education" (Bourgonjon et al., 2013, p. 22).

To sum up, the adoption and the usefulness of an instructional method may be greatly affected by teachers' perceptions (Vandercruysse et al., 2013), therefore understanding "teachers' perceptions and beliefs that underlie their decision-making processes" (Bourgonjon et al., 2013, p. 21) is of great significance. However, studies on users' perceptions towards gamified learning tools are scarce (teachers as users are also included) (Rapp, 2015; Van Rosmalen et al., 2015).

Since gamification is a concept that is closely related with digital games, teachers are possible to perceive gamified learning as the use of digital games in the classroom. A number of studies investigated how teachers perceive game-based learning from various aspects. For example, Huizenga et al. (2017) propose that "teachers' perceptions of the usefulness of digital games might be a reason for the limited application of digital games in education" (p. 105), so they conducted a semi-structured interview to explore how 43 secondary school teachers view the values and barriers of implementing game-based learning in education. The study of Noraddin and Kian (2015) specified digital games as "motivational, collaborative and instructional tools" (p. 143), and collected questionnaires from 273 university teachers. Other studies majorly focused on teachers' intention to adopt digital games in formal education (De Grove, Bourgonjon, & Van Looy, 2012; Koh, Kin, Wadhwa, & Lim, 2012; Lijuan Wang & Ha, 2009).

Only a limited number of studies chosen gamified learning rather than digital games as the research objectives and the majority of them focused on factors affecting teachers' intention to implement gamification in classes. Sánchez-Mena and Martí-Parreño (2017) conducted interviews to investigate the drivers and barriers for teachers to implement

gamification methods in their classes. The result reported that the top four reasons driving teachers to use gamification are its functions on attention-motivation, entertainment, interactivity, and easiness to learn. As to barriers, the top four frequently-mentioned elements are the lack of resources, students' apathy, subject differences, and classroom dynamics. Adukaite, van Zyl, Er, and Cantoni (2017) studied whether six factors are affecting the adoption intention of gamification in learning contexts, and the six factors are perceptions about playfulness, curriculum fit, learning opportunities, challenge, self-efficacy and computer anxiety. There are also other studies investigating how different users view gamification in education. Ziesemer, Müller, and Silveira (2013) focused on game elements and concluded that "users are not aware of some game elements and they have distinct motivation and knowledge about gamification". The study of Denden, Tlili, Essalmi, and Jemni (2017) furtherly extended the research since it considered individual differences and the effects on users' perception of gamification and game elements. The results showed "gender and gaming frequency does not affect the perception of gamification but they do affect the perception of using game elements" (Denden et al., 2017, p. 1).

To sum up, gamification as a novel concept is understudied on how it was perceived by teachers. The study of how teachers review gamified learning apps on the perceived usefulness and perceived ease of use is even scarce.

2.3.2. Advantages of GBL and gamified learning: Teacher perception

Top keywords were clustered in five themes that emerged as the main drivers for teachers' use of gamification in schools: engagement and motivation, risk-free environment, visualised feedback, learning autonomy, and inclusivity (see Table 10).

The first and foremost theme is the potential to engage learners, which triggers intrinsic or extrinsic motivation (Emel'yanenko, Vetoshko, Malinnikov, Malashenko, & Vetoshko, 2016; Huizenga et al., 2017; José Martí-Parreño et al., 2016; Sánchez-Mena & Martí-Parreño, 2017). Gamification that employed game elements is reported to be attention-catching, entertaining, interactive and fun (Sánchez-Mena & Martí-Parreño, 2017). Wingo et al. (2019) conducted two focused group studies involving 13 students to investigate their perception of a gamification platform Kaizen. The Kaizen system provides multiple-choice questions in a game format, encourage competition among individuals and teams, with results displayed in leaderboards and badges. The results show that the students like the extrinsic rewards such as points, badges and improved ranking, as well as the intrinsic rewards such as personal challenges and educational enhancement.

Gamified learning provides a risk-free way to learn, which free students from the worries of making mistakes (Emel'yanenko et al., 2016; José Martí-Parreño et al., 2016; Sánchez-Mena & Martí-Parreño, 2017; Wingo et al., 2019). According to Huizenga et al. (2017), games enable learners to try things in simulated situations without being exposed to real dangers. The concept of safety can be a broad one, such as learners being able to learn from mistakes without threats. The same theory applies to gamification as well: gamification provides students with the freedom to fail without fears by providing a trial-and-error learning process (Emel'yanenko et al., 2016; Lee & Hammer, 2011).

GBL or gamified learning also provides immediate feedback in a visualised way with the use of game elements such as points, badges, leaderboards and sound effects (Huizenga et al., 2017; Kapp, 2012; Koh et al., 2012; J Martí-Parreño, Méndez-Ibáñez, & Alonso-Arroyo, 2016; Sánchez-Mena & Martí-Parreño, 2017). It enables the students to experience the consequence of their choices.

GBL or gamified learning also enables students to take control of their own learning, which furtherly triggers learning autonomy. It enables the students to master their learning pace by providing student-centred ad adaptive learning materials (Hanus & Fox, 2015; Wingo et al., 2019). The findings were consistent with constructivist and adult learning principles (Kirkley & Kirkley, 2005) in that students wanted more control and tended to describe the game

in terms of how it enhanced their learning (Wingo et al., 2019). Accordingly, students who are not typically active in academic activities are provided with a chance to demonstrate how active they could be (Emel'yanenko et al., 2016).

Inclusivity is the next theme reported as the advantages of GBL or gamified learning in previous studies. The study of Wang and Ha (2009) reported that GBL helps deal with students' individual differences effectively by including students with different skill levels. Some of the game-based or gamified learning tools also allow teachers to modify the tool to include more students (Wang & Ha, 2009). By customizing the learning content based on each individual student's needs, GBL or gamified learning includes more students in the learning process (Hanus & Fox, 2015; Marino, Israel, Beecher, & Basham, 2013; J Martí-Parreño et al., 2016).

At the same time, game-based learning or gamified learning also provides social interaction (Wingo et al., 2019), more learning opportunities in class and in after school hours (Huizenga et al., 2017; Ismail & Ibrahim, 2018), natural attractiveness to students of the new generations that have grown in the age of video games (José Martí-Parreño et al., 2016; Sánchez-Mena & Martí-Parreño, 2017). Though not common, previous findings revealed other advantages of GBL or gamified learning; for example, according to Sánchez-Mena and Martí-Parreño (2017), the games for GBL are easy to use for students.

Table 10. Advantages of GBL and gamified learning

#	Theme	Reference	
		GBL	Gamification
1	Bring engagement and motivation	(Huizenga et al., 2017)	(Emel'yanenko et al., 2016; José Martí-Parreño et al., 2016; Sánchez-Mena & Martí-Parreño, 2017)
2	Risk-free environment	(Emel'yanenko et al., 2016; Huizenga et al., 2017; Sánchez-Mena & Martí-Parreño, 2017)	(José Martí-Parreño et al., 2016; Sánchez-Mena & Martí-Parreño, 2017; Wingo et al., 2019)
3	Visualised feedback	(Huizenga et al., 2017; Koh et al., 2012; Sánchez-Mena & Martí-Parreño, 2017)	(Kapp, 2012; J Martí-Parreño et al., 2016; Sánchez-Mena & Martí-Parreño, 2017)
4	Learning autonomy	(Hanus & Fox, 2015; Marino et al., 2013; Wingo et al., 2019)	
5	Inclusivity	(Emel'yanenko et al., 2016; Hanus & Fox, 2015; Sánchez-Mena & Martí-Parreño, 2017; Wang & Ha, 2009)	(Hanus & Fox, 2015; José Martí-Parreño et al., 2016; Sánchez-Mena & Martí-Parreño, 2017)

2.3.3. Factors contributing to teachers' acceptance intention

Notably, the advantages of GBL or gamification are not necessarily the factors contributing to teachers' acceptance intention. Previous empirical studies indicate that when considering implementing GBL or gamified learning, teachers are influenced by various factors, which could be categorised into three domains: teacher, technology and context (H.-K. Wu, Hsu, & Hwang, 2008). Similarly, Berg Marklund (2015) proposed that three aspects of educational settings should be taken into consideration during educational game integration and use: human factors, technological factors, and organisational structures and praxis.

Considering the current context, the three domains were paraphrased as teacher factors, tool factors, and external support. The specific factors of each theme are shown below and in Table 11.

Teacher factors

Negative attitude. Even though the construct 'attitude' is in the Technology Acceptance Model (TAM) (Davis, Bagozzi, & Warshaw, 1989), the attitude does not indicate positive or negative, while the two types of attitude impact differently to users' final acceptance intention. Discussions in previous studies focused on the negative attitude that contributes to teachers' acceptance intention to GBL or gamified learning, which consists of two main themes: the negative attitude caused by the scepticism of the function and the negative attitude caused by the challenged teachers' roles.

Firstly, digital games that entail playfulness is often considered an oxymoron with learning (De Grove et al., 2012). Using digital games in the classroom requires a mindset that is independent of the traditional beliefs, assumptions and values about teaching and learning (Koh et al., 2012); however, researchers reported that digital games in a classroom are often "looked upon with scepticism by teachers" (De Grove et al., 2012, p. 2025). According to Demirbilek and Tamer (2010), some of the teachers indicated that "computer games should not be fun-oriented" (p. 713), or the students will focus on game playing rather than learning. The overall "gaming mindset", which refers to the teacher's experience and attitude toward games, is hindering teachers' acceptance intention to GBL (Koh et al., 2012, p. 59). As noted by Koh et al. (2012), "it seems that playing is akin to doing something wrong", therefore "teachers generally perceived gaming as a good form of learning for students" (p. 59). However, it is notable that the aforementioned issue is reported in GBL practices only rather than gamified learning activities.

The second theme is the challenged teachers' roles. GBL tools and gamified learning tools are providing students with an option of learning that enables students to do without a teacher in the educational process (Emel'yanenko et al., 2016), accordingly, some authors believe that gamification changes the role of the teacher from leaders to partners, which weakens teachers' leading role (Emel'yanenko et al., 2016; Tuparova et al., 2018). Meanwhile, the use of technological innovations can be a cause of anxiety, as well as GBL or gamified learning (Sánchez-Mena & Martí-Parreño, 2017).

Teachers' skills. Several studies have pointed out the importance of skill as an acceptance intention determinant (Bourgonjon et al., 2013; De Grove et al., 2012; Sánchez-Mena & Martí-Parreño, 2017; Tuparova et al., 2018). The skills can be interpreted twofold: the skills to solve technical problems and the skills to use the GBL or gamification tools to achieve the learning goals. While the former can be easily addressed by providing technique supports and information technology (IT) training, the latter is more complicated, which require teacher training to improve the knowledge in applying the new approach (Sánchez-Mena & Martí-Parreño, 2017).

Implementation difficulties. In implementing GBL or gamified learning in schools, teachers come across difficulties in classroom management as students are easy to lose control (Demirbilek & Tamer, 2010; Tuparova et al., 2018; Wang & Ha, 2009). Moreover, Both GBL and gamification are reported not easy to fit into the existing teaching routines, or, they can even prevent teachers from complying with the teaching schedule (Becker, 2007; De Grove et al., 2012; Sánchez-Mena & Martí-Parreño, 2017). It reported that classes in some countries or regions are only 30 minutes per time and a few times each week, thus "it is impossible to use digital device(s) during the class hour" (Ismail & Ibrahim, 2018, p. 38) and "implement digital educational game(s) in the class" (p. 39). Ismail and Ibrahim (2018) then proposed to use digital games in after-school hours or at leisure time, which raises concern on monitoring and the misuse of the device. The fixed physical classroom setting is also reported as a barrier for gamified classes (Sánchez-Mena & Martí-Parreño, 2017).

What's more, previous studies reported that more time is needed in both preparing and implementing a GBL or gamified class. Teachers must put more time and effort to analyse conditions, aware of users' ability, choose the proper tool and adjust the tool (Platonova & Bērziša, 2017). The requested time is even longer due to the lack of

related information (Wang & Ha, 2009) and the time-consuming process of implementing it in classes (Koh et al., 2012; Platonova & Bērziša, 2017; Wang & Ha, 2009).

Tool factors

The tool factor mainly focuses on whether the tools provide learning opportunities, whether they are integrated with the curriculum, whether available and whether they bring negative effects.

Learning opportunities. The first and foremost concern is to what extend the GBL or gamified learning tools provide learning opportunities (De Grove et al., 2012). The study of Bourgonjon, Valcke, Soetaert, and Schellens (2010) investigated students' perceptions about the use of video games in the classroom, which revealed that learning opportunities are positively associated with games' perceived usefulness in terms of performance and productivity in the classroom. After that, De Grove et al. (2012) assume that if digital games are of better value in providing more learning opportunities, teachers' acceptance intention will be higher. Notably, the capability of GBL or gamified learning in providing learning opportunities is doubted. Lim (2008) finds that one major barrier in using games in the classroom is teachers' concern about standards, grades and outcome measures. Namely, games might not provide the expected outcomes that align with the needs of test-oriented education, therefore teachers believe that "(educational computer games') educational potential is very low" (Demirbilek & Tamer, 2010, p. 713).

There is also a lack of assessment measuring the capability of GBL or gamified learning in providing learning opportunities, as well as the problems of transferring skills learned in virtual games into real-life educational contexts (Emel'yanenko et al., 2016; Huizenga et al., 2017), therefore, even though the teachers do "have the impression that their students learned from playing or creating games", they hesitate to assert the educational potential of the tools (Huizenga et al., 2017, p. 111).

While the aforementioned studies are about the acceptance intention of digital games, the study of Sánchez-Mena and Martí-Parreño (2017) confirmed the importance of providing learning opportunities for gamified learning, as the authors discussed gamification's benefits, performance and expected outcomes.

Curriculum integration. Digital games have been reported ill-integrated with the curriculum content, which is outlined as one main negative influence on teachers' acceptance of educational video games (Koh et al., 2012; McFarlane, Sparrowhawk, & Heald, 2002; Prensky, 2001; Proctor & Marks, 2013; Tuparova et al., 2018). However, Koh et al. (2012) stressed that this finding is limited to digital games. The integration issue applies to both GBL and gamified learning is the integration with specific subjects being taught, as highlighted by Tuparova et al. (2018) and Sánchez-Mena and Martí-Parreño (2017) respectively.

Resource availability. Previous studies reported that there is a limited resource such as suitable games as supporting materials, high-quality educational games and time to implement GBL or gamified learning (Koh et al., 2012; Sánchez-Mena & Martí-Parreño, 2017; Tuparova et al., 2018). One participant in the study of Wang and Ha (2009) stressed that "I felt the greatest challenge was designing the game" (p. 419) since the games should be both fun and educational that address the characteristics of students' age, students' skill level, class size, difficulty level, game tactics, discipline and so on.

The heavy reliance on technique supports during the development of gamified learning tools is another constrain. Gamified learning tools or other digital pedagogical tools should be based on humanitarian ideas of education as an addition to the traditional forms of teaching; at the same time, the use of new technologies is possible to bring negative impacts on learners (Emel'yanenko et al., 2016). Therefore, the development of those tools should be aware of both the educational features and technological side effects. However, there is a gap in the cooperation between teachers and software developers as the tools created by teachers tend to be weak on the ease of use of the tools, while those

initialled by software developers tend to be weaker in achieving educational functions. To promote the flourishment of gamified learning tools, Wingo et al. (2019) proposed to involve both instructional designers and software developers in the process of tool development.

Negative effects. Previous studies highlighted the negative effects of GBL or gamified learning. Since the higher-order thinking skills (e.g. application, analysis, synthesis, evaluation and creating) (Bloom, 1984; Conklin, 2005) are difficult to be integrated with game features, gamification tools tend to address lower-level learning goals such as remembering and understanding (Sánchez-Mena & Martí-Parreño, 2017). In addition, it has been criticised that GBL often concerns "repetition of cyclic content" that provokes "persistent reengagement" (Sánchez-Mena & Martí-Parreño, 2017, p. 434).

It is mentioned that gamification only creates temporary motivation, and its transformation into long-term behaviour is difficult (Emel'yanenko et al., 2016). Findings of a study by de los Arcos, Faems, Comas-Quinn, and Pulker (2017) explained that the gamification tool is encouraging learners to give "quick and unconsidered responses" by providing game elements (badges and ranking) that are not needed or valued by every learner (p. 135). What's more, the temporary motivation tends to be extrinsic rather than intrinsic, which in turns harm intrinsic motivation: when the students are with intrinsic motivation such as "I want to learn", the tangible and rewards provided by gamification will shift the motivations to extrinsic ones such as "I want to earn the reward" (Hanus & Fox, 2015).

Wang and Ha (2009) proposed that GBL requires students to be equipped with preconditions such as fundamental skills, relevant game knowledge moderate self-control and high motivation in participation, so the use of GBL is more complicated than the traditional direct teaching.

External support

External support is another factor contributing to teachers' acceptance intention to GBL or gamified learning, which includes parents', students' and administrative support.

Parents' negative attitude serves as a factor determining teachers' acceptance intention to GBL or gamified learning (Koh et al., 2012; Tuparova et al., 2018). Parents' main concerns are losing control of the learning process and the misuse of digital devices (Ismail & Ibrahim, 2018). The same principles apply to teachers' negative attitudes (Demirbilek & Tamer, 2010; Wang & Ha, 2009). Unexpectedly, though literature highlights the potential of engaging students, empirical studies revealed that students are possible to lack interest in gamified courses (Sánchez-Mena & Martí-Parreño, 2017). Naming the phenomenon as "students apathy", Sánchez-Mena and Martí-Parreño (2017) explained that students are possible to be not interested because they lack the perceived usefulness of gamified courses, therefore they would perceive the use of gamification as a waste of time (p. 438).

Administrative support is an important push to teachers' acceptance of GBL or gamified learning, which can be divided into three categories: financial support, hardware support and policy support. The cost of educational games or gamified learning tools is a determinant reported by various researchers (Koh et al., 2012; Sánchez-Mena & Martí-Parreño, 2017; Tuparova et al., 2018); what's more, the implementation requires further support on the learning environment (Koh et al., 2012), equipment (Koh et al., 2012), technical infrastructure (Demirbilek & Tamer, 2010) and technical supports (Koh et al., 2012). Besides, policies supporting the use of GBL or gamification is also of great significance (Koh et al., 2012; Sánchez-Mena & Martí-Parreño, 2017).

Table 11. Factors contributing to teachers' acceptance intention of GBL and gamified learning

#	Theme	Reference	
		GBL	Gamification
Tool factor			
1	Learning opportunities	(De Grove et al., 2012; Demirbilek & Tamer, 2010; Emel'yanenko et al., 2016; Huizenga et al., 2017; Lim, 2008)	(Sánchez-Mena & Martí-Parreño, 2017)
2	Curriculum integration	(Koh et al., 2012; McFarlane et al., 2002; Prensky, 2001; Proctor & Marks, 2013; Tuparova et al., 2018)	(Sánchez-Mena & Martí-Parreño, 2017)
3	Recourse availability	(Emel'yanenko et al., 2016; Koh et al., 2012; Tuparova et al., 2018; Wang & Ha, 2009)	(Emel'yanenko et al., 2016; Sánchez-Mena & Martí-Parreño, 2017; Wingo et al., 2019)
4	Negative effects	(Baek, 2008; Wang & Ha, 2009)	(Emel'yanenko et al., 2016; Platonova & Bērziša, 2017; Sánchez-Mena & Martí-Parreño, 2017)
Teacher factor			
5	Mind-set (negative attitude)	(De Grove et al., 2012; Demirbilek & Tamer, 2010; Koh et al., 2012; Tuparova et al., 2018)	(Sánchez-Mena & Martí-Parreño, 2017)
6	Teachers' skill	(De Grove et al., 2012; Tuparova et al., 2018)	(Bourgonjon et al., 2013; Sánchez-Mena & Martí-Parreño, 2017)
7	Implementation difficulties	(Baek, 2008; Becker, 2007; De Grove et al., 2012; Demirbilek & Tamer, 2010; Koh et al., 2012; Tuparova et al., 2018; Wang & Ha, 2009)	(Platonova & Bērziša, 2017; Sánchez-Mena & Martí-Parreño, 2017)
External support			
8	Parent support (process monitoring, device misuse)	(Koh et al., 2012; Tuparova et al., 2018)	n/a
9	Teachers support	(Demirbilek & Tamer, 2010; Ismail & Ibrahim, 2018)	n/a
10	Student support	n/a	(Sánchez-Mena & Martí-Parreño, 2017)
11	School support (financial, hardware and policy)	(Baek, 2008; Demirbilek & Tamer, 2010; Koh et al., 2012; Tuparova et al., 2018; Wang & Ha, 2009)	(Sánchez-Mena & Martí-Parreño, 2017)

2.4. Technology acceptance intention

2.4.1. General technology acceptance intention frameworks

Various frameworks have been established and validated by researchers to illustrate and predict an individual's behavioural intention to new technologies, such as the theory of reasoned action (TRA), the Technology Acceptance Model (TAM), TAM2, TAM3 and the unified theory of acceptance and use of technology (UTAUT). The predecessors of these frameworks are the theory of reasoned action (TRA) and the theory of planned behaviour (TBP).

The theory of reasoned action (TRA) (Fisbein & Ajzen, 1975) assumes that human behaviour is a result of behavioural intention, which in turn depends on attitude toward behaviour and subjective norms. Attitude toward a behaviour is usually defined as "permanent mental or neutral willingness gained from the experience, making the directive or dynamic influence" on an individual's response to a certain behaviour (Ham, Jeger, & Frajman Ivković, 2015, p. 739). Subjective norm refers to a "person's perception that most people who are important to him think he should or should not perform the behaviour in question" (Fisbein & Ajzen, 1975, p. 302), which drives an individual to behave in a certain way to comply with other people's views (Ham et al., 2015). Intention is the central factor in the theory of reasoned action, which is assumed to capture the motivational factors that influence behaviour. Intentions are "indications of how hard people are willing to try…in order to perform the behaviour" (Ajzen, 1991, p. 181). Generally speaking, strong intention indicates a high possibility of the actual behaviour.

The theory of planned behaviour (TBP) involved one more factor on the foundation of TRA: the perceived behavioural control (see Figure 5) (Ajzen, 1991). It is believed that the importance of actual behavioural control is self-evident; however, in predicting intentions and actions, research focus should be on the perception of behavioural control and its impact on intentions and actions (Ajzen, 1991). According to Ham et al. (2015), perceived behavioural control is defined as "a combination of locus of control (belief about the amount of control that a person has over events and outcomes in his life) and self-efficacy (perceived ability to perform the task)" (pp. 740-741). Both the TRA and TBP serve as the basis of most conceptual frameworks that study behaviour or technology adoption (De Grove et al., 2012).

Ajzen (1991) furtherly explained that behavioural achievement "depends jointly on" motivation and ability, in which motivation is equivalent to behavioural intention and ability is equivalent to behavioural control. Therefore, if the users are with sufficient behavioural control and strong behavioural intention, they will not only trigger the final behaviours but also the possible behavioural achievements.

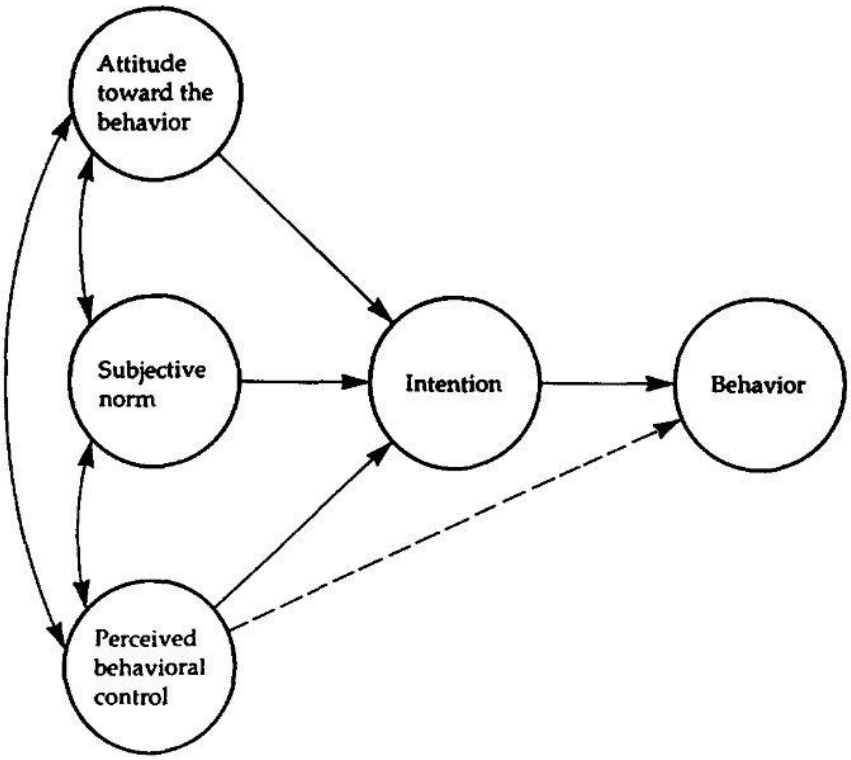

Figure 5. Theory of Planned Behaviour (TPB) (Ajzen, 1991, p. 182).

On the foundation of TRA, Davis (1989) proposed the Technology Acceptance Model (TAM) which is a representative theory illustrating and predicting an individual's acceptance of information technologies (see Figure 6) (Deng, Hong, Ren, Zhang, & Xiang, 2018; Park, 2009; Wong, 2016). In TAM, an individual's actual technology use is affected by behavioural intention, while behavioural intention is impacted by attitude toward using. Two main factors construct an individual's attitude toward technology using: perceived usefulness and perceived ease of use, while the perceived usefulness measures "the degree to which a person believes that using a particular system would enhance his or her job performance" (Davis, 1989, p. 320) and the perceived ease of use is defined as the degree "the degree to which a person believes that using a particular system would be free of effort" (Davis, 1989, p. 320). In TAM, the contributing factors of the perceived usefulness and perceived ease of use are concluded as one word: the external variables.

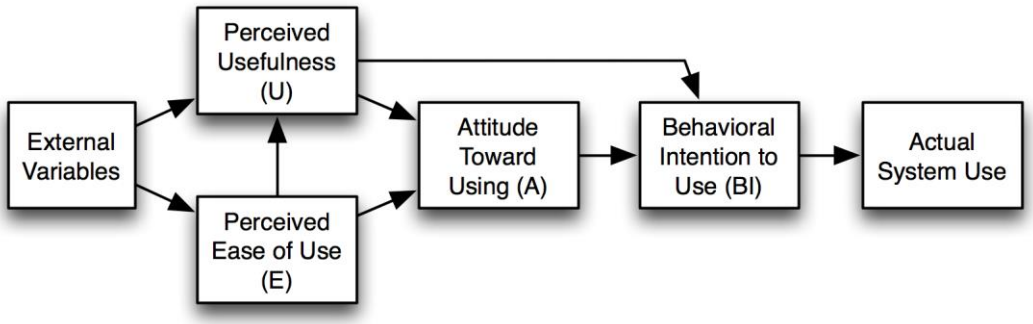

Figure 6. The Technology Acceptance Model (TAM) (Davis et al., 1989)

However, TAM has been criticised for being too "parsimonious" to "give practical advice on how to improve the perceived usefulness and perceived ease of use of technology" (Wong, 2016, p. 316); therefore, researchers have attempted to extend the TAM into more specific ones.

The TAM2 proposed by Venkatesh and Davis (2000) kept the main body of the TAM theory (perceived usefulness, perceived ease of use, intention to use, and usage behaviour) while adding five factors that contribute to the perceived usefulness: subjective norms, image, relevant, output quality and result demonstrability. Experience and voluntariness were added as moderating variables (see Figure 7).

Venkatesh and Davis (2000) stressed the impacts of social factors impinging on an individual when facing the decision of accepting or rejecting a new technology: subjective norm, voluntariness, and image. Venkatesh and Davis (2000) theorised that in system usage settings, there will be a direct compliance-based effect of subjective norm on intention, which "over and above perceived usefulness and perceived ease of use" of system use (p.188). In other words, subjective norm impacts differently on participants in mandatory settings and those in voluntary settings: even when the users perceive a system useful and easy to use, their usage intentions vary because some users are "unwilling to comply with" mandates (Venkatesh & Davis, 2000, p. 188). Meanwhile, an individual tends to establish or maintain a favourable image within a reference group: if important members of a person's social group believe that behaviour should be performed, then "performing it will tend to elevate his or her standing within the group" (Venkatesh & Davis, 2000, p. 189). Named as image, which refers to "the degree to which use of an innovation is perceived to enhance one's…status in one's social system" (Moore & Benbasat, 1991, p. 195), the factor is also playing an important role in affecting the acceptance intention of innovations.

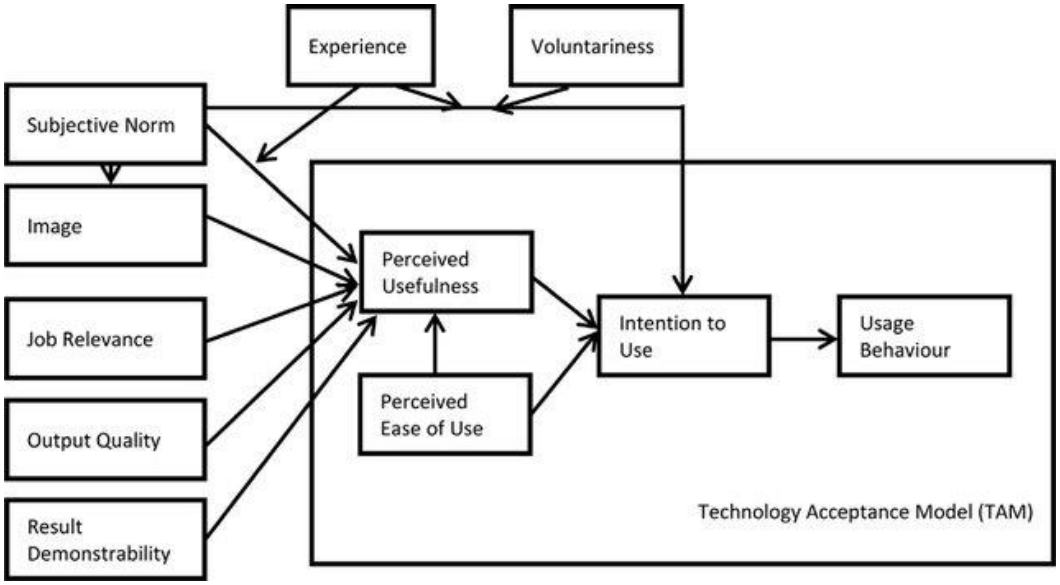

Figure 7. The TAM2 framework (Venkatesh & Davis, 2000, p. 188)

On the basis of TAM2, Venkatesh and Bala (2008) added six more factors that contribute to the perceived ease of use in the TAM3 framework: computer self-efficacy, perception of external control, computer anxiety, computer playfulness, perceived enjoyment, and objective usability (see Figure 8).

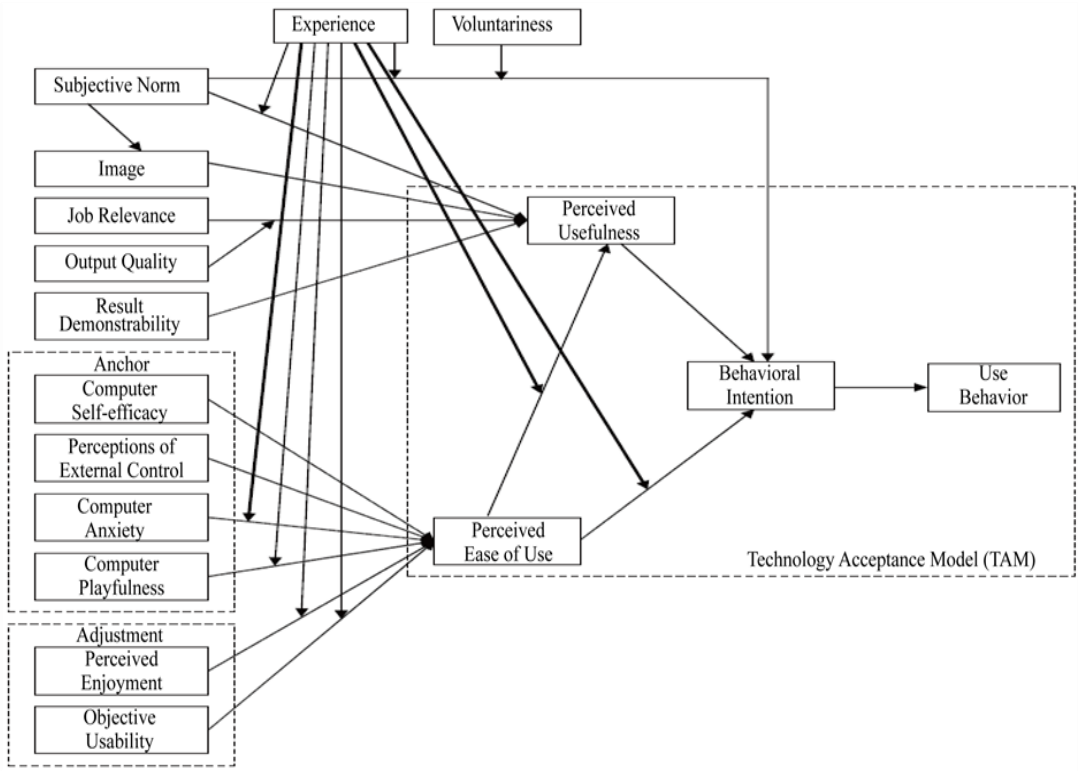

Figure 8. The TAM3 framework (Venkatesh & Bala, 2008)

As pointed by De Grove et al. (2012), the trend to "look for additional factors to increase predictive power" is crystallised in the unified theory of acceptance and use of technology (UTAUT) (see Figure 9) (p. 2024). The UTAUT model established by Venkatesh, Morris, Davis, and Davis (2003) indicated that there are seven determinants to behavioural intention (performance expectancy, effort expectancy, social influence, gender, age, experience and voluntariness of use). Together with facilitating conditions, they contribute to the occurrence of actual behaviours.

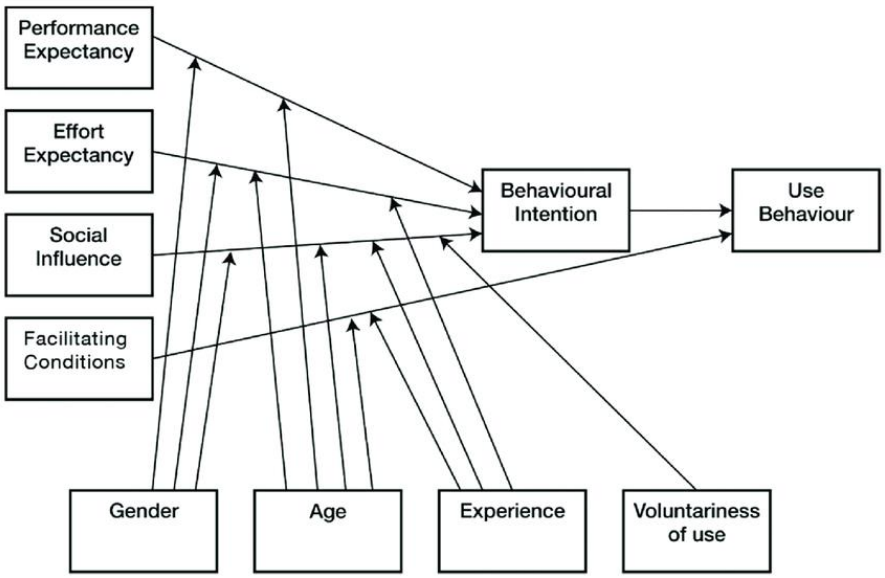

Figure 9. The UTAUT framework (Venkatesh et al., 2003)

2.4.2. TAM-based frameworks with additional factors

TAM has been criticised for being too "parsimonious" to "give practical advice on how to improve the perceived usefulness and perceived ease of use of technology" (Wong, 2016, p. 316). Therefore, researchers have attempted to extend the TAM model into more specific ones by adding factors based on the core of the TAM model (PU, PEoU, BI, and attitude), as summarised in Table 12. The involved factors in Table 12 served for framework establishment in the current study, as detailed in the next section.

Later studies have widely adopted or revised the aforementioned frameworks. For example, Wong (2016) revised a model to investigate the behavioural intention of Hong Kong primary teachers to a new system Direct Subsidy Scheme, which adds the factor "facilitating conditions" from UTAUT to the TAM model; Huang and Liaw (2018) studied learners' adoption intention toward virtual reality learning, in which added three factors: perceived-self efficacy, perceived interaction, and learning motivation;

Cigdem and Ozturk (2016) proposed a theoretical framework to model students' adoption intention toward learning management systems (LMS) at a Turkish post-secondary vocational school. Quite similar with Huang and Liaw (2018) who consider self-efficacy and interactivity, Cigdem and Ozturk (2016) added one more factor to address the learning opportunity a new technology provides: multimedia instruction.

Beggs (2000) investigated the factor contributing to the use of technology. Different from other TAM-related frameworks, Beggs (2000) did not enlist PU as a core factor; instead, Beggs (2000) expanded the meaning of PU, such as improved student learning, advantages over traditional teaching, increased student interest, and personal comfort. It is notable that ease of use in the study of Beggs (2000) was not limited to technical issues but implementation ones, such as equipment availability, compatibility with discipline, time needed to learning, and compatibility with materials. Organisational factors were also considered, such as administrative support, colleague use, and training.

Y. Sun and Jeyaraj (2013) conducted a longitudinal study that involved 132 users to investigate individuals' behavioural intention and continuance to information technology. Similar to Beggs (2000), Y. Sun and Jeyaraj (2013) valued the factor "perceived compatibility". Perceived compatibility was linked to whether the technology fits well with the way the participant would like to study in the course and the participant's work style. Besides, Y. Sun and Jeyaraj (2013) nominated the factor "expertise", which refers to participants' skills in using digital devices, the expertise to use information technologies, and the expertise in using different information systems. The same with compatibility and expertise, four factors were added: personal innovativeness, self-efficacy, facilitating conditions, and social influence.

Park (2009) analysed university students' behavioural intention to use e-learning. The results showed that besides the core determinants of TAM (PU, PEoU, attitude, BI), three key determinants are contributing to the final e-learning use intention: self-efficacy, subjective norm, and system accessibility. A consecutive study by Park, Nam, and Cha (2012) narrowed the research objectives from e-learning to a more specific concept "mobile learning". The TAM-related factors were named endogenous latent variables, which include PU, PEoU, attitude, and behavioural intention; the TAM-extended factors were named exogenous latent variables, which include mobile learning self-efficacy, major relevance, system accessibility, and subjective norm. compared with the framework established in the study of Park (2009), one factor was added in the theoretical framework: major relevance, namely the content relevance.

M.-C. Lee (2010) synthesised the expectation–confirmation model (ECM), the technology acceptance model (TAM), the theory of planned behaviour (TPB), and the flow theory to construct a theoretical framework to explain user intention to continue using e-learning. By combining the key factors in those frameworks into one, M.-C. Lee (2010)

proposed that five main constructs are contributing to users' continued intention of using e-learning: satisfaction, attitude, flow, subjective norm, and perceived behavioural control. Under the construct flow, two sub-constructs were listed: perceived enjoyment and concentration.

Varannai, Sasvári, and Urbanovics (2017) investigated university students' acceptance intention of gamification with the use of a gamified learning tool named Kahoot!. Though the proposed research model was based on the TAM model by Davis (1989), PU the core construct was not included. Instead, four constructs were nominated as parallel important factors contributing to acceptance intention of gamification: perceived utility, PEoU, experience, and availability (of gamified learning tools).

Yusoff, Crowder, and Gilbert (2010) conceptualised a framework describing the factors contributing to students' acceptance intention to serious games, in which contains four main constructs: transfer of learnt skills, learner control, situated and authentic learning, and reward.

Also focusing on university students' acceptance intention of gamified learning, Vanduhe et al. (2020) focused on the social characteristics of gamification: social influence and social recognition. Social influence indicates that other people's views influence an individual's intentions to use a particular technology, which "is a significant factor in determining the acceptance and use of technology" (Vanduhe et al., 2020, p. 21476); social recognition is to provide rewards to an individual to help understand his or her abilities and skills. Task-technology fit was also nominated, which refers to the degree of integration of educational content and gamification features.

While researchers like Herzig, Strahringer, and Ameling (2012) tend to focus on entertainment features (e.g., enjoyment, flow), Bourgonjon, Valcke, Soetaert, and Schellens (2009) laid more emphasis on educational functions the video games can bring to students. Bourgonjon et al. (2009) explored the acceptance of video games in the classroom by secondary school students. The authors criticised that it is too narrow to focus on results only when investigating performance and perceived usefulness, because "education means more than mere outcome" (p. 653). Accordingly, Bourgonjon et al. (2009) divided the perceived usefulness into "perceived usefulness-product" (PU-Product) and "perceived usefulness-process" (PU-Process), with PU-Product referring to the perceived usefulness on academic performance and PU-Process referring to the perceived usefulness on the learning process.

Since earlier studies demonstrated that experienced video gamers are more likely to be in favour of game-based learning than the students who seldom play, game experience was included in the study of Bourgonjon et al. (2009). Similarly, (De Grove et al., 2012) investigated teachers' adoption intention of digital games in formal education, which contains three main factors besides the core ones in TAM: experience, learning opportunities, and curriculum-relatedness. It is notable that though game experience was evidenced important in influencing user acceptance of game-based learning and was included as one contributing factor in some e-learning studies (Revythi & Tselios, 2019) and gamification acceptance studies (Varannai et al., 2017), whether it applies to gamified learning needs further confirmation.

When comes to technologies for educational purposes, the proposed frameworks tend to value the product's educational potential. Ibrahim and Jaafar (2011) conceptualised an educational games acceptance model, which is similar to the UTAUT framework: they replaced the factor "facilitating condition" with "learning opportunity" and "enjoyment". As a whole, Ibrahim and Jaafar (2011) proposed that five factors are contributing to behavioural intention, including performance expectancy, effort expectancy, social influence, learning opportunities and enjoyment.

In gamified learning studies, the integration of game features and pedagogical needs is always a challenging difficulty (Bourgonjon et al., 2013; Koh et al., 2012; Tuparova et al., 2018; Van Rosmalen et al., 2015), therefore De Grove et al. (2012) took "curriculum-relatedness" into consideration. Similarly, Park et al. (2012) used the term "major

relevance" to indicate the relatedness of a technology to pedagogical contents. There are similar words, such as knowledge learning (Lin & Chen, 2013), instruction (Cigdem & Ozturk, 2016), learning opportunity (Ibrahim & Jaafar, 2011).

Researchers also aimed to specify the adoption intention of specific technologies, such as digital games or digital-game-based learning tools. Frameworks of this type lay focus on the effects of sensory enjoyment brought by game-playing. For example, Holsapple and Wu (2007) investigated users' adoption intention to entertainment-oriented information technology such as virtual worlds, in which involves fantasy, escapism and arousal; Lin and Chen (2013) summarised four factors constructing behavioural intention to use, which contains hedonic experience, namely the experience of gaining pleasure and enjoyment (Higgins, 2006).

Among various factors added to the basic model TAM and UTAUT, this author found two important ones for our research context: subjective norm and perceived self-efficacy. Subjective norm refers to "person's perception that most people who are important to him (or her) think he (or she) should or should not perform the behavior in question" (Fisbein & Ajzen, 1975, p. 302). For example, users might adopt new technology to meet social needs, to prepare for future jobs, to copy with peer pressure or to meet others' expectations (Park et al., 2012; Wong, 2016), which directly contribute to the perceived usefulness of a technology. Self-efficacy concerns "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3). There are also terms relating self-efficacy to specific contexts, such as computer self-efficacy and e-learning self-efficacy. High self-efficacy enables the user to perform the technologies with confidence, and vice versa. Self-efficacy influences the users' perceived ease of use toward a technology.

In the study investigating 345 university students' acceptance intention of learning management systems (LMSs), (Revythi & Tselios, 2019) stressed the importance of system accessibility. The structural modelling equation test confirmed that system accessibility is significantly contributing to the perceived usefulness, perceived ease of use, attitude, and behavioural intention of LMSs.

P. P. Sun and Mei (2020) focused on preservice Chinese-as-a-second/foreign-language (L2 Chinese) teachers to investigate the factors influencing their intention to use educational technologies in classrooms. Besides attitude, perceived usefulness, and intention to use, P. P. Sun and Mei (2020) proposed that three are three more factors: technology self-efficacy, facilitating conditions, and experience of technology use. While lots of other students regard experience as a mediating factor, the study of P. P. Sun and Mei (2020) is one of few that list it as one main construct of technology acceptance intention.

Teo, Lee, and Chai (2008) proposed that facilitating conditions are contributing to better perceived ease of use, subjective norms enhance the perceived usefulness, and the four factors contribute to computer attitude.

Ranellucci, Rosenberg, and Poitras (2020) examined pre-service teachers' attitudes towards technology from the value and cost perspective. According to Ranellucci et al. (2020), there are eight factors influencing teachers' behavioural intentions, including four values (expectancy, attainment value, intrinsic value, and utility value) and four costs (task effort cost, outside effort cost, loss of valued alternatives, and emotional cost) (see Figure 10).

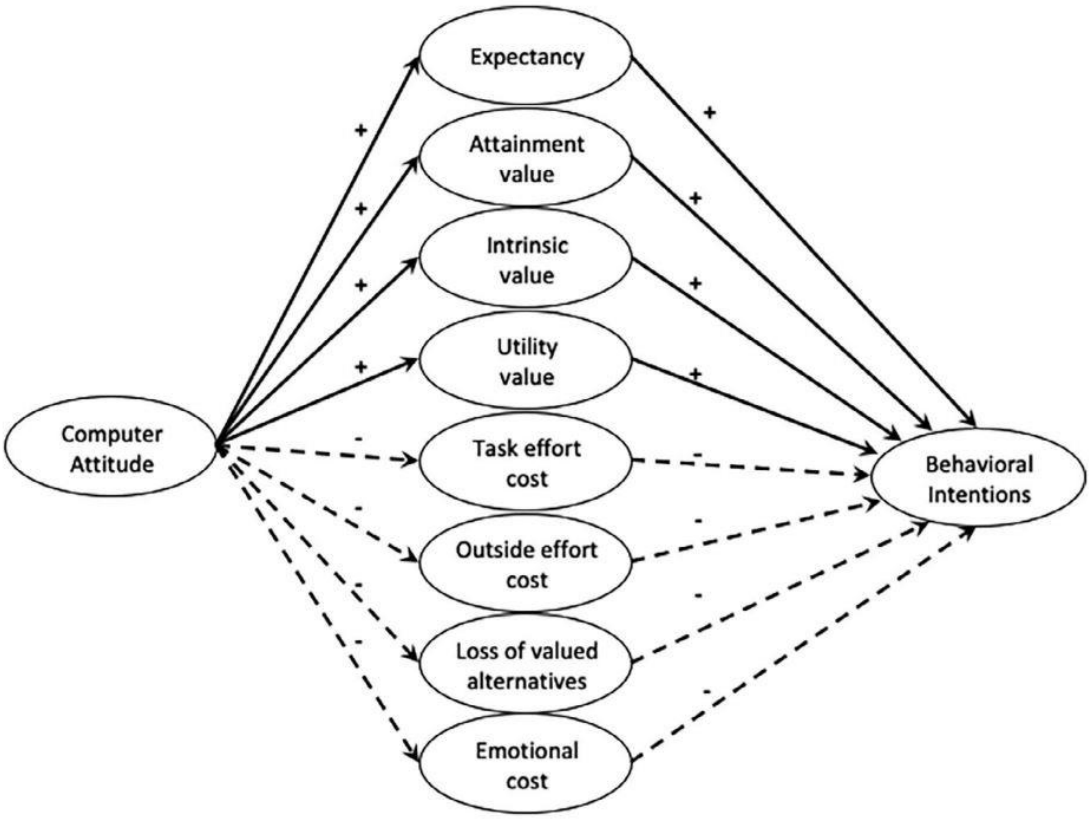

Figure 10. Framework proposed by Ranellucci et al. (2020)

Note: solid lines indicate positive relations and broken lines indicate negative relations.

Table 12. Summary of previous TAM-related frameworks

Source	TAM-expanding factor	TAM core factor	Framework
Venkatesh and Davis (2000)	Subjective norms, image, relevant, output quality, result demonstrability, experience, voluntariness	PU, PEoU, intention to use, usage behaviour	Technology acceptance model 2 (TAM2)
Venkatesh and Bala (2008)	Subjective norms, image, relevant, output quality, result demonstrability, experience, voluntariness, computer self-efficacy, perception of external control, computer anxiety, computer playfulness, perceived enjoyment, objective usability	PU, PEoU, intention to use, usage behaviour	Technology acceptance model 3 (TAM3)
Venkatesh et al. (2003)	Performance expectancy, effort expectancy, social influence, facilitating conditions, gender, age, experience, voluntariness of use	BI, use behaviour	The unified theory of acceptance and use of technology (UTAUT)
Ibrahim and Jaafar (2011)	Performance expectancy, effort expectancy, social influence, learning opportunity, enjoyment	BI	Educational games acceptance model
Beggs (2000)	Improved student learning, advantage over traditional teaching, equipment availability, increased student interest, compatibility with discipline, time needed to learn, materials in discipline, compatibility with materials, training, administrative support, personal comfort, colleague use	Ease of use	Factors influencing the use of technology
Y. Sun and Jeyaraj (2013)	Perceived compatibility, perceived ease of use, expertise, personal innovativeness, self-efficacy, facilitating conditions, social influence	PU, PEoU	Individuals' behavioural intentions to information technology
Park (2009)	e-learning self-efficacy, subjective norm, system accessibility	PU, PEoU, attitude, intention to use	University students' behavioural intention to use e-learning
Park et al. (2012)	Mobile learning self-efficacy, major relevance, system accessibility, subjective norm	PU, PEoU, attitude, BI	University students' behavioural intention to use mobile learning
Revythi and Tselios (2019)	e-learning self-efficacy, social norm, system accessibility, year	PU, PEoU, attitude, BI	University students' acceptance intention to e-learning (learning management systems)
Varannai et al. (2017)	Perceived utility, experience, availability	PEoU, attitude, intention	University students' acceptance intention to gamification
Vanduhe et al. (2020)	Task-technology fit, social influence, social recognition	PU, PEoU, attitude, continuance intention	University students' continuance intention to gamification
Yusoff et al. (2010)	Transfer of learnt skills, learner control, situated and authentic learning, reward	PU, PEoU, BI	Serious games
Bourgonjon et al. (2009)	Usefulness-process (PU-Process), usefulness-product (PU-Product), experience	PEoU, BI	The acceptance of video games in the

			classroom by secondary school students
Wong (2016)	Facilitating conditions	PU, PEoU, attitude, BI	Primary teachers' behavioural intention to educational technologies
Lin and Chen (2013)	Knowledge learning, self-efficacy, social interaction, hedonic experience	BI	Senior game-technology acceptance model (SGTAM)
Herzig et al. (2012)	Flow, enjoyment	PU, PEoU, BI	Gamification effects on user acceptance
Cigdem and Ozturk (2016)	Multimedia instruction, interactivity in learning management system (LMS), LMS self-efficacy	PU, PEoU, BI, perceived satisfaction	Students' behavioural intention to use LMS
Huang and Liaw (2018)	Perceived self-efficacy, perceived interaction, learning motivation	PU, PEoU, intention to use	Learners' intention towards virtual reality learning
Zhang, Yin, Luo, and Yan (2017)	Perceived learner control, e-learning self-efficacy, personal innovativeness in IT	PU, PEoU, intention	Adoption intention for MOOC learning
De Grove et al. (2012)	Experience, learning opportunities, curriculum-relatedness	Usefulness, ease of use, BI	Teachers' adoption intention of digital games in formal education
Deng et al. (2018)	Perceived risk, trust, control variables (gender, age, education and chronic diseases)	PU, PEoU, adoption intention	Patients' adoption intention toward mobile-assisted health services
M.-C. Lee (2010)	Perceived enjoyment, confirmation, concentration, satisfaction, attitude, subjective norm, perceived behavioural control	PU, PEoU, continued intention	Continuance intention towards e-learning
Sugar et al. (2004)	Attitude toward the behaviour, subjective norm, perceived behavioural control	BI	Teachers' decision to adopt new technology
Al Kurdi, Alshurideh, Salloum, Obeidat, and Al-dweeri (2020)	Social influence, perceived enjoyment, self-efficacy	PU, PEoU, acceptance	University students' e-learning system acceptance
Teo et al. (2008)	Facilitating conditions, subjective norms	Ease of use, usefulness, computer attitude	Pre-service teachers' computer attitude
P. P. Sun and Mei (2020)	Technology self-efficacy, facilitating conditions, experience of technology use	Intention, PU, attitude	Teachers' adoption of educational technology
Ranellucci et al. (2020)	Expectancy, attainment value, intrinsic value, utility value, task effort cost, outside effort cost, loss of valued alternatives, emotional cost	BI	Pre-service teachers' use of technology

Note: PU: perceived usefulness; PEoU: perceived ease of use; BI: behaviour intention.

2.4.3. Other frameworks

The same with the UTAUT, some frameworks investigated the factor contributing to behavioural intention, in which PU and PEoU are absent. Namely, the authors directly link the "external variables" with attitude or behavioural intention. For example, Khlaif (2018) established a theoretical framework describing the factors contributing to teachers' acceptance intention to mobile technologies in K-12 settings, which contains four main determinants: performance expectancy, effort expectancy, facilitating conditions, and social influence. Though frameworks of this type provide valuable insights as well, this author did not include them in framework selection because they involve factors of high variance.

Researchers nominated other frameworks, such as the characteristics of innovation acceptance by Rogers (2010). According to Rogers (2010), there are five main stages in accepting an innovation, including knowledge, persuasion, decision, implementation, and confirmation. The knowledge refers to users' understanding of the innovation, which is affected by users' socioeconomic characteristics, personality variables, and communication behaviours. During the persuasion stage, five attributes of innovations impact the final acceptance: the relative advantages of the innovation compared with the previous ones; the compatibility of the innovation with the existing values, past experience, and needs of potential adopters; the complexity of the innovation (the degree two which an innovation is perceived as difficult to understand and use); the trialability (the degree to which an innovation may be experimented with on a limited basis), and the observability of the innovation's results (Rogers, 2010).

Diffusion of innovations is a theory that seeks to explore how, why, at what rate new technology spread. Rogers (2010) proposed that four main elements influence the spread of a new idea: the innovation itself, communication channels, time, and a social system. Innovation is a broad concept, which refers to any idea, practice, or object that is perceived as new by an individual or other unit of adoption could be considered an innovation available for study; communication channels enable information to be transferred among adopters; social system is the combination of both external influences and internal influences (Rogers, 2010). The four elements can be regarded as the prerequisites of an IT innovation's popularity. However, the four main elements are general concepts, so more studies are needed to detail the specific connotations in varied contexts.

Attributes of innovations

Relative advantage: Degree to which an innovation is perceived as better than the idea it supersedes.

Compatibility: Degree to which an innovation is perceived as being consistent with the existing values, past experience, and needs of potential adopters.

Complexity: Degree two which an innovation is perceived as difficult to understand and use.

Trialability: Degree to which an innovation may be experimented with on a limited basis.

Observability: Degree to which the results of an innovation are visible to others.

Emphasising the importance of personal innovativeness (the degree to which an individual adopts a new idea), Rogers (2010) categorised users into five groups: innovators, early adopters, early majority, late majority, and laggards. According to Rogers (2010), a small portion of people will first accept new information or technologies because they are high in personal innovativeness (approximately 2.5%), followed by early adopters (13.5% of the target people). The early majority and late majority tend to follow when other people have already accepted the new information or technologies (34% for each type of majority), while 16% of the laggards tend to be resistant to change. Though the proportion of each type of people may vary in different contexts, namely the exact percentage may be different, the theory proposed by Rogers (2010) stressed the importance of personal innovativeness in technology acceptance.

Some other frameworks attempted to explain technology adoption and use from other perspectives. For example, Koh et al. (2012) proposed that three determinants are contributing to the final adoption of video games for educational purposes: demographic factors, push factors, and pull factors. Push factors are like the facilitating conditions provided to teachers to facilitate the adoption process, which include policies, curriculum, and environmental support; pull factors are related to teachers' positive attitude, which can be explained by the diffusion of innovations theory by Rogers (2010): in accepting new technologies, the users can be categorised as innovators, early adopters, early majority, late majority, and laggards. The criterion for the adopter categorisation is personally innovativeness, which can be determined by users' demographic characteristics, such as the users' teaching level, years of experience, and teaching subject. The diffusion of innovations theory clarified what to support to enhance user technology acceptance, which is of great significance; however, this framework is still at the initial stage, so more studies are needed to push the frontier further. Due to the immaturity of the theory, the diffusion of innovations theory was not included in the current study.

2.5. Determinants of PU and PEoU

2.5.1. Importance of investigating the factor determinants

Ajzen (1991) highlighted that researchers should not only predict human behaviours but also explain them by dealing with the antecedents of factors determining intentions and actions. However, there is a dearth of research on what precisely makes a learning tool perceived useful or perceived easy to use (Aljukhadar & Senecal, 2015; Alsabawy, Cater-Steel, & Soar, 2016); namely, little attention has been paid to the determinants of PU and PEoU, neither their antecedents (Aziz & Macredie, 2005; Reinicke & Marakas, 2005; Venkatesh & Davis, 1996). With this gap existing, it is understudied that what contributes to users' resistance in using a specific educational technology, as well as how to address the hindrances. The lack of study is even more obvious when comes to gamification, as it is a new research area.

It is of great significance to better understand PU and PEoU, the two main determinants of technology acceptance, in educational gamification contexts. To improve users' acceptance and utilisation of gamified learning tools, "it is necessary to better understand the antecedents and determinants of key acceptance constructs" (Venkatesh & Davis, 1996). What's more, Gamification as an immature-yet research area originated from video game concepts, so to some extent, it shares similarities with video games. The use of video games for educational purposes is criticised for various reasons, such as irrelevance with educational content (Kirriemuir & McFarlane, 2004; Van Rosmalen et al., 2015) and time-demandingness in implementation (Kirriemuir & McFarlane, 2004), which may also apply in some gamification practices. Consequently, gamification as an educational technology cannot be regarded as a general one, which highlights the necessity of examing the determinants of PU and PEoU in gamification contexts. Besides, Baaren, Van De Wijngaert, and Huizer (2008) highlighted that "it is important to realize that the meaning of perceived usefulness … has an expiry date", so even though the determinants of PU and PEoU of a technology have been investigated (which is not in gamification studies), latest research examining the determiants should also be followed.

2.5.2. Necessity of investigating the determinants of PU and PEoU

Though the TAM model is a well-known framework for technology acceptance intention, it has been criticised for being too "parsimonious" to "give practical advice on how to improve the perceived usefulness and perceived ease of use of technology" (Wong, 2016, p. 316); therefore, it is a trend to explore the determinants of technology acceptance intention in specific contexts. In another word, it is a trend to interpret and enrich the "external variables" in specific contexts.

For example, Raza, Umer, and Shah (2017) proposed that there are four new determinants of adoption intention of mobile banking, including resistance, perceived risks, computability, and awareness; in the field of mobile wallet, Mei and Aun (2019) suggested that three factors are contributing to the adoption intention (convenience, confidentiality, and social influence); focusing on the adoption intention of an organisation's website, Aljukhadar and Senecal (2015) concluded that PU and PEoU consist of four determinants including site aesthetics, information quality, site interactivity, and site trust.

It is obvious that the connotations of PU and PEoU in distinguished contexts are of significant differences, so the exploration of those for gamified foreign language learning is essential. Since there is a lack of closely related studies, this author reviewed articles that working on the general technology concept, such as digital games, instructional technologies, e-learning, mobile learning, and digital educational interventions.

The involved determinants can be categorised into four types: the parallel determinants with PU or/and PEoU, the determinants of both PU and PEoU, the determinants of PU, and the determinants of PEoU. The specified determinants are as detailed in Tables 13, 14,15, and 16.

2.5.3. Parallel determinants of PU and/or PEoU

In previous frameworks that relate to technology acceptance intention, researchers proposed parallel determinants with PU and PEoU (see Table 13).

Some researchers focused on technical issues when nominating the potential determinants.

Y. Sun and Jeyaraj (2013) attempted to separate the technology acceptance process into two parts: the intention to adopt an IT innovation at the early stage, and the intention to adopt or continue using an IT innovation at the later stage. Accordingly, the technology acceptance studies should be separated into two parts: the acceptance intention at the early stage, and the intention to adopt or continue at the later stages. According to Y. Sun and Jeyaraj (2013), the determinants are different at these two stages. At the early stage, individual characteristics affect the decision-making process, which include the user's expertise, personal innovativeness, and self-efficacy; at the later stage, contextual factors affect the intention to use or continue using the innovation, which include facilitating conditions and social influence. Throughout the technology-adoption process, the innovation attributes of the IT innovation play a role, which include PU, PEoU, and the perceived compatibility.

In the frameworks proposed by Y. Sun and Jeyaraj (2013), perceived compatibility is one sub-theme of the innovation attributes, together with PU and PEoU; personal innovativeness and social influence are the sub-themes for individual characteristic and contextual factors respectively. Malaquias, Malaquias, and Hwang (2018) adopted these two sub-themes as independent determinants. According to Malaquias et al. (2018), there are four independent determinants deciding whether students use a serious game, including PU, PEoU, personal innovativeness in IT, and social influence.

Researchers also nominated determinates for specific contexts. For example, virtual reality learning highlights the interactions between humans and computers, so in exploring individuals' acceptance intention to virtual reality learning, perceived interaction was list as a determinant (Huang & Liaw, 2018). The determinants that apply to gamification contexts need further exploration.

In gamification studies, it has long been criticised that the integration of game features and pedagogical content is poor. Therefore, researchers valued the educational impact of games or gamification for educational purposes. Determinants that are for education were specified, such as "learning opportunities" (De Grove et al., 2012), "curriculum relatedness" (De Grove et al., 2012), and "compatibility with discipline" (Beggs, 2000).

Table 13. Parallel determinants of PU and/or PEoU in previous studies

Determinant	Objective	Technology	Reference
Learning opportunities, experience, curriculum relatedness	Teachers	Digital games	De Grove et al. (2012)
Personal innovativeness in IT, social influence	Students	Serious games	Malaquias et al. (2018)
Improved students learning, advantage over traditional teaching, equipment availability, increased student interest, ease of use, compatibility with discipline, time needed to learn, materials in discipline, computability with material, training, administrative support, personal comfort, colleague use	Faculties	Instructional technologies	Beggs (2000)
Perceived compatibility	Individuals	Information technology	Y. Sun and Jeyaraj (2013)
Locus of control, ambiguity tolerance, risk-taking propensity, experience, education, professionalism, computer playfulness, general computer self-efficacy, self-esteem, and computer anxiety	Users	Information system	Reinicke and Marakas (2005)
e-learning self-efficacy, personal innovativeness in IT, perceived learner control	Learners	Massive Open Online Courses (MOOCs)	Zhang et al. (2017)
Perceived self-efficacy and perceived interaction	Learners	Virtual reality learning	Huang and Liaw (2018)

2.5.4. Determinants of both PU and PEoU

Researchers proposed that some variables can be the determinants of both PU and PEoU. The variables include learning curriculum experience, experience, facilitating conditions, self-efficacy, subjective norm, system accessibility, major relevance, instruction, and interactivity, as detailed in Table 14.

Table 14. Determinants of both PU and PEoU in previous studies

Determinant	Objective	Technology	Reference
Learning opportunities, curriculum relatedness, experience	Teachers	Digital games	De Grove et al. (2012)
Experience	Secondary school students	Video games	Bourgonjon et al. (2009)
Facilitating conditions	Teachers	Educational technology	Wong (2016)
e-learning self-efficacy, subjective norm, system accessibility	University students,	e-learning	Park (2009)
Mobile learning self-efficacy, major relevance, system accessibility, subjective norm	University students	Mobile learning	Park et al. (2012)
Multimedia instruction, interactivity, LMS self-efficacy	Students	Learning management systems (LMS)	Cigdem and Ozturk (2016)

2.5.5. Determinants of PU

Since gamification is evidenced promising in engaging users by providing enhanced experience, the providence of enjoyment is regarded as one main function of gamification. The ultimate status of enjoyment is absolute immersion, or named flow, so flow was proposed as one determinant of PU for gamification (Herzig et al., 2012). Terms with similar connotations were proposed, such as "affective needs" (Hashim, Tan, & Rashid, 2015) and "enjoyment" Rahmi, Birgoren, and Aktepe (2018).

Education-related concepts were listed, too, such as learning opportunities (Baptista & Oliveira, 2019; Bourgonjon et al., 2013) and learning environments (Teo, 2011b) (see Table 15).

Baptista and Oliveira (2019) conducted a meta-analysis based on the 586 variables found in 54 studies and 59 datasets working on factors contributing to gamification or serious games. Besides Z value, the standard score that gives information about how far a data point is from the mean, the authors also calculated the weight of each variable. Weight was calculated as the result of "the number of times a relationship is found to be statically significance divided by the number of times it appears in the literature" (p.309). The weight results identified four best predictors of users' acceptance intention to gamification or serious games and six promising predictors. The best predictors are: attitude on intention, enjoyment on intention, PU on intention, and PEoU on PU; the promising predictors are: intention on brand attitude, enjoyment on brand attitude, PU on brand attitude, hedonic value on behavioural intention, learning opportunities on PU, and socialness on PU. The authors furtherly summarised a theoretical framework based on the weight and meta-analysis. In the proposed model, four factors are contributing to both brand attitude and behavioural intention: the perceived usefulness, hedonic value, attitude, and enjoyment. Besides, the perceived usefulness contains three determinants: the perceived ease of use, learning opportunities, and socialness (see Figure 11).

It is interesting that the proposed theoretical framework by Baptista and Oliveira (2019) makes attitude a parallel factor with PU, while most of the TAM-extending frameworks regard PU as the contributing factor of attitude. Besides, the results showed that PEoU is set as a contributing factor to PU, which is also rare in other studies (see Figure 11). Teo (2011a) hold the same viewpoint: according to Teo (2011a), PU consists of two factors: subjective norm and perceived ease of use.

The survey items measuring PU in the TAM model also demonstrated that there are different determinants of PU. For example, Proctor and Marks (2013) used the TAM survey items to measure PU, which suggests that the PU is expected to indicate the perceived usefulness on five sub-categories: the usefulness in accomplishing more, the PU in improving performance, the PU in increasing productivity, the PU in making job easier, and the perceived usefulness in job (from a general perspective). Though always being mentioned together, the survey items are measuring different aspects of PU (e.g., productivity and easiness of the job). It confirms the importance and necessity of exploring the connotations of PU in different contexts.

Alsabawy et al. (2016) explored the determinants of PU of e-learning systems with a focus on technical service quality. The framework indicates that service delivery quality is the only factor that directly contributes to the PU of e-learning systems, which was influenced by three factors: system quality, information quality, and IT infrastructure services (Figure 12). It can be inferred that if view PU from technical perspective, the determinants will be significantly different. As will be discussed in the research findings, the connotations of PU and/or PEoU vary in different contexts from different perspectives, which furtherly confirms the necessity of exploring the determinants of PU and PEoU in the current research contexts.

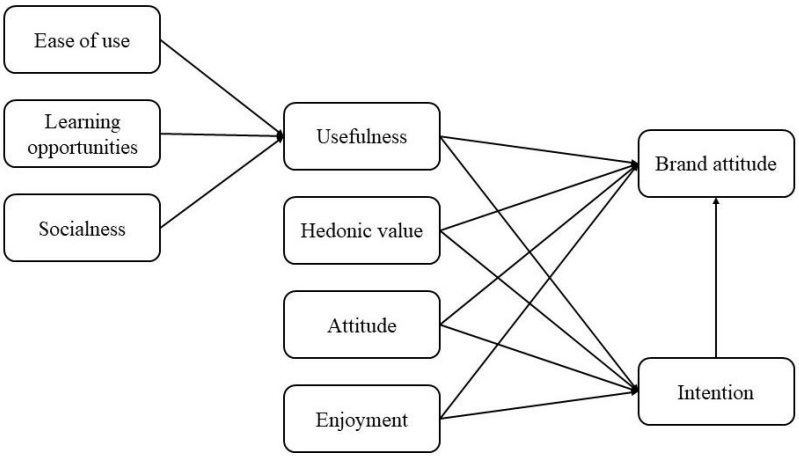

Figure 11. Factors contributing to brand attitude and intention (Baptista & Oliveira, 2019, p. 310)

Table 15. Determinants of PU in previous studies

Determinant	Objective	Tool	Reference
Flow	Users	Gamification	Herzig et al. (2012)
Ease of use, learning opportunities, socialness	n/a	Gamification and serious games	Baptista and Oliveira (2019)
System quality, information quality, and IT infrastructure services	Students	e-learning systems	Alsabawy et al. (2016)
Critical mass, learning opportunities, subjective norm, personal innovativeness, experience, complexity	Secondary school teachers	Game-based learning	Bourgonjon et al. (2013)
Cognitive needs, affective needs, social needs	Adult learners	Mobile learning	Hashim et al. (2015)
Self-efficacy, subjective norm, interaction, enjoyment, compatibility	Users	e-learning	Rahmi et al. (2018)
Learning environment, course delivery, tutor attribute, facilitating conditions	Teachers	e-learning	Teo (2011b)
Subjective norm, perceived ease of use	Teachers	Technology	Teo (2011a)
Ease of use, learning opportunities, socialness	Users	Gamification and serious games	Baptista and Oliveira (2019)

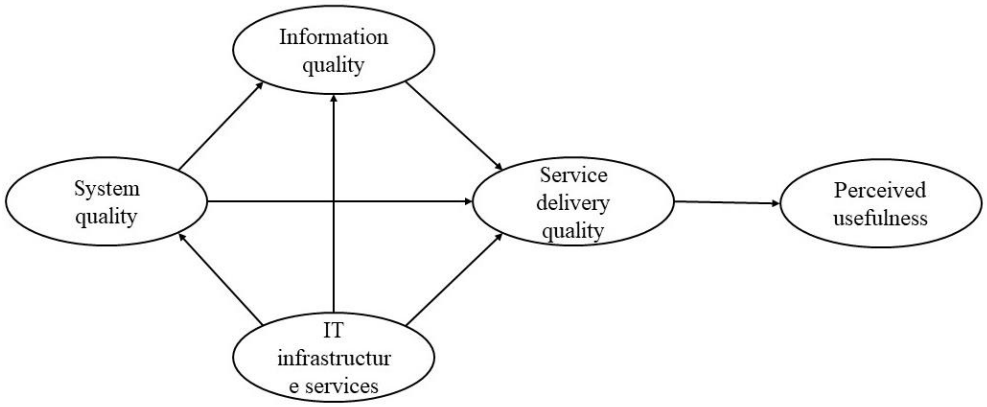

Figure 12. Determinants of PU (Alsabawy et al., 2016, p. 847)

2.5.6. Determinants of PEoU

Determinants of PEoU proposed in previous studies overlap with those of PU, such as enjoyment (Herzig et al., 2012; Rahmi et al., 2018), flow (Herzig et al., 2012), interaction (Rahmi et al., 2018), and facilitating conditions Teo (2011a) (Table 16).

Some concepts are obviously connected with the perceived ease of use of digital devices, such as computer anxiety (Brown, 2002; Rahmi et al., 2018) and usability (Venkatesh & Davis, 1996).

Some other concepts are potentially related to PU, such as effectiveness and efficiency (Aziz & Macredie, 2005).

Table 16. Determinants of PEoU in previous studies

Determinant	Objective	Tool	Reference
Enjoyment, flow	Users	Gamification	Herzig et al. (2012)
Self-efficacy, subjective norm, interaction, enjoyment, anxiety	Users	e-learning	Rahmi et al. (2018)
Facilitating conditions	Teachers	Technology	Teo (2011a)
Efficiency, effectiveness, user satisfaction	Users	Information systems	Aziz and Macredie (2005)
Ease of finding, ease of understanding, self-efficacy, and computer anxiety	Students	Web-based learning technologies	Brown (2002)
Computer self-efficacy, objective usability, direct experience	Users	Training interventions	Venkatesh and Davis (1996)

CHAPTER 3

Literature Review-2

(Effectiveness of educational gamification)

03

Literature Review
(Effectiveness of educational gamification)

The three-to-four-year doctoral study was an exploratory experience during which successful attempts were not guaranteed. For this author, gamification studies mainly focused on two aspects: the gamification design and effectiveness assessment. Therefore, when firstly cut into this new research area, this author made connections between gamification study and effectiveness assessment.

Literature reading related to the effectiveness of gamification revealed several difficulties, such as heavy reliance on computer science skills and methodological limitations. Together with implementation difficulties, this author abandoned effectiveness studies of gamification.

Instead, this author considered a similar concept perceived usefulness, which emphasised user perception rather than the concrete contributions of gamified learning tools. As digging deeper into this topic, this author noticed the technology acceptance model (TAM) (Davis, 1989), which contains two main constructs: perceived usefulness (PU) and perceived ease of use (PEoU). Therefore, this thesis eventually turned to factors that contribute to secondary-school teachers' acceptance intention to gamified learning tools.

However, literature reading on the topic of gamification's effectiveness revealed important and interesting findings, which are also of great importance to explain users' acceptance intention. After all, in the TAM model, it is proposed that both PU and PEoU construct the final acceptance intention, while PEoU is also contributing to PU. In another word, PU can be the dominant factor influencing the final technology adoption. As discussed before, the PU (perceived usefulness) of gamified learning tools can be highly related to their effectiveness, so factors influencing effectiveness also provide insightful information for the technology acceptance intention.

Bearing the aforementioned concerns in mind, literature related to gamification's effectiveness was also included in the current thesis, as detailed in the current chapter.

CHAPTER THREE: LITERATURE REVIEW-2

Note:

This chapter reports the literatures about the effectiveness of educational gamification.

As mentioned before, when first cut into the gamification area, researchers tend to focus on the impact of gamification on learning (Baydas & Cicek, 2019). This author failed to continue studying the topic due to practical limitations, so the research focus has changed to technology acceptance intention.

Though seemed irrelevant with technology acceptance intention, this chapter was included too due to its value in providing insights and inspirations for future studies.

3.1. The effectiveness of gamified learning

As supported by Hakulinen, Auvinen, and Korhonen (2013) and Denny (2013), though gamification has been gaining increasingly more attention, "there is little empirical evidence to support their effectiveness" on engaging students or improving academic performances (Denny, 2013, p. 763). As a result, the effectiveness of gamification for learning "remains controversial" (Silpasuwanchai, Ma, Shigemasu, & Ren, 2016, p. 459).

Previous empirical studies tested the effectiveness of educational gamification, and reported varied results (positive, negative, mixed, partial, slight, and no impact), as detailed below.

3.2.1. Gamified learning with positive effects

A comparatively large number of empirical studies reported positive effects of gamification on learning.

Thom, Millen, and DiMicco (2012) proposed to assess the effectiveness of gamification in a social networking service (SNS) by removing the points system including points and badges. The results suggest that the removal of the points system has a "significant negative impact" on user activity (Thom et al., 2012, p. 1069). Namely, the points system helps improve user activity. Similarly, Hamari (2017) investigated the effectiveness of an individual mechanic 'badges' on a trading website *Sharetribe*. The results show that users in the gamification group are "significantly" more likely to take part in online activities (p. 469). J. C.-Y. Sun and Hsieh (2018) attempted to exam whether a gamified interactive response system enhances EFL (Egnlish-as-a-foreign-lagnauge) learners' motivation, engagement and attention. The results showed that the gamified learning tool effectively stimulates interest, intrinsic motivation, attention and emotional engagement, while there were no significant differences in behavioural and cognitive engagement. The research results of other scholars also support the positive effect of gamified learning (Fernandes et al., 2012; González et al., 2016; Hamari, 2013; Ibanez, Di-Serio, & Delgado-Kloos, 2014; Osipov, Volinsky, Nikulchev, & Prasikova, 2016; Papastergiou, 2009; Rai & Beck, 2017; Riva, Camerini, Allam, & Schulz, 2014; Stanculescu, Bozzon, Sips, & Houben, 2016; Su & Cheng, 2013).

Kim and Lee (2015) proposed a dynamical model of educational effectiveness for the gamification of learning, which compares the effectiveness of gamified learning and traditional learning. According to Kim and Lee (2015), gamified learning or game-based learning starts with lower educational effectiveness than the traditional way due to hindrances (e.g., adaptation time), and then "increases rapidly and exceeds the educational effectiveness of traditional learning" (p. 8489). Eventually, gamified learning helps achieve better learning outcomes than the traditional way.

It is noteworthy that some empirical studies that confirmed the positive effect of gamified learning also imply that gamified learning should be implemented with more considerations. For example, Denny (2013) surveyed to explore whether a gamification element 'badge' engages students in online learning settings. The overall data suggests a positive result, however, as stated by Denny (2013), "not all students are motivated… to the same extent" (p. 770). In other words, a certain group of students are not engaged in gamified learning though it is not revealed in the data.

3.2.2. Gamified learning with negative effects

For the studies reporting negative impact of gamified learning, one main concern is that gamified learning may harm students' intrinsic motivation.

Hanus and Fox (2015) believed that one central goal of education is to get students intrinsically motivated, which occurs when the desire to learn comes from within the students (Deci & Ryan, 2000). However, the gamification mechanisms available in gamified learning tools tend to provide extrinsic incentives or rewards, such as badges, points and competition. According to Hanus and Fox (2015), providing "tangible" and "expected" rewards to learners who are already interested in the current task "may cause them to shift motivations from intrinsic (e.g., because they wanted to) to extrinsic (e.g., because they want to earn a reward)". This opinion was supported by various scholars such as Deci and Ryan (2000), Lepper, Greene, and Nisbett (1973) and Tang and Hall (1995).

Hanus and Fox (2015) tested their hypotheses on 71 students across two courses during a 16-week semester in educational settings. The results showed that the students in the gamified group were less motivated, satisfied or empowered than students in the controlled group. At the same time, students in the gamified course gained lower scores in the final examination.

The study of Y. Chen, Burton, Mihaela, and Whittinghill (2015) revealed similar results, as users felt disengaged and confused about the customised gamification system. As analysed by Y. Chen et al. (2015), the failure to engage students was largely caused by technical and logistical implementation issues. The ease of use is crucial for perceived effective gamified learning tools, as well as the gamification design.

3.2.3. Gamified learning with mixed or partial effects

Consistent with the conclusion of Adukaite et al. (2017), there is an emerging body of research suggesting that gamified learning is not always beneficial or effective. Besides those who prove gamification is not significantly effective from the general educational perspective, some researchers aim to dig into the detailed differences of the impact, such as the difference on quantitative results and qualitative opinions (DomíNguez et al., 2013; Ibanez et al., 2014), the differences on overall scores and specific areas (DomíNguez et al., 2013), the differences on varied individuals (Denny, 2013), and the differences on academic outputs and behaviour engagement (De-Marcos, Domínguez, Saenz-de-Navarrete, & Pagés, 2014; Hakulinen et al., 2013).

The study of Ibanez et al. (2014) reported mixed results on the effectiveness of gamified learning activities targeted at the learning of the C-programming language. Besides analysing students' login information from log files, the

authors requested the 22 participants to complete five questionnaires, which contained four open-ended questions focusing on students' reasons to get involves and students' main focus during the learning activity. The login information and questionnaire results suggest general "positive" effects on student engagement toward gamified learning activities and a "moderate" improvement in learning outcomes (p. 291); though, the open-ended question responses present a portion of negative answers. Participants leave comments such as "There were no new things to explore", "I did not feel engaged" and "I feel bored" (Ibanez et al., 2014, pp. 297, 299).

J. C.-Y. Sun and Hsieh (2018) attempted to exam whether a gamified interactive response system enhances EFL learners' motivation, engagement and attention. The results showed that the gamified learning tool effectively stimulates interest, intrinsic motivation, attention and emotional engagement, while there were no significant differences in behavioural and cognitive engagement. The underlying reasons resulting in this result need further exploration.

Reynolds and Taylor (2020) reported mixed results in examining the impact of a gamified EFL tool (Kahoot!) on user experiences and vocabulary knowledge. Weaknesses of the involved gamified learning tool were investigated, among which an important one was the incompatibility of gamification to "some students" (p. 81): while a majority of students presented positive reactions, several others were demotivated in the experiment. Besides, even though the instructors generally held a positive attitude towards the gamified learning tool, they were "not entirely convinced" of its efficacy related to instructions (Reynolds & Taylor, 2020, p. 79). C.-M. Chen et al. (2020) investigated the impact of a web-based collaborative reading annotation system on reading performance. The results showed that though the experimental group generated more annotations, there was no difference in reading comprehension performance. It can be inferred that behavioural investment does not equivalent to improvements in academic achievement.

C.-M. Chen et al. (2020) investigated the impact of a web-based collaborative reading annotation system on reading performance. The results showed that though the experimental group generated more annotations, there was no difference in reading comprehension performance. It can be inferred that behavioural investment does not equivalent to improvements in academic achievement.

Similarly, the qualitative comments in the study of DomíNguez et al. (2013) indicate a negative impact of gamification techniques on students' engagement. According to DomíNguez et al. (2013), good results do not apply to every individual. The system was "not motivating enough", not fun, and even "discouraging" (p. 391). DomíNguez et al. (2013) designed a gamification plugin containing PBL for a well-known e-learning platform and tested students' achievement and attitude. They further analysed the data and summarised a specific conclusion that gamification partially impacts on students as students in gamified group "got better scores in practical assignments and in overall score", but they "performed poorly on written assignments and participated less on class activities, although their initial motivation was higher" (DomíNguez et al., 2013, p. 380).

The same with DomíNguez et al. (2013), other empirical studies reported that gamified learning does not impact every aspect of learning: gamified and traditional learning have different functions and impacts. For example, De-Marcos et al. (2014) found that students in the gamified learning group performed better on practical assignments but worse in knowledge assessments. At the same time, even though students in the gamified group are positive in attitude, they gain comparatively low scores and rates. With similar results, Hakulinen et al. (2013) attempted to differentiate behaviour engagement from general engagement, as the results of their study indicate that badges affect the "behaviour of students" but not the academic "grading" of the students (p. 47).

3.2.4. Gamified learning with slight or no effects

Surprisingly, there are also studies reporting that gamification brings slight or even no effect on learning.

Both DomíNguez et al. (2013) and De-Marcos et al. (2014) found that in a final examination for the assessment of conceptual knowledge, students in the gamified gained similar academic scores with students in the controlled group. DomíNguez et al. (2013) concluded that the "cognitive impact of gamification over students is not very significant" (p. 391).

Dindar, Ren, and Järvenoja (2020) conducted an experimental study on the effects of gamified cooperation elements and gamified competition elements on English vocabulary learning, then concluded that there were no differences regards to behavioural engagement (task effort), academic achievement, and motivation. However, it is notable that the study did not involve an un-gamified control group, so the effectiveness result should be interpreted with caution.

In examining other indicators, such as academic efficacy, there was also a study reporting no difference (Rachels & Rockinson-Szapkiw, 2018).

Generally speaking, previous studies reported significantly different results regarding the effectiveness of gamification in education. The ultimate goal of the current study is not only to form a state-of-art of the effectiveness study results but also to sort out the methods to improve the effectiveness of gamification in education. To achieve so, this author needs to dig deeper into how effectiveness can be measured and what factors contribute to varied effectiveness of gamification studies, as discussed in the following sections.

3.3. Measurement of the effectiveness of gamified learning in previous studies

Though researchers propose that there are many benefits of using gamified learning tools in education, one of the most common goals is improving students' engagement in learning (Landers & Armstrong, 2017), which "has been proved a positively correlated with outcomes of student success" such as satisfaction, academic achievement and self-efficacy (Ibanez et al., 2014, p. 292). As summarised by Ibanez et al. (2014), among the various conceptualizations of engagement, a frequently-used theory is the three-component engagement: behavioural engagement, affective engagement and cognitive engagement (see Table 26).

Behaviour engagement. Behaviour engagement is the task engagement we can observe based on users' behaviours. In general education, behavioural engagement relates to time on task, homework completion, attendance, classroom activity participation (Appleton, Christenson, & Furlong, 2008; Appleton, Christenson, Kim, & Reschly, 2006; Ge & Ifenthaler, 2018; Ibanez et al., 2014). In gamification that involves e-learning, behavioural engagement relates to more calculation-convenient measures such as time on task and number of attempts (Silpasuwanchai et al., 2016).

Affective engagement. Affective engagement is the task engagement linked to emotional reactions. As explained by Boykin and Noguera (2011), "the greater the student's interest level, positive affect, positive attitude, positive value held, curiosity, and task absorption (and the less the anxiety, sadness, stress, and boredom), the greater the affective engagement" (p. 43).

Cognitive engagement. Rotgans and Schmidt (2011) define cognitive engagement as a psychological state in which students are willing to put time and effort to "truly understand a topic" and in which students "persist studying over a long period of time" (p. 465).

The definition of cognitive engagement is complex with the definition of behaviour engagement as Appleton et al. (2006) state that cognitive engagement has been traditionally measured by homework completion, class attendance and interactions with the teachers, which are also "indicators of behavioral engagement" (p. 429). We agree with the further discussion of Appleton et al. (2006), which explains that cognitive engagement includes "less observable" and "more internal" indicators, such as self-regulation, value of learning, personal goals and learning autonomy (p. 429), which was later simplified by Silpasuwanchai et al. (2016) into "learners' psychological investment in learning" (p. 462). Previous studies mainly assess the effectiveness of gamified learning tools or services by measuring students' behavioural and affective engagement.

Researchers measure behavioural engagement with indicators such as time of use (Osipov et al., 2016; Silpasuwanchai et al., 2016; Stanculescu et al., 2016), the number of answers or connections (Denny, 2013; Hamari, 2017; Osipov et al., 2016; Stanculescu et al., 2016), earned points (Osipov et al., 2016), or simply the participation (DomíNguez et al., 2013; Hamari, 2013; Thom et al., 2012). A large number of researchers collect behavioural engagement data via observation or from the log files in the gamified learning system.

Researchers also measured affective engagement with factors such as achievement (Fleming, 2001; Su & Cheng, 2013), attitude (De-Marcos et al., 2014; DomíNguez et al., 2013; González et al., 2016; Rai & Beck, 2017), efficacy (Buckley & Doyle, 2017), empowerment (Hanus & Fox, 2015; Riva et al., 2014), enjoyment (Silpasuwanchai et al., 2016), motivation (Fernandes et al., 2012; Hanus & Fox, 2015), satisfaction (Fernandes et al., 2012; Hanus & Fox, 2015) and the sense of encouragement (Y. Chen et al., 2015). There are also other indicators such as the easiness to understand, easiness to play, perceived usefulness, domination, immersion, self-determination and amusement (Fernandes et al., 2012; Fleming, 2001; Koivisto & Hamari, 2014; Riva et al., 2014). The main measures of these indicators are self-report questionnaires or interviews (Hew, Huang, Chu, & Chiu, 2016).

Notably, some researchers connect academic performance with cognitive engagement, therefore, they assess the marks in exams (De-Marcos et al., 2014; DomíNguez et al., 2013; Hanus & Fox, 2015), number of correct responses or number of recalled items (Silpasuwanchai et al., 2016). Academic assessment indeed involves cognitive activities, however, we believe that cognitive involvement is not equivalent to cognitive engagement. If studies assessing academic performance as the indicators of cognitive engagement are excluded, we find that cognitive engagement remains under-investigated. As stated by Silpasuwanchai et al. (2016), "it remains inconclusive" whether gamification can help students "think deeply" and "reflect" (p. 462).

Table 17. The effectiveness of gamified learning practices

#	Domain	Indicator	Instrument
1	Behavioural engagement	Time input, number of interactions, earned points, online contributions, task completion rate, surveyed behavioural engagement	Log analysis, self-report survey, scale, observation
2	Affective engagement	Confidence, immersion experience, anxiety, curiosity, attitude, interest, enjoyment	Open-ended survey report, scale, interview
		General learning experience (fun, immersion, usability experience)	
		Surveyed emotional engagement	
3	Cognitive engagement	Surveyed cognitive engagement	Scale
4	Academic achievement	Pre-and-post test	Test, quiz
5	Others	Motivation, academic efficacy, attention, usability, enthusiasm	Observation, survey, self-report journal

3.4. Factors influencing the effectiveness results of gamified learning tools

3.4.1. Biases or barriers impeding the effectiveness of gamified learning

Previous studies revealed that the development of gamification in education comes across with challenges, which include but are not limited to seven aspects: inappropriate or superficial use of gamification techniques**, a**mbiguous or broad definition of the term**,** poor integration of game features and pedagogical content**,** difficulties of designing adaptive gamification tools for education/technical limitations, heavy reliance on computer science skills, methodological limitations, and research biases caused by the experimental setting (see Table 27).

Table 18. Barriers of effectiveness studies in gamification studies

Barriers of gamification studies
1) Inappropriate or superficial use of gamification techniques.
2) Ambiguous or broad definition of the term.
3) The poor integration of game features and pedagogical content.
4) The difficulty of designing adaptive gamification tools for education/Technical limitations.
5) Heavy reliance on computer science skills.
6) Methodological limitations.
7) Research biases caused by the experimental setting.
8) Data interpretation biases.

5.4.1.1. Inappropriate or superficial use of gamification techniques

Inappropriate or superficial use of gamification techniques was reported as one factor influencing the effectiveness results. Gamification has been defined as "using gamification mechanisms or game elements to non-game contexts", which indicates that gamification is the "use of design (rather than game-based technology or other game-related practices) elements (rather than fully developed games)" (Deterding, Sicart, et al., 2011, p. 5). However, for a long time in practice, the tools for gamification studies are customised video games for teaching purposes. There are several concerns of using video games in education, such as their poor integration with educational content, irrelevant tasks, homogenous content, waste of classroom time and possibilities of getting students addicted (Kim & Lee, 2015; B. Morris et al., 2013; Van Rosmalen et al., 2015). As a result, the video games and school "each make(s) the other worse" (Lee & Hammer, 2011, p. 4). Later on, researchers begin to focus on educational websites or mobile applications with the use of game elements, which is aligned with the gamification definition. However, though there are dozens of game elements or mechanisms, the most explored are limited to points, badges and leaderboard only. In this way, gamification is more likely to be "pointsfication" only.

5.4.1.2. Misinterpretation of the gamification concept

Previous studies presented a tendency to misinterpret the gamification concept or failures in achieving meaningful gamification. Some other researchers also applied video games in their studies even though the related keyword was "gamification". For example, Tamtama, Suryanto, and Suyoto (2020) specified "The *gamification* method uses a mobile phone-based application with video *games*" (p. 150), and Palomo-Duarte, Berns, Dodero, and Cejas (2014) used the word "game" to refer to the involved tool and "player" to refer to the participant (p. 382).

At the same time, Chou (2019) stressed that gamification is not the simple adding of game elements; instead, it should be a playful process that motivates users to achieve the expected behaviours by triggering their human desires (e.g. gain social influence, avoid failure). However, some previous studies layered game elements directly onto the conventional learning activities (Angelia, 2019; Heryadi & Muliamin, 2016; Hong, Hwang, Liu, & Tai, 2020), which lacks signs of triggering meaningful gamification experience: if the learners do not know the answers to the questions, they cannot gain any positive feedback or experience. A typical addition of simple game elements was named "pointsification", which refers to the simple adding of points-related elements in non-game contexts, including points, badges, experience score (XP), leaderboard, etc. (Crow & Parsons, 2015). Though of no flaw according to the

definition (the use of game design elements in non-game context), it has been criticised for being too simple to provide a fun experience, neither triggering meaningful gamification (Hong et al., 2020). Pointsficaiton, the example of suboptimal use of gamification concept, is an obvious tendency constraining the potential of gamified learning tools.

Some researchers applied video games in their studies even though the related keyword was "gamification". For example, Tamtama et al. (2020) specified "The gamification method uses a mobile phone-based application with video games" (p. 150), and Palomo-Duarte et al. (2014) used the word "game" to refer to the involved tool and "player" to refer to the participant (p. 382); Y. Chen et al. (2015) conducted focused groups and interviews to explore students' experience with a virtual currency system named Cogent. Authors of that article believed that Cogent is a system with meaningful gamification designs, in which students can "create a successful business and accumulate individual monetary value" (p. 41). Participants report that they were confused. One participant describes the system as the "elephant in the room" in which everyone "acknowledge(s) it, but no one consistently explained or understood it" (p. 42).

A paragraph stated by Lee and Hammer (2011) explained why the above-mentioned phenomenon exists. According to Lee and Hammer (2011), the traditional schools are already using game-like elements: the grades students receive for their assignments can be translated to "points" or "badges", the upgrade to the next level can be translated to "level up", and the rewards or punishments for the desired/undesired behaviours can be regarded as "common currency" or "a reward system". Similarly, the influential scholar McGonigal (2011) paints a blueprint in which life and work can be regarded as motivating games. Basically, these works use gamification concepts to describe or interpret the traditional teaching methods or tools. It is worth noticing that the key of gamification study is not whether there are game elements/mechanisms in the design or how to interpret the world from the gamification perspective; instead, the key is whether and how we provide new gamified experience.

5.4.1.3. Poor integration of game features and pedagogical content

Of course, studies are working on gamified learning experiences; and of course, there still are difficulties. As a result, there is normally a poor integration of game features and pedagogical content. According to Lee and Hammer (2011), one key feature of playfulness (or game experience) is freedom: the freedom "to experiment, to fail, to explore multiple identities, (and) to control one's own investment and experience" (p. 4). However, teaching is kind of making play mandatory by setting rules and goals. By putting these two things together, the expected "chocolate and peanut butter" turn out to be "chocolate-covered broccoli" (Lee & Hammer, 2011, p. 4). Simply "dressing up" an activity as a gamified one is far from being intrinsically engaging (Lieberoth, 2015). This statement explains why gamification normally integrates poorly with pedagogical content and why some gamified attempts ended up with negative outcomes.

5.4.1.4. The use of different or unsuitable game elements in varied context

Another theme that emerged from the content analysis was the use of different or unsuitable game elements in varied contexts. Werbach (2014) highlighted the challenges of inherent comparability of empirical studies: even though Deterding, Dixon, et al. (2011) defined gamification as the use of "game design elements" in non-game contexts (p. 10), there is "no universal list of game elements" (Werbach, 2014, p. 267); therefore, different studies would test *different* game elements with *different* definitions into *different* educational contexts, which makes the comparison like comparing apples with bananas. Researchers also proposed that certain game elements might not suitable for pedagogical activities: Abrams and Walsh (2014) complained that some game design characteristics (e.g. shooting)

distracted learners from vocabulary learning; Kurniawan, Sitohang, and Rukmono (2019) found three elements only brought limited effects to language learning (rules, story, and avatar). In summary, the use of different game elements raises comparability issues, and the use of unsuitable game elements is possible to bring mixing or even opposite results.

5.4.1.5. Difficulty of designing adaptive gamification tools for education/technical limitations

Technical limitations impeded the potential of gamified learning tools, too. Rego (2015) criticised that the program developers do not always provide learners with a tool that meet needs, either due to a lack of understanding towards students' needs or due to technical limitations in design. Similarly, it was commented that a gamified tool can also be designed too simple to trigger the real gamification experience in empirical studies (B. J. Morris, Dragovich, Todaro, Balci, & Dalton, 2019), or too immature in the development of dedicated language processing technologies (Purgina et al., 2020).

The lack of adaptation is one main criticism of gamification caused by immature technical supports (Buckley & Doyle, 2017; Landers & Armstrong, 2017). Most current gamified learning tools serve for average students, which are hardly engaging for high-performing or left-behind students. At the same time, a limited number of gamified learning tools can provide learners with tasks "just barely out of reach" to trigger flow status (Cain & Piascik, 2015, p. 3).

There are high entry barriers to design gamified learning tools (Cain & Piascik, 2015; Y. Chen et al., 2015). Educators with little or no training in computer science are very possible to create "bad games" "with little or no effect on learning" (Cain & Piascik, 2015, p. 4).

There are high entry barriers to design gamified learning tools (Cain & Piascik, 2015; Y. Chen et al., 2015). Educators with little or no training in computer science are very possible to create "bad games" "with little or no effect on learning" (Cain & Piascik, 2015, p. 4).

5.4.1.6. Suboptimal measurement

There was an issue of suboptimal measurement: what has been measured might not be what the researchers wish to obtain. A typical example is a large number of studies that measured participants' behavioural investments to test whether gamified learning tools are effective, while the "effectiveness" was connected with cognitive activities or academic achievement. Unfortunately, results "may imply a gap" between behavioural engagement and the final learning outcomes (C.-M. Chen et al., 2020, p. 13). B. J. Morris et al. (2019) also suggested that behaviour engagement (e.g., the number of earned badges) is a representation of effort input rather than cognitive improvement in learning activities, while the two concepts have not been proofed equivalent or identical.

Researchers proposed that more measurements should be involved in assessing the effectiveness of gamified learning tools, such as users' psychological processes (Dindar et al., 2020), users' psychological characteristics (C.-M. Chen et al., 2020), fun learning experience (Heryadi & Muliamin, 2016), and benefits of kinesthetic activities (Reynolds & Taylor, 2020).

5.4.1.7. Methodological limitations

A considerable number of researchers criticise that there are methodological limitations in most existing studies of gamification in education, including short experiment timeframe, small sample size, absence of control groups, the

involvement of one single gamification mechanism and the lack of both qualitative and quantitative analysis (Y. Chen et al., 2015; De-Marcos et al., 2014; Seaborn & Fels, 2015; Stanculescu et al., 2016). We do not consider these methodological limitations are as fatal as the aforementioned ones.

Researchers reflected methodological limitations in empirical studies, such as the lack of a control group (Dindar et al., 2020), the lack of long-term investigation (Y.-L. Chen & Hsu, 2020; Kurniawan et al., 2019), the lack of strict control of variables (B. J. Morris et al., 2019), and small sample size (Dindar et al., 2020; Licorish, Owen, Daniel, & George, 2018; B. J. Morris et al., 2019). What's more, Dindar et al. (2020) mentioned ceiling effects in academic achievement tests: if the test is too simple, participants in both the experiential and the controlled group will be scored high, which narrows the differences. Researchers criticised improper assessment designs; for example, Dindar et al. (2020) pointed out they measured learning achievement with a vocabulary test rather than a comprehensive one covering other language learning abilities such as reading and writing; similarly, Y. W. Lam, K. Hew, and K. F. Chiu (2018) retrospect the research process and inferred a possible suboptimal research design: their assessment emphasised measuring one aspect of language skills, with another important one unassessed.

5.4.1.8. Research biases caused by the experimental setting

Similar to but different from methodological limitations, the experimental setting contributed as a factor bringing research biases. As figured out previously, there is a lack of longitudinal study investigating the long-term impact of gamification on foreign language learning, so the experimental group might generate abnormal or deceptive data due to the novelty effect (C.-M. Chen, Li, & Chen, 2018; Kurniawan et al., 2019; Silpasuwanchai et al., 2016). Even though the experimental group reported significant effects of the gamified learning tools, there will be a lack of evidence showing the impact can be maintained (Rueckert et al., 2020). Other problems were reported, too, including the unauthentic experience provided in experimental settings (Gafni, Achituv, & Rahmani, 2017; B. J. Morris et al., 2019) and the involuntariness of participants in learning the target language (even though they gave consent to take part in the experiment) (Loewen et al., 2019).

5.4.1.9. Data interpretation biases

Besides, this author noticed data interpretation biases due to varied perspectives. For example, Murad, Wang, Turnbull, and Wang (2018) surveyed participants' user experience by asking them to select "agree", "neutral" or "disagree" to questions. When asked whether the tool was enjoyable or easy to use, none of the participants ticked "disagree" (0%); notably, when asked whether the gamified tool was helpful for vocabulary acquisition or pronunciation improvement, there was a significant number of participants selecting "neutral" or "disagree" (46.6% and 33.4% respectively). It can be inferred that the provided gamified learning tool can be emotionally engaging but comparatively less sufficient in bringing academic benefits. However, the authors concluded that the selected gamified foreign-language-learning (FLL) tool "helped them to improve pronunciation" while "fewer reported … (it) helped with vocabulary learning". The interpretation is true, but it did not highlight the 33.4% of participants who were not positive about the tool's impact on pronunciation, which brings a different conclusion to the effectiveness of the involved gamified learning tool.

3.4.2. Moderating factors impeding the effectiveness of gamified learning

The above-mentioned issues are barriers of gamification studies in achieving the optimal result, which tend to be limitations or biases. Previous findings also revealed factors that influence the effectiveness results, which tend to be mediator factors. The factors including demographic differences, game experience, individual differences (e.g., personality traits or learning styles), and resistance to new technologies (see Table 28).

Demographic differences were frequently mentioned, which include gender and age (Yee, 2006): they have been shown to be the main contributing factors affecting users' perception of digital games and in some cases, gamification shares mutual features with digital games (Silpasuwanchai et al., 2016). The results indicated that age does not bring significant differences in the effectiveness of gamified learning (Koivisto & Hamari, 2014), while gender and game experience do (Koivisto & Hamari, 2014; Silpasuwanchai et al., 2016). On average, students of different gender perceive the benefits of gamified learning differently, and students with more game experience are more possible to benefits more from the gamification design (Koivisto & Hamari, 2014; Landers & Armstrong, 2017).

Game experience was reported as one factor deciding whether the user would be positive about accepting gamified learning tools.

Learning style differences or personality traits is an important research branch in studying the factors influencing the effectiveness of gamification in education, as gamified learning does not impact every student. Researchers study the impact of individual differences from various perspectives including learning style and personality traits (Buckley & Doyle, 2017), player types (Barata, Gama, Jorge, & Gonçalves, 2017; DomíNguez et al., 2013), personal preferences (Denny, 2013), player motivation (Yee, 2006), etc.

If the study is an empirical one involving a specific gamified learning tool, the resistance from new technologies is a frequent-mentioned factor affecting the result. The gamified instruction may lead to poorer outcomes if students are not comfortable with the new technology (Landers & Armstrong, 2017) or students feel the tool is difficult to use (DomíNguez et al., 2013). The situation is even more severe when students holding poor attitudes towards game-like tools (unfortunately some are) for training purposes.

Table 19. Factors influencing the effectiveness of gamified learning

Moderating factors
1) Demographic difference (e.g. age, gender).
2) Game experience.
3) Personality traits or styles.
4) Resistance to new technologies.

Barriers of gamification studies

1) Inappropriate or superficial use of gamification techniques
2) Misinterpretation of the gamification concept
3) The poor integration of game features and pedagogical content
4) The use of different or unsuitable game elements in varied context;
5) The difficulty of designing adaptive gamification tools for education/Technical limitations
6) Heavy reliance on computer science skills
7) Methodological limitations
8) Research biases caused by the experimental setting
9) Data interpretation biases

CHAPTER 4

Methodology

04

Methodology

This study aimed to develop and validate a scale measuring secondary school teachers' acceptance intention to gamified learning tools in formal education contexts, using a combination of qualitative and quantitative methods.

This section introduces the method selection, participants, research instrument, data collection, data analysis, and reliability and validity

CHAPTER FOUR: METHODOLOGY

4.1. Introduction: Aims and objectives

This study aimed to develop and validate a scale measuring secondary school teachers' acceptance intention to gamified learning tools in formal education contexts, using a combination of qualitative and quantitative methods.

Notably, changes were applied during the research process. In the first round of data collection, findings revealed that two main constructs are contributing to teachers' acceptance intention to gamified learning tools (PU and PEoU); however, both of the two constructs have varied connotations in the gamification contexts. To be specific, in the original framework, PU refers to the perceived usefulness that relates to achievements, but the PU in the current research context contains the perceived usefulness for both academic achievements and benefits of the enjoyable process; at the same time, while the PEoU was traditionally referring the perceived ease of use of the technology or the digital devices, PEoU of gamified learning tools in secondary-school contexts is connected with not only the cumbersomeness in using digital devices but also the effort-demandingness of dealing with correspondent issues such as classroom discipline management. Therefore, the specific connotations of PU and PEoU in new contexts need further exploration. Accordingly, a second round of survey construction and validation was followed, which aimed to develop two scales that measure PU and PEoU of gamified learning tools respectively.

The current study covered four main aspects:

1) Summarise factors that contribute to users' acceptance intention to new technology through review of the literature.
2) Establish a theoretical framework based on the acceptance-intention-factors, which applies to gamification contexts.
3) Construct a scale measure secondary school teachers' acceptance intention to gamified learning tools. The scale is abbreviated as TAI-GLT.
4) Validate the scale. Use the typical scale evaluation techniques proposed by DeVellis (2003) to test and improve the reliability and validity of the developed scale. If needed, construct a shorter version one with fewer survey items.

(After obtained the results of the first-round scale development findings, it was found that PU and PEoU have more than one connotation in the current research context.)

5) Analyse what factors are contributing to secondary school teachers' PU towards gamified learning tools (GLT).
6) Design a scale based on the analysed factors and the interview responses. The scale is abbreviated as PU-GLT.
7) Analyse what factors are contributing to secondary school teachers' PEoU towards gamified learning tools (GLT) in formal education.
8) Design a scale based on the analysed factors and the interview responses. The scale is abbreviated as PEoU-GLT.
9) Validate the two scales PU-GLT and PEoU-GLT.

Consequently, the main outputs of the current study were one scale measuring teachers' acceptance intention towards gamified learning tools, and two sub-scales measuring their perceived usefulness and perceived ease of use toward gamified learning tools.

4.2. Overall method selection: The exploratory sequential mixed method

This is a cross-sectional study that employed the exploratory sequential mixed method, which combines the collection and analysis of both qualitative and quantitative data in a sequence of phases (Creswell & Clark, 2017). In a typical exploratory sequential mixed method study, results of the qualitative data analysis function as a basis to construct the quantitative phase, which are particularly significant for scale development studies (Baydas & Cicek, 2019; Mihas, 2019; Moral-Bofill et al., 2020).

In the current study, this author collected qualitative data via interviews and quantitative data via survey. The interview responses provided directions for survey item generation, survey item analysis, and content validity establishment, while the quantitative survey responses were analysed following the typical scale evaluation procedures.

In most of the extant scale-development studies, the qualitative data collection process was skipped: the researchers constructed the scale based on researchers' experience or based on literature review. This author insists that qualitative data collection is appropriate and essential for scale development. The qualitative data, interview responses in the current study, help explore the dynamic, holistic and individual perceptions of the target group. It also captures data which are meaningful to the target group.

Flick (2018) suggested that qualitative and quantitative can be integrated, sequenced or triangulated in one study.

Quantitative research begins with assumptions and the use of theoretical frameworks that inform the study of research problems addressing the meaning of the target groups related to a social or human problem (Creswell & Poth, 2016). Qualitative research contains characteristics such as being conducted in a natural setting, relying on the researcher as the key instrument in data collection, involving the use of multiple methods, involving complex reasoning, focusing on participants' multiple perspectives and meaning, being situated within the context or setting of participants, involving an emergent and evolving design, being reflective and interpretive of researcher's background influence, and presenting a holistic and complex picture (Creswell & Poth, 2016, p. 45).

4.3. Qualitative method selection: Grounded theory

Grounded theory is a systematic method of data analysis for theory development (Lyons & Coyle, 2007), which is used when researchers wish to "move beyond description" of participants' individual stories and to "generate or discover a theory" (Creswell & Poth, 2016, p. 82). The grounded theory helps establish a unified theoretical explanation for a process or an action based on the data collected from participants.

According to Vollstedt and Rezat (2019) and Birks and Mills (2015), grounded theory is suitable when one or more of the three phenomena was met:

1) Little is known about what is being studied.
2) The relationships among the concepts are not elaborated enough.
3) The relevance of the concepts and their relationships has not been corroborated for the population of the context.

This thesis meets the aforementioned three criteria. Though the acceptance intention of educational technologies has been widely studied, studies about gamified learning or gamified learning tools are comparatively scarce. What's more, the specific factors contributing to users' acceptance intention of gamified learning tools remain understudied, as well as the relationships among these factors.

One characteristic of grounded theory analysis is that data collection, data analysis, and theory development "are not successive steps" but are "intertwined and interdependent" (Vollstedt & Rezat, 2019, p. 85).

As stated by Flick (2018), it is a dynamic process in analysing qualitative data with the use of grounded theory: the researcher will experience "a spiral of cycles of data collection, coding, analysis, writing, design, theoretical categorization and data collection" (Flick, 2018, p. 449), namely the theory establishment process based upon categories developed from ongoing data analysis. The sample size is determined by the completeness of the theories: more samples should be recruited if there is a need for more demographic representatives, and the data collected can be halted if no additional information is needed for the framework (Flick, 2018).

The same with the analysis-and-participant-recruitment process, the coding process is also a dynamic spiral cycle that involves dynamic shifts between open coding, selective coding and theoretical coding. Coding is a process of "breaking data down into much smaller components and labelling those components" as codes (Sbaraini, Carter, Evans, & Blinkhorn, 2011, p. 3), while codes refer to "a form of shorthand that researchers repeatedly use to identify conceptual reoccurrences and similarities in the patterns in the data" (Birks & Mills, 2015, p. 89). The coding processes were described with the use of different languages, such as open coding, selective coding, axial coding, initial coding, and theoretical coding. The coding process for the current thesis was a combination of multiple rounds of open coding, selective coding, and theoretical coding (Table 17).

Open coding is usually the first step in grounded theory analyses, which focuses on the conceptualization and categorization of phenomena through an intensive analysis of the data (Vollstedt & Rezat, 2019). Open coding aims to "identify incidents in the data that appear to indicate one or more concepts and labelling these using one or two salient words per code" (Holton & Walsh, 2016, p. 81). Open coding is essential because it transforms ethnographic descriptions to themes applicable for further analysis or comparison. In open coding, qualitative data with similar concepts can be labelled with the same code, and one piece of information can be coded for more than one concept (Holton & Walsh, 2016; Vollstedt & Rezat, 2019). At this stage, the interview texts were analysed phrase-by-phrase and given conceptual names. The interview responses were repeatedly reviewed for multi times for enhanced reliability.

Selective coding is the process of choosing one category to be the core category, and relating the similar or highly relevant categories to the core one. In other words, it is a deductive process that aims to relate the codes generated from the open coding process to "the core category" to "produce a parsimonious theory" (Holton & Walsh, 2016, p. 84). According to Strauss and Corbin (1990), each code can be regarded as a *concept*, and the concepts of higher order can be named as *category*. A category is the combination of concepts (codes) describing different aspects of one phenomenon. By putting the open-coded data into categories, the selective coding process helps trim the data into more parsimonious results, which eventually leads to the generation of a new theory.

Theoretical coding is followed after selective coding. Theoretical coding, also named axial coding (Vollstedt & Rezat, 2019), refers to "the modeling of the relationships between and among the core categories and related concepts as a fully integrated theory" (Holton & Walsh, 2016, p. 86). It is the process of relating codes to each other via inductive and deductive thinking, which eventually fit the involved codes into a basic frame of generic relationships.

Table 20. Coding in grounded theory (Holton & Walsh, 2016, p. 87)

Types of coding		Description	Purpose
Substantive	Open	Capturing what is going on or conceptualizing what is happening.	To identify incidents in the data that appear to indicate one or more concepts.
	Selective	Coding around the core theory.	To identify properties and dimensions of the core category.
Theoretical (Axial)		Coding to model the relationships between and among the core category and related concepts.	To shape and integrate the theory.

4.4. Quantitative method selection: Scale development

Scale evaluation is important because "failure to design measuring instruments scientifically will result in poor-quality data and ultimately poor decision-making" (Robertson, 2017, p. 2). Conversely, according to Robertson (2017), a scientifically developed and standardised measuring instrument brings benefits as follows:

1) It provides new dimensions to surveys.
2) It is "cost-effective" to use once it developed (p. 3).
3) It is flexible to be used as an individual scale or a part of an existing survey.
4) It provides a quick way to assess gamified learning apps.
5) It allows "better cross-country and cross-survey comparison", which is of high compatibility (p. 3).
6) It is highly sustainable since it can be commented on or criticised by other researchers and therefore make adjustments.

It also provides other benefits such as:

7) It enables more people to make a quick decision in the selection of gamified learning tools. Schoolmasters, practising teachers and parents do not need to read complicated theories to reach a rational decision.
8) It enlists new dimensions that are important in designing gamified learning tools. Therefore, tool developers who are most probably with programming backgrounds can think from a pedagogy perspective at the first beginning of software/app design.
9) It enables website/app developers to evaluate their products and make modifications.

Due to the complexity in scale development and scale evaluation, the specific procedure and techniques would be detailed in the next chapter.

4.5. Participants

4.5.1. Participants of the qualitative study

To collect qualitative responses, this author conducted two rounds of data collection: a survey with open questions and a semi-structured interview. Both rounds of the data collection involved the same questions and the convenience sampling method. The survey study helps gather a large number of key concepts, and the interview was designed to reveal in-depth opinions and comments.

In data collection, 516 secondary school teachers in China responded to the survey with 347 responses considered valid. As shown in Table 18, 57% were female (n=199) and 43% were male (n=148), with the majority aged 26 to

35 (n=183, 53%). The interviews involved 14 secondary school teachers, with an equal number of male and female teachers (n=7). Seven participants were aged 20 to 25 (50%), followed by four teachers aged 36-45 (29%), two teachers aged 36 to 45 (14%) and one aged 46 or more (7%).

Table 21. Participant information (interview)

Category		Email survey (n=347)		Face-to-face interview (n=14)	
		Number	Proportion	Number	Proportion
Gender	Male	148	43%	7	50%
	Female	199	57%	7	50%
Age	20-25	84	24%	7	50%
	26-35	183	53%	2	14%
	36-45	60	17%	4	29%
	46 or more	20	6%	1	7%

4.5.2. Participants of the quantitative study

Sample size is an issue for scale evaluation since the involved factor analysis techniques will generate more stable results when used based on a large sample size than a smaller one (DeVellis, 2003). Comrey (1988) suggested that a sample size of 200 is reasonably good for ordinary factor-analytic work with 40 or fewer variables; the 5:1 or 10: 1 ratio of subject to items are two widely used sample size standards, namely at least 10 subjects or five subjects are required for every item presented in a scale (Anthoine, Moret, Regnault, Sébille, & Hardouin, 2014; Baggaley, 1983; Everitt, 1975; Gorsuch, 1983; Tinsley & Tinsley, 1987); Item communalities also play a role (DeVellis, 2003). Fabrigar and Wegener (2011) proposed that when the item commonalities are between .40 and .70 while there are at least three items measuring one construct, a sample of 200 is sufficient. Taking the aforementioned three standards into consideration, this study aimed to achieve a sample size larger than 200.

The survey study for scale version one recruited 512 teachers (see Table 19). This author designed two validation questions to filter the surveys that are of potential in decreasing the overall reliability. The two validation questions asked the participants to select “Disagree/2”. While the answers for other items tended to be “agree”, namely the option of “5” (Slightly agree), “6” (Agree), or “7” (Strongly agree), it is assumed that participants tend to choose 5, 6 or 7 as well if they do not read the survey item carefully. The surveys with abnormal answers were also excluded, such as the ones selected “Agree” for almost all the items. This author obtained 701 responses. Eventually, the valid responses for the data collection for scale version one was 512 with a valid response rate of 73%.

Scale version two was divided into two parts, with each one measure the perceived usefulness of gamified learning tools (PU-GLT) and the perceived ease of use of gamified learning tools (PEoU-GLT) respectively. The original PU-GLT and PEoU-GLT contained a large number of survey items (27 and 23 items). To minimise research biases caused by participant fatigue, this author collected the data of the two scales separately. As shown in Table 19, this author obtained 446 responses for the PU-GLT scale and 364 responses for the PEoU-GLT scale. After removing surveys that did not pass the two validation questions selected questions (“Please select ‘Disagree/2’ for validation purpose”), there were 263 and 239 valid surveys for the two scales. The valid response rate for the two scales were 59% and 66%.

Table 22. Participant information (the overall study)

Method	Data type	Participant	N (all)	N (valid)	Valid response rate
Email interview	Qualitative	Secondary-school teachers	516	347	67%
Face-to-face interview	Qualitative	Secondary-school teachers	14	14	100%
Expert interview	Qualitative	Experts	2	2	100%
Survey (TAI-GLT)	Quantitative	Secondary-school teachers	701	512	73%
Survey (PU-GLT)	Quantitative	Secondary-school teachers	446	263	59%
Survey (PEoU-GLT)	Quantitative	Secondary-school teachers	364	239	66%

4.6. Research instrument

Both the open-ended-question survey and face-to-face in-depth interview at the qualitative data collection stage involved the same interview questions.

As suggested by Seidman (2006), this author did not ask the exact question "what factors are influencing your gamification adoption intention" to avoid biased answered brought by suggestive questions. Since this study involved the inductive analysis in which the research results "bear little relation to the specific questions that were asked of the participants" (Braun & Clarke, 2006, p. 83), indirect questions are applicable. The three questions about the teachers' experience are:

1) the advantages and disadvantages of gamified learning.
2) the motive and barriers in implementing gamified learning.
3) the characteristics of good gamified learning tools.

There were three sets of surveys for the two rounds of survey studies, which were for the measurement of teachers' overall acceptance intention, teachers' PU of gamified learning tools, and teachers' PEoU of gamified learning tools respectively. The construction of survey items was based on analysis results of the qualitative interview responses, which will be detailed in the section of Findings. The specific survey items were as shown in the Appendices.

4.7. Data collection and ethics consideration

Both the qualitative and quantitative data were collected with the convenience sampling method.

To collect interview responses, this author contacted the principals of the target secondary school in China for perception, then sent email invitations to teachers. Teachers can either fill the open questions in an online survey anonymously or contact this author for a face-to-face interview. The survey study followed similar procedures. The expert selection for the expert interview was based on the participants' expertise.

Ethical issues were addressed before, during, and after the data collection, such as seeking university approval, avoiding deceiving participants, using appropriate security measures to store the data, avoiding siding with participants and disclosing only positive results, respecting participants' privacy, etc. The specific measures to address ethical issues were employed following the instructions listed by Creswell and Poth (2016, pp. 55-56).

4.8. Data analysis

4.8.1. Data analysis- Qualitative data

Code the survey responses. Before data analysis, this author divided the survey responses into segments. Even though the survey responses only reported key information, there were still participants reporting more than one concern in one survey. The 347 participants presented 418 statements regarding the factors contributing to their acceptance intention to gamified learning tools.

Code the interview responses. When dealing with the interview data, this author imported the transcribed responses into the software Nvivo, followed by a thorough inspection of each interview response in the entire data set. When identified a sentence or a paragraph that is possibly related to the contributing factors of gamification adoption intention, this author selected it as a new code. After getting all possible codes from the data set, this author contacted a second-rater to view the codes and categorise them into groups, which were later named *patterns*. Together with the second-rater, this author checked the patterns with the original dataset to make sure the patterns were inclusive, coherent, consistent, and distinctive.

The whole pattern-generating process does not involve theoretic frameworks or theories, following the grounded theory analysis guidelines by Birks and Mills (2015) and Holton and Walsh (2017) (see Section 3.3. Qualitative method selection: Grounded theory).

During the data analysis process, the checklist for rigorous thematic analysis (Braun & Clarke, 2006) was also valued. Grounded theory and thematic analysis are two varied qualitative-data-analysis approaches; however, some techniques of the thematic analysis can be used for grounded theory. Thematic analysis is a qualitative data analysis method for "identifying, analysing and reporting patterns (themes) within data" (Braun & Clarke, 2006, p. 79). The patterns or themes capture important information concerning the research questions, which make the data easier to present or read. There are multiple advantages of the thematic analysis approach. According to Braun and Clarke (2006), the thematic analysis is flexible, relatively easy to learn, able to summarise key features of a large body of data, able to highlight similarities and differences across the dataset, and able to generate unanticipated insights. Braun and Clarke (2006) provided specific procedures, checklists, and assessment criteria of a rigorous thematic analysis, which can be used for grounded theory, too. Therefore, during the data analysis process, this author valued the coding criteria proposed by Braun and Clarke (2006), as shown below:

1) Transcribed the interview responses in detail, and checked the transcripts against the record for accuracy;

2) During the coding process, each piece of information was given equal attention;
3) The generated themes were concluded with considerations of internal coherency, consistency and distinctiveness;
4) All relevant extracts were collated;
5) Themes were checked against each other to make sure they are distinctive from each other. Themes were also compared with the original dataset to ensure there were no themes missing or new themes emerging;
6) The themes were internal coherency, consistency and distinctiveness. Though have been mentioned in the third step, this step is essential because thematic analysis could be a dynamic process that keeps changing in Steps four and step five;
7) Data were "analysed… rather than just paraphrased or described" (Braun & Clarke, 2006, p. 96). During this process, the analysis and data matched each other, and the categorization of the interview responses was "convincing and well-organized" (Braun & Clarke, 2006, p. 96).
8) Overall, enough time was allocated to complete all phases of the aforementioned analysis.

During the qualitative data analysis, this author coded the qualitative responses as individual statements with each one containing one piece of key information. After removing the irrelevant statements that were off-topic, in total this author obtained 368 responses from surveys and 108 from interviews.

In reporting the interview responses, "P" was used to indicate "participant", followed by the number of the participant; for example, P2 indicates participant number two.

4.8.2. Data analysis- Quantitative data

After collecting quantitative data via the established surveys based on interview responses, the data were analysed using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) and model fit (To be revised).

Factor analysis is a "data reduction technique" used primarily for questionnaire or scale development (Pallant, 2016, p. 182). Factor analysis enables researchers to determine how many latent variables underlie a set of items and condenses a large set of variables into a smaller set of factors or sub-scales (DeVellis, 2003).

In general, the survey conceptualisation and validation procedures were guided by the principles of measurement put forward by Robertson (2017) and DeVellis (2003). Since scale evaluation involves multiple techniques, detailed information will be introduced in the next chapter.

4.9. Reliability and validity

A good research practice is expected to use a variety of research methods to demonstrate the robustness of the findings (Carlson & Herdman, 2012), so reliability and validity need to be paid special attention to. This author addresses reliability and validity as follows.

4.9.1. Reliability concerns in qualitative studies

As summarised by Whittemore, Chase, and Mandle (2001), reliability refers to the *stability* of findings, while validity represents the *truthfulness* of findings (p. 523).

To guarantee the reliability of the coding process, this author employed the coding reliability approach, which is to use multiple coders to independently code the data and compare the results of different coders. Cohen's Kappa was used as the measurement of the inter-rater agreement at the qualitative data analysis stage. Cohen's kappa measures the agreement between two coders who each classify items. According to McHugh (2012), the Cohen's Kappa value 0.01–0.20 indicating none to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement. The result turned out to be 0.834 and 0.928 for the survey responses and interview responses respectively ($p < .001$). The Cohen's Kappa results indicated good inter-coder reliability of the coding process as the value between 0.81 to 1.00 indicates perfect agreement (McHugh, 2012). Since the data collection was conducted in China, the survey has been translated into Chinese and back-translated into English to minimise information mis-delivery.

4.9.2. Validity concerns in qualitative studies

It is challenging to establish validity criteria in qualitative research since qualitative research incorporate rigour, subjectivity and creativity in the research process (Creswell & Miller, 2000; Whittemore et al., 2001). In addition, different qualitative methods espouse varied evaluation criteria (Whittemore et al., 2001). Whittemore et al. (2001) proposed that the primary criteria in qualitative research include credibility, authenticity, criticality, and integrity, while the secondary criteria to address are explicitness, vividness, creativity, thoroughness, congruence, and sensitivity (p.534) (see Table 20).

Table 23. Validity criteria in qualitative studies (Whittemore et al., 2001)

Category	Criterion
Primary validity	1) Credibility
	2) Authenticity
	3) Criticality
	4) Integrity
Secondary validity	1) Explicitness
	2) Vividness
	3) Creativity
	4) Thoroughness
	5) Congruence
	6) Sensitivity

The main validity criteria considered in the current thesis were credibility, authenticity, criticality, integrity, thoroughness, and explicitness.

Credibility refers to "the conscious effort to establish confidence in an accurate interpretation of the meaning of the data", therefore the assurance of credibility in qualitative research include the considerations on whether the results reflect the experience of participants, whether the results reflect the context in a believable way, and does the explanations fit the descriptions (Whittemore et al., 2001, p. 530). *Authenticity*, closely linked to credibility, "reflects the meanings and experiences that are lived and perceived by the participants" (Whittemore et al., 2001, p. 530). Because of multivocality (different people make interpretations from varied perspectives), phenomenon or situation is important in qualitative studies. *Criticality* addresses whether the research process demonstrates "evidence of critical appraisal" (Whittemore et al., 2001, p. 534). *Integrity* in qualitative research refers to the "recursive and repetitive checks of validity as well as a humble presentation of findings" (Whittemore et al., 2001, p. 534). *Thoroughness* in qualitative research refers to "sampling and data adequacy as well as the comprehensiveness of approach and analysis" (Whittemore et al., 2001, p. 532). Thoroughness implies the "full development of ideas" (Whittemore et al., 2001, p. 532). *Explicitness* requires the researchers to specify the idea without ambiguity (Whittemore et al., 2001).

Validation techniques in the current study

This thesis employed the following techniques for demonstrating validity, which were selected from the list proposed by Whittemore et al. (2001):

1) Employing triangulation. Triangulation refers to the practice of using multiple sources of data (data triangulation), multiple evaluations in observations or data analysis (investigator triangulation), multiple theoretical perspectives (theory triangulation) or multiple methods to study a single problem (methodological triangulation), which helps enhance the credibility of a research study (Salkind, 2010).

2) Member checking. Being described as "the most crucial techniques for establishing credibility" (Lincoln & Guba, 1985, p. 314), member checking requests researchers to take the data and interpretations back to the participants for them to confirm. In this way, the participants have a chance to "react to both the data and the final narrative" (Creswell & Miller, 2000, p. 127), which add credibility to the qualitative study. In the current study, member checking was applied when collecting interview responses.

3) Expert checking. Expert checking is achieved by "engaging with other researchers to reduce research bias" (Noble & Smith, 2015, p. 35). Expert checking

4) Negative case analysis. In the first round of qualitative data analysis, findings emerge patterns and trends within participants' responses, and the negative cases are those which do not fit the patterns or trends. If the patterns and trends are seen as the preliminary "rules" for how participants are responding, the negative cases may "broaden the rule, change the rule, or cast doubt on the rule altogether" , which makes negative case analysis of great significance.

Birks and Mills (2015) commented that the credibility of a grounded theory study is dependent on three main factors: researcher expertise, methodological congruence, and procedural precision. This author addressed the potential concerns by taking the suggestions of Birks and Mills (2015), including extensively reading the available literature, being committed to achieving quality outcomes, feeling an affinity for the topic, being honest about the limitations

of the research, employing “memoing” in support of the study (“the records of thoughts, feelings, insights and ideas in relation to a research project”) (p.39), making logical connections between the data and abstractions, etc.

Other techniques listed by Whittemore et al. (2001) were also considered, including performing a literature review, using computer programs (SPSS, Microsoft Excel, Amos, Nvivo), testing hypothesis in data analysis, “drawing data reduction tables”, and “providing thick descriptions” (p.533).

4.9.3. Survey and survey item reliability

There are good surveys and bad ones: bad surveys produce data that are unreliable, irreproducible or invalid, while the good ones provide insights into the topic of interest (Fink & Litwin, 1995). It is essential to guarantee the reliability and validity of a scale before data analysis.

Reliability is “a statistical measure of how reproducible the survey instrument's data are” (Fink & Litwin, 1995, p. 6). According to Fink and Litwin (1995), the reliability of a scale is commonly assessed in three forms: test-retest, alternate-form, and internal consistency.

Test-retest reliability is the most commonly used indicator of survey instrument reliability, which is measured by having the same set of respondents complete the survey “at two different points in time” to see “how stable the responses are” (Fink & Litwin, 1995, p. 6). It is suggested to conduct a correlation study to compare the two sets of responses, in which the correlation coefficients (or *r* values) are expected to be equal to or exceed 0.70 (Fink & Litwin, 1995).

Alternate-form reliability involves the use of differently-worded items to measure the same concept (Fink & Litwin, 1995) to minimise content sampling error (Urbina, 2014). Scale designers can also change the order of survey items.

Internal consistency reliability is a commonly used measure in assessing instruments and scales (Fink & Litwin, 1995), which is to assess how well the different survey items measure the same concept. Internal consistency is measured by calculating a statistic known as Cronbach's coefficient alpha.

The following table (Table 21) summarises the types of reliability and the characteristics, which was borrowed from statements by Fink and Litwin (1995).

Since alternate-form reliability needs equivalent item wording, which is demanding on language proficiency, this author as an English-as-the-second-language learner did not consider the alternate-form reliability in evaluating the scale. Since the current study used the convenience sampling method, test-retest reliability does not apply to the current study. The eventually used reliability assessment is internal consistency reliability.

Table 24. Types of reliability (Fink & Litwin, 1995, p. 32)

Type of reliability	Characteristic
Test-retest	“Measures the stability of responses over time, typically in the same group of respondents.”
Alternative-form	“Uses differently worded stems or response sets to obtain the same information about a specific topic.”
Internal consistency	“Measures how well several items in a scale vary together in a sample.”

4.9.4. Survey and survey item validity

Validity information indicates how well a test measures a given area with a given topic (Dick & Hagerty, 1971). According to Dick and Hagerty (1971), there are seven types of validity including content, concurrent, construct, predictive, face, curricular and differential. Others also concluded that there are three types of validity external validity, internal validity and instrument validity (Table 22). However, the categorisation is not important: the type of validity which should be estimated for a test depends on the aims of the test, so not every type of validity is required to be tested.

Face validity is based on "a cursory review of items by untrained judges" (Fink & Litwin, 1995, p. 35), which is to assess whether the survey items seem to measure what they claim to measure. According to Fink and Litwin (1995), face validity is the "least scientific measure of all the validity measures", and in fact "many researchers do not consider face validity a measure of validity at all" (Fink & Litwin, 1995, pp. 34-35).

Content validity assessed whether the survey "includes everything it should and does not include anything it shouldn't" by investigating the "overall opinion of a group of trained judges" (Fink & Litwin, 1995, p. 35). Content validity is not quantified with statistics; instead, it is a subjective measure that can be assessed by involving reviewers who have some knowledge of the subject matter.

Criterion validity is a measure of "how well one instrument stacks up against another instrument or predictor" (Fink & Litwin, 1995, p. 36), Criterion validity can be divided into two components: concurrent validity and predictive validity (Streiner, Norman, & Cairney, 2015).

- *Concurrent validity* measures how well a new test compares to a well-established test. It can also refer to the practice of concurrently testing two groups at the same time, or asking two different groups of people to take the same test. It is important to select a good and relevant standard to judge the new survey instrument (Fink & Litwin, 1995). Concurrent validity can be assessed by examining the correlation coefficient of the data.
- *Predictive validity* is "the ability of a survey instrument to forecast future events, behaviors, attitudes, or outcomes" (Fink & Litwin, 1995, p. 37). Like concurrent validity, predictive validity is calculated as a correlation coefficient between the initial test and the secondary outcome.

Construct validity is the measure of how meaningful the scale is in practical use. Construct validity is "the most valuable yet most difficult way of assessing a survey instrument", which is "difficult to understand, to measure, and to report" (Fink & Litwin, 1995, p. 39). Construct validity is often thought to comprise two other forms of validity: convergent validity and divergent validity.

- *Convergent validity* "implies that several different methods for obtaining the same information about a given trait or concept produce similar results" (Fink & Litwin, 1995, p. 45); or, it reflects "the extent to which two measures capture a common construct" (Carlson & Herdman, 2012, p. 18). Evaluating convergent validity is analogous to measuring alternate-form reliability, except that the former is more theoretical and requires a great deal of work, usually by multiple investigators with different approaches" (Fink & Litwin, 1995, p. 45). According to Carlson and Herdman (2012), if two measures are assumed to represent the same construct, there should be a strong correlation between them. Correlations closer to 1 indicate strong convergent validity; or, they indicate that the measures do not capture the intended construct well (Carlson & Herdman, 2012). However, it is notable that strong convergent validity is not sufficient for construct validity, since two measures with strong convergent validity may measure the same wrong things (Carlson & Herdman, 2012; Nunnally, 1978), so convergent validity cannot be the only measure to address in scale development.
- *Divergent (discriminant) validity* is "another theoretically based way of thinking about the ability of a measure to estimate the underlying truth in a given area. For a survey instrument to have divergent validity,

it must be shown not to correlate too closely with similar but distinct concepts or traits" (Fink & Litwin, 1995, p. 45).

Table 25. Validity types and measures

Validity type	Sub-type	Description
Content validity	n/a	"Formal expert review of how good an item or series of items appears" (Fink & Litwin, 1995, p. 45). Or,
		The degree to which game content adequately covers the dimension of the target construct it aims to associate with (Graafland, Schraagen, & Schijven, 2012)
Face validity	n/a	"Casual review of how good an item or group of items appear."
Criterion	Concurrent validity	"Concordance of study results using a concept instrument (e.g. game) and study results on an established instrument or method, believed to measure the same medical theoretical construct" (Graafland et al., 2012).
	Predictive validity	The degree of concordance of a concept instrument (e.g. game) outcome and task performance in reality, based on a validated scoring system (Graafland et al., 2012).
Construct validity	Convergent validity	It reflects "the extent to which two measures capture a common construct" (Carlson & Herdman, 2012, p. 18).
	Divergent validity (discriminant validity)	"For a survey instrument to have divergent validity, it must be shown not to correlate too closely with similar but distinct concepts or traits" (Fink & Litwin, 1995, p. 45).

According to Urbina (2014), there is a more comprehensive way of dealing with a large number of constructs through a series of statistical procedures knowledge collectively as factor analysis. The main goal of factor analysis is to "reduce the number of dimensions needed to describe data derived from a set of measured variables", and to "investigate the structure that accounts for the interrelationships between the variables" (Urbina, 2014, p. 186).

From the publications on scale development, the common validity test measures are:

1) Convergent factor validity
2) Discriminant factor validity

Test-retest reliability, alternate form reliability, and internal consistency are popular considerations in reliability tests.

1) Test-retest reliability. Ask the same group of participants to take the same test for a second time after a period of time. The score for tests 1 and 2 should be highly correlated with each other. The test-retest method to measure reliability is considered as "an index of the stability of the test" (Dick & Hagerty, 1971, p. 19).
2) Alternate forms reliability also named parallel forms reliability. Ask students to fill in two different forms of the same test. The score from the two forms of the test should be highly correlated (Dick & Hagerty, 1971).
3) Internal consistency (Cronbach's Alpha, Split half reliability)

CHAPTER 5

Scale evaluation techniques

05

Scale evaluation techniques

Though usually described in the methodology section, the procedures and techniques for scale evaluation were separated as an individual chapter since they are rich in content.

While the Methodology section mainly introduces the typical methodological concerns such as research design, participants, data collection, and data analysis, this chapter (Chapter 5) presents overall information on how a scale can be evaluated.

CHAPTER FIVE: SCALE EVALUATION TECHNIQUES

This study follows the typical procedures of scale development proposed by DeVellis (2003), which contains eight steps as shown in Table 26 and as detailed in the following sections.

Table 26. Eight steps of scale development (DeVellis, 2003)

#	Description
Step 1	Determine what to measure
Step 2	Generate an item pool
Step 3	Determine the format for measurement
Step 4	Have the initial item pool reviewed by experts
Step 5	Consider the inclusion of validation items
Step 6	Administer items to a development sample
Step 7	Evaluate the items
Step 8	Optimise scale length

5.1. Step 1: Determining what to measure

DeVellis (2003) criticised that though researchers "think they have a clear idea of what they wish to measure", many of them fail to clearly define it (p. 60). Therefore, the first and foremost thing to consider in scale development is determining what to measure. The involved issues include whether to develop scales based on theories, what should be included in a measure, how specific should the measure be, should some aspects of the phenomenon be emphasised more than others, etc.

At this step, a theoretical framework serves as an aid to clarity, as researchers are required to be clear about the construct being measured (DeVellis, 2003; Robertson, 2017). According to Robertson (2017), key considerations at this stage include the identification of all the main constructs of interest, the specification on what will and will not be measured, the providence of precise definition of the selected constructs, and the involvement of a conceptual or theoretical model to explain the phenomena of interest. In choosing the theoretical model, the scale designers should make a decision between general models and specific models. The former ones help develop general and global measures that apply to more contexts, while the latter ones intend to relate to specific behaviours or situations (DeVellis, 2003).

If there is a lack of extant theories that directly contribute to the establishment of scale item pool, the researchers can draft the survey items based on their experience, understanding or exptertise. However, the decision should be based on "at least a tentative theoretical model that will serve as a guide to scale development" (DeVellis, 2003, p. 61). In this case, researchers are expected to detail how the new construct relates to existing phenomena and how the construct is defined (operationalisation).

As guided by DeVellis (2003), this author firstly searched relevant frameworks that illustrate teachers' acceptance intention to gamified learning tools. Failed to obtain related frameworks, this author considered draft the survey items based on personal experience and understanding, and ended the plan in consideration of research reliability and validity.

As a compromise, this author decided to modify the survey items that are drafted based on theoretical framworks about technology acceptance intention: as there is a lack of theoretical models focusing on the acceptance intention of gamification in the specific context (secondary schools in test-oriented countries), whereas gamification could be regarded as one option of general educational technologies. In other words, if take gamification as one general educational technology, there are numerous frameworks that could be used or adapted as the theoretical foundation; however, if take research contexts into considerations, there will be a lack of appropriate theory related to the measurement problem at hand, therefore this author needs to conceptualise the framework by choosing a tentative one, testing it on participants, and making modifications.

In determining what to measure, one notable issue is what to include in the scale. According to DeVellis (2003), it is possible that similar items tap quite different constructs. In such cases, the designed scale may be sensitive to one phenomenon even though it is designed to measure another one. Though researchers can interview experts for guaranteed face validity, which refers to the extent to which a scale appears to measure what it claims to measure, further statistical analysis techniques are still needed. The detailed analysis processes are introduced in Step 7 and Step 8.

5.2. Step 2: Generating an item pool

Hinkin, Tracey, and Enz (1997) indicated that the item pool generation can be done inductively or deductively.

Inductive scale development refers to generating items first and then derive the items, which is usually used when exploring "an unfamiliar phenomenon where little theory may exist" (p. 102). For example, scale designers can ask experts in the field to provide descriptions or comments on the target research objective, conduct a content analysis on the responses, classify the responses into several categories, and then derive items from the categorised responses (Hinkin et al., 1997).

Deductive scale development "uses a theoretical definition of a construct", which serves as a guide for the item generation (Hinkin et al., 1997, p. 102). The deductive approach requires an understanding of the relevant literature and the target research objectives for guaranteed content adequacy in the final scale. Hinkin et al. (1997) suggest that "in most situations where some theory exists, the deductive approach would be most appropriate" (p. 102).

Similar to the inductive-or-deductive suggestion, Streiner et al. (2015) suggested that once the scale purpose has been clearly defined, scale developers can devise the items by:

1) Selecting and/or modifying existing items from previous studies; or

2) Creating a list based on qualitative responses of the target group of people.

Selecting and or modifying existing items from previous studies (related literature or questionnaires) is advantageous as it saves work in constructing and validation (Robertson, 2017). What's more, at least in some contexts (clinic

assessment), the ways to ask about a specific problem is limited, so turning to existing items is a good option (Streiner et al., 2015).

Disadvantages also apply, as the items in previous studies might be outdated, and the varied research purposes are inadequate of being used (Streiner et al., 2015). In this case, designers are expected to create new items based on different resources, such as the subject themselves, observation, theory, research and expert opinions (Streiner et al., 2015). Similarly, Robertson (2017) detailed that scale designers can generate the item pool from some forms of qualitative research, such as observation, focus groups, in-depth interviews and ethnography. In the current thesis, both approaches were employed.

Churchill Jr (1979) emphasised that items at the early stages of item generation should "tap each of the dimensions of the construct at issue", which are with "slightly different shades of meaning" (p. 68). Researchers are expected to incorporate statements with slightly different nuances of meaning, which provides a better foundation for the eventual measure (Churchill Jr, 1979).

Theoretically, a good item pool contains items randomly selected from "the universe of items relating to the construct of interest" (DeVellis, 2003, p. 64), which makes the scale over-inclusive or redundant. The number of items in the initial item pool can be large, as "it is not unusual to begin with a pool of items that is three or four times as large as the final scale" (DeVellis, 2003, p. 66), or "at least twice as many items" as what needed for the final scale (Hinkin et al., 1997, p. 104). The current study does not limit the number of items; instead, it aims to involve all possible items to the author's knowledge when building the initial item pool.

As directed by Hinkin et al. (1997) and DeVellis (2003), 'bad' items were removed, including those which are ambiguous, exceptionally lengthy, high in reading difficulty level, and double-barrelled (that convey two or more ideas, so the item might refer to either or both ideas).

Negative-wording items were not considered in the current study. The intent of wording items both positively and negatively within the same scale is to "avoid an acquiescence, affirmation, or agreement bias", which refer to a responder's "tendency to agree with items, irrespective of their content" (DeVellis, 2003, p. 69). However, respondents may become so confused that they give inconsistent answers, no responses or lower average scores on negative items (DeVellis, 2003; Salazar, 2015), which generates a loss of scale reliability by lowering levels of correlation between the scores for an item and the total score (Salazar, 2015). As a consequence, DeVellis (2003) asserted that the disadvantages of wording items in an opposite direction overweight advantages; Dalal and Carter (2014) commented that negative-wording items are appropriate when they are to measure substantive traits (e.g., anxiety, negative attitudes), or they could be not necessary; Salazar (2015) conducted a comparison study and summarises that the use of positive-wording items does bring acquiescence bias, however, "a remedy to this situation only made matters worse, as the combination of positive and negative items seriously affected the internal consistency of the scales" (p. 192).

5.3. Step 3: Determining the format for measurement

There are various formats for measurement, such as Likert scale, semantic differential scaling, visual analog scale, binary options and item time frames (Robertson, 2017). Among the varied formats of measurement, Likert scale is one of the most common item formats in scale development (Hinkin et al., 1997).

When a Likert scale is used, "the item is presented as a declarative sentence, followed by response options that indicate varying degrees of agreement with or endorsement of the statement" (DeVellis, 2003, pp. 78-79). A common application of Likert scale is to include several possible responses such as "strongly disagree, moderately disagree", "slightly disagree", "neutral", "slightly agree", "moderately agree" and "strongly agree", which "form(s) a continuum from strongly disagree to strongly agree" with or without a neutral midpoint (DeVellis, 2003, p. 79). Likert scales are widely used in instruments measuring opinions, beliefs and attitudes (DeVellis, 2003).

In the last decades, there are debates on the optimal number of response categories in rating scales (Cox III, 1980). Dawes (2002) examined whether the number of scale points used in research surveys affects research findings. By comparing the results (Mean, variance, kurtosis, skewness) of a 5-point scale and an 11-point zero-to-ten scale, Dawes (2002) found that the data produced is essentially the same.

In a following study, Dawes (2008) evaluated whether the varied "scale format" (which refers to "scales with differing numbers of response categories") changes data characteristics by comparing the aggregate-level data characteristics derived from attitudinal questions with either 5-point, 7-point or 10-point numerical scales (p. 61). Findings show that the 10-point scale is statistically and significantly different to the other two formats on the mean score (0.3 point lower when all the three scales been rescaled to scores out of 10), and the three different formats exhibit no appreciable differences in terms of standard variation, skewness and kurtosis. Dawes (2008) concluded that 5-point scale can be transferred equivalently to a 7-point one, and either 5-point, 7-point or 10-point scales are comparable for analytical tools such as confirmatory factor analysis or structural equation modelling.

Preston and Colman (2000) compared the differences of 12 scale formats, with 11 of them range from 2 to 11 and the remaining one contains 101 points. The findings show that the two-point, three-point and four-point scales "performed relatively poorly", and indices were "significantly higher" for scales with more response categories up to about seven (Preston & Colman, 2000, p. 1). Respondents in the study reported that they prefer the 10-point scale, closely followed by the 7-point scale.

While the aforementioned studies discussing the equivalent effect of 5-point and 7-point scales, some others indicate 7-point scales could be more ideal. Maydeu-Olivares, Fairchild, and Hall (2017) examined the effect of a varied number of responses in rating scales in regard to the goodness of fit in factor analysis. They "recommend using a large number of response alternative (≥ 5)" to increase the power to detect incorrect substantive models (p. 495). Finstad (2010) investigated whether the 5-point and 7-point Likert scale impact differently on users reports on two software applications' perceived usability. Finstad (2010) concluded that "7-point Likert items provide a more accurate measure of a participant's true evaluation and are more appropriate for electronically-distributed … questionnaires" (p. 104). In earlier studies, Lewis (1993) confirmed that "mean differences from 7-point scales correlate more strongly with observed significance levels than those for 5-point scales" (p. 383), and Colman, Norris, and Preston (1997) said that 7-point rating scales explain more variances and fit the data better. With the aforementioned considerations in mind, this author selected 7-point Likert scale in the current study.

5.4. Step 4: Having the initial item pool reviewed by experts

DeVellis (2003) suggested having the initial item pool reviewed by experts to confirm or invalidate the definition of the phenomenon, evaluate the items' clarity and conciseness, and point out "ways of tapping the phenomenon that you have failed to include" (p. 86). Robertson (2017) concluded that at this step experts' opinions are needed for face validity and content validity, as followed below:

Face validity "simply indicates whether, on the face of it, the instrument appears to be assessing the desired qualities" (Streiner et al., 2015, p. 5), which is to guarantee whether the items on the scale are adequate for measuring the construct they claimed to measure (da Silva Brito, Contreras Pinochet, Luiz Lopes, & de Oliveira, 2018). Normally, face validation is based on subjective judgement by one or more experts in the field with no empirical approaches used (Streiner et al., 2015).

Content validity addresses the issue of whether all facets of the construct of interest have been measured (Robertson, 2017). The same with face validity, content validity is usually assessed by subject matter experts, which is not "measured" but "assured" by the invited experts (Robertson, 2017, p. 13). Robertson (2017) adds that efforts should be made to ensure that each item reflects the underlying construct the designer mean to measure, which can be enhanced by continually referring to the theoretical underpinnings of the framework. The specific content assessment methods include requiring respondents to categorise or sort items based on similarity to construct definitions, or asking respondents to rate the extent to which items corresponded with construct definitions (Hinkin et al., 1997).

5.5. Step 5: Considering inclusion of validation items

DeVellis (2003) suggested adding validation items to help detect undesirable response tendencies. The first type is the items detecting participants' social desirability, which are normally with answers regarded as negative in the society. For example, the social desirability items could be "would you smile at people every time you meet them", "would you ever lie to people" (Haghighat, 2007), "I have doubts about my ability to succeed in life", and "I say things that hurt others' feelings" (He et al., 2015, p. 234). If the participants are strongly motivated to present themselves in a way that society regards as positive, the responses may be distorted.

Similarly, participants are also possible to have "prestige bias" by "faking good" (Streiner et al., 2015, p. 75): when asked what the participants would use televisions for, the participants tend to answer watching educational and cultural programmes instead of watching game shows or soap operas, even though the latter ones are obviously much more popular. There are also other types of bias, such as deviation ("the tendency to respond to test items with deviant responses"), faking bad (which occurs when "a person feels he or she may avoid an unpleasant situation (such as the military draft) by looking bad"), yea-saying or acquiescence bias ("the tendency to give positive responses to a question"), end-aversion bias or central tendency bias ("the reluctance of some people to use the extreme categories of a scale") (Streiner et al., 2015, pp. 75-79). When design the scale, the aforementioned bias were paid special attention to.

This author designed two validation questions to filter the surveys that are of potential in decreasing the overall reliability. The two validation questions asked the participants to select "Disagree/2". While the answers for other items tended to be "agree", namely the option of "5" (Slightly agree), "6" (Agree), or "7" (Strongly agree), it is assumed that participants have a tendency of choosing 5, 6 or 7 as well if they do not read the survey item carefully. The surveys with abnormal answers were also excluded, such as the ones selected "Agree" for almost all the items. This technique filtered about 30% to 40% invalid survey responses.

5.6. Step 6: Administering items to a development sample

Step 6 is mainly about data collection with the consideration of an adequate sample size. According to DeVellis (2003), the factor analysis based on a large-sample data collection will be generally more stable than that based on a smaller sample, therefore the sample size for a scale evaluation study is an issue.

Both absolute and the relative (i.e., a ratio of subjects to items) number of subjects were considered by previous researchers. Nunnally (1978) suggests that 300 is an adequate number, and the scales can be developed successfully with a smaller sample size when the number of items and the number of scales to be extracted is small. Comrey (1973) proposes that a sample size of 100 is poor, 200 fair, 300 good, 500 very good, and 1000 excellent. After that, Comrey (1988) adds more details such as a sample size of 200 is reasonably good for ordinary factor-analytic work with 40 or fewer variables, and a sample size of 400 makes a factor analysis with few problems.

However, these "simplistic guidelines" have been criticised for not considering the numbers and types of variables being analyzed" (Tinsley & Tinsley, 1987, p. 415), therefore, researchers also sought for a standard ratio of subjects to items. A widely used standard is the ratio of subject to items (5 to 1 or 10 to 1), namely at least 10 subjects or five subjects are required for every item presented on a scale (Anthoine et al., 2014; Baggaley, 1983; Everitt, 1975; Gorsuch, 1983; Nunnally, 1978; Tinsley & Tinsley, 1987), while the ratio can be relaxed when the sample size reached 300 (Tinsley & Tinsley, 1987).

Item communalities and factor loading also play roles (DeVellis, 2003; Pallant, 2016). Fabrigar and Wegener (2011) proposed that reasonable results can also occur with a substantially smaller sample when the data is in good conditions. To be specific, under optimal conditions (communalities of 0.70 or greater and three to five measured variables loading on each factor), a sample of 100 "can be adequate" (p. 26), and under moderately good conditions (communities of 0.40 to 0.70 and at least three measured variables loading on each factor), a sample of 200 is sufficient. Tabachnick, Fidell, and Ullman (2007) suggested a sample size of 300 would be recommended, but it can be reduced to a smaller one (e.g. 150) if the factor loading is above 0.80.

It is notable that the scale development could be a trial-and-error process with failures, namely not each designed scale can be validated successfully; accordingly, this author was prepared to modify the scales and conduct another round of data collection when necessary. Therefore, though 400 or 500 is the ideal sample size, this author planned to follow the standard of Comrey (1988) and Fabrigar and Wegener (2011) to limit the sample size to 200 at the exploratory stage. When the scales were confirmed, another round of data collection with a large sample size (400 to 500) could be pursued.

5.7. Step 7 and Step 8: Scale evaluation

After developing the scale and administering it to an appropriately large and representative sample, scale designers should evaluate the performance of the survey items, which is the heart of the scale development process (DeVellis, 2003). In the book *Scale Development: Theory and Applications* by R. DeVellis (2003), the seventh and eighth steps are named "evaluate the items" (p. 104) and "optimise scale length" (p. 110), which are named as "scale evaluation" by this author. Scale evaluation consists of two main steps: data adequacy check before the formal data analysis and the factor analysis.

5.7.1. Factor analysis: Function and definition

The seventh and eighth steps (evaluate the items and optimise scale length) for scale development can be named scale evaluation, which is usually achieved by using the factor analysis approach.

According to Urbina (2014), there is a more comprehensive way of dealing with a large number of constructs through a series of statistical procedures knowledge collectively as factor analysis. Factor analysis serves several purposes, in which the most primary function is to help determine how many latent variables underlie a set of items (DeVellis, 2003). Alternatively speaking, factor analysis is a technique used to "identify or confirm a smaller number of factors or latent constructs from a large number of observed variables (or items)" (Worthington & Whittaker, 2006, p. 807). Fabrigar and Wegener (2011) specify that researchers often encounter a large set of observations or measures for a group of people (the research objects), which follows with two concerns: whether this large set of observations can be "more parsimoniously represented", and whether it is possible to figure out the "underlying structure of associations for the set of measures" (p. 1). It is possible that there is no parsimonious representation of the measures at all, certain measure reflects "something uniquely its own that has little relation to any other measure in the set of observations", or, the set of measures reflect a single underlying construct or a few distinct constructs (Fabrigar & Wegener, 2011, p. 1).

There are two main ways to conduct factor analyse: the exploratory factor analysis (EFA) that discover which factors or dimensions underlie the measures, and the confirmatory factor analysis (CFA) aiming to test hypotheses or to confirm theories.

Correlation is an easy-and-quick technique to examine the aforementioned issues: if all the measures are examining one construct, the correlations among them should be strong and of similar strength; if two subsets of measures examine two different and unrelated constructs, measures of each subset are expected to strongly correlated with one another and unrelated to measures from the other subset; and if each measure assesses a different construct, the correlations among the measures should be not significant (Fabrigar & Wegener, 2011). However, unfortunately, real data, especially the ones with a large number of measures, produce complicated patterns that are difficult for visual inspection (Fabrigar & Wegener, 2011).

5.7.2. Suitability assessment

Prior to the data analysis, scale developers are suggested to conduct an initial examination of items' performance to confirm whether the data is suitable for factor analysis (DeVellis, 2003; Fabrigar, Wegener, MacCallum, & Strahan, 1999). According to Pallant (2016), there are two main concerns affecting the results: sample size and the strength of the inter-correlations among items.

The adequate sample size for factor analysis has been discussed in Step 6 in the section "Procedure of scale development", which concluded that the sample size should be larger than 200.

FUNCTION OF FACTOR ANALYSIS

Identify or confirm a smaller number of factors from a large number of survey items.

Concerning the strength of the inter-correlations among items, Tabachnick and Fidell (2013) recommend a benchmark of .30 as to the coefficients in the correlation matrix. A set of highly inter-correlated items indicates that each individual item should be correlated substantially with the collection of remaining items, too. Therefore, scale developers can compute item-scale correlation (DeVellis, 2003). According to (DeVellis, 2003), an item with a high value for the item-scale correlation is "more desirable than an item with a low value" (p. 93), while the items with low values should be "eliminated first" (p. 98).

Pallant (2016) suggests using two statistical measures generated by IBM SPSSS to help assess the factorability of the data: Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. Bartlett's test of sphericity should be "significant ($p < .05$) for the factor analysis to be considered appropriate" (Pallant, 2016, p. 184). The KMO index ranges from 0 to 1, and it is expected to be higher than .60 for good factor analyses (Worthington & Whittaker, 2006). The specific measures and standards were shown in Table 24.

Table 27. Measures and standards in assessing the suitability of data for factor analysis

Concern	#	Measure	Standard
Sample size	1	Sample size	Larger than 200 in the current study
Strength of the relationship among variables	2	Coefficients in the correlation matrix	Greater than .30
	3	Bartlett's test of sphericity	Significant ($p < .05$)
	4	KMO index	Greater than .60

5.7.3. Factor analysis: The selection of EFA or CFA

There are two main categories of factor analysis: exploratory and confirmatory (see Table 25). If researchers are not sure about the number of constructs, the recommended procedure is to conduct *exploratory factor analysis* (EFA). In contrast, if researchers have a theory that "clearly specifies a precise number of factors and exactly which measured variables each factor should influence", a *confirmatory factor analysis* (CFA) is preferred (Fabrigar & Wegener, 2011, p. 28). The selection of EFA or CFA is an issue as the researchers' expectations "will fall somewhere between these two extremes" (Fabrigar & Wegener, 2011, p. 29).

If the researchers specify two or three competing models instead of one single model as the target theoretical foundation, as long as the researchers have "sufficient theoretical foundation to precisely specify two or three competing models" which are with adequately developed theories, a CFA can be applied (Fabrigar & Wegener, 2011, p. 29).

Researchers are recommended to adopt EFA when the researchers are generating new items to measure a particular construct, or when the researchers have a general idea of the number of factors and the correspondent contributing variables but not supported by a "sufficiently developed" theory or a theory providing "a high level of confidence" (Fabrigar & Wegener, 2011, p. 28).

Worthington and Whittaker (2006) indicate that researchers typically conduct CFA on an instrument after it "has already been assessed using EFA" (p. 815). Holding the same viewpoint, Fabrigar et al. (1999) comment that it is useful to use both EFA and CFA in evaluating an instrument, as the EFA can be conducted in an initial study to "provide a basis" for specifying a CFA model in a "subsequent" study (p. 277). During this process, researchers are expected to use a new sample for enhanced validation. However, if the sample size in one study is large enough, the

sample could be randomly split into two parts, with one part for EFA and the other half for CFA (Fabrigar et al., 1999).

Table 28. The application of EFA and CFA

Method	Application
Exploratory factor analysis (EFA)	When the researchers are generating new items to measure a particular construct. When the researchers' expectations are not supported by a "sufficiently developed" theory or a theory providing "a high level of confidence" in regard to the exact number of factors and the corresponding contributing variables (Fabrigar & Wegener, 2011, p. 28).
Confirmatory factor analysis (CFA)	"investigate the structure that accounts for the interrelationships between the variables" (Urbina, 2014, p. 186).

5.7.4. Factor analysis: Factor extraction

Factor extraction is mainly to determine the "smallest number of factors" that can be used to present the interrelationships among the involved variables (Pallant, 2016, p. 184). The IBM SPSS provides the popular extraction techniques including principal components, unweighted least squares, generalised least squares, maximum likelihood factoring, principal axis factoring, alpha factoring and image factoring.

There are two conflicts in factor analysis: researchers are expected to "find a solution with as few factors as possible" to make the scale concise, and the factors need to explain "as much of the variance in the original data set as possible" (Pallant, 2016, p. 184). Alternatively speaking, scale designers need to find a balance between "parsimony (i.e., a model with relatively few common factors)" against the need for "plausibility (i.e., a model with a sufficient number of common factors to adequately account for the correlations among measured variables)" (Fabrigar et al., 1999, pp. 277-278).

Therefore, there is a need to determine the number of extracted factors from the data set. Tabachnick and Fidell (2013) recommend researchers adopt an exploratory approach, in which experiment with different numbers of factors until found a satisfactory solution. There are several techniques to assist this decision-making process, including Kaiser's criterion, scree test, and parallel analysis (Pallant, 2016). The current study adopted Kaiser's criterion in scale evaluation.

5.7.5. Factor analysis: Item removal

DeVellis (2003) proposed to optimise scale length, which is to help determine whether to keep the validated survey long or short, and whether the shortened version is appropriate at the price of "a bit less reliability" (DeVellis, 2003, p. 97).

Scales with a large number of items "are generally more reliable" however, a questionnaire taking a long time to complete, such as "longer than about 15 to 30 minutes might become problematic (Worthington & Whittaker, 2006, p. 824).

According to Worthington and Whittaker (2006, p. 824), when a factor contains more than the desired number of items, the survey designers can choose to delete items that:

1) “Have the lowest factor loadings”,

2) “Have the highest cross-loadings”,

3) “Contribute the least to the internal consistency of the scale scores”, and

4) “Have low conceptual consistency with other items on the factor”.

Following the guidance, this author shortened the first version of the scale from 31 items to 11 items; for the scale measuring teachers’ PU of gamified learning tools, the survey items have been reduced from 27 to 15; for the survey focusing on teachers’ PEoU of gamified learning tools, the number is 23 to 11. The specific process can be found in the section of Findings and Appendix.

CHAPTER 6

Findings (1)

06

Teacher perceptions

It has been concluded that many top-down attempts to introduce innovative technologies in education have failed to achieve long-term effects (Cuban, 1986), and one reason can be the lack of awareness of teachers attitudes. Teacher attitude plays a crucial role in technology acceptance: if teachers do not see the value of the technology, the tool selection and implementation would be significantly restrained (Ale & Chib, 2011; Huizenga et al., 2017; José Martí-Parreño et al., 2016).

This chapter presents findings and discussion of the qualitative data collected from the interview on secondary school teachers in China. The interview responses were analysed with the use of the grounded theory techniques. This chapter aims to answer the first research question (RQ):

RQ 1: What are the factors contributing to teachers' acceptance intention of gamified learning tools in secondary schools?

Experts' comments were also presented.

CHAPTER SIX: TEACHER PERCEPTIONS

6.1. Teacher perceptions: Findings of the grounded theory analysis

Findings of the interview response analysis regarding teachers' perceptions of gamified learning tools in secondary schools are detailed below and in the following two tables (Table 29 and Table 30).

Table 29. Factors contributing to teachers' acceptance intention to gamified learning tools

Factor	Theme	Survey responses (out of 368)		Interview responses (out of 113)	
		F	%	F	%
Teacher factor	Teacher control	11	3%	8	7%
	Teacher skill	0	0%	6	5%
	Teacher attitude	0	0%	5	4%
	Risk avoidance	4	1%	4	4%
	Effort expectancy (time)	0	0%	5	4%
Tool factor-educational perspective	Resource availability	3	1%	12	11%
	Compatibility	10	3%	11	10%
	Curriculum relatedness	41	11%	7	6%
	Perceived usefulness on process (PU-Process)	39	11%	14	12%
	Perceived usefulness on academic outcomes (PU-Product)	68	18%	7	6%
	Cost-effectiveness	8	2%	6	5%
Tool factor-IT perspective	Perceived ease of use	82	22%	4	4%
	Cost	16	4%	2	2%
	Negative effects	36	10%	6	5%
	Gamification design	7	2%	2	2%
	Interface aesthetics	16	4%	0	0%
External support	Students', parental and administrative support	27	7%	14	12%

6.1.1. Teacher factor

Teachers' perceived control. The responses also revealed that teachers were in great concern of teacher control in adopting gamified learning tools (n=11). Teachers were sceptical to students' self-control, so they worried whether the students will use the digital devices for other purposes, whether the students will be distracted from learning content, and whether the students will be addicted. Two responders expressed that they expect anti-addiction or anti-distraction functions such as parental or teacher control. Parental or teacher control refers to a setting that allows parents or teachers to restrict the access of content to the students, which roughly be achieved by filtering content, limiting usage, enforcing the use of certain software and monitoring the usage (Baumrind, 1966).

While teachers in the survey focused on device usage control, teachers in interviews raised concerns about teacher control on the learning process. In face-to-face classroom teaching, teachers tended to refuse using tools with fixed knowledge system, in which teachers' roles are marginalised. In another word, a gamified tool that can be used for independent learning is not preferred by secondary school teachers. Participants expressed their preference to "design(ing) the pedagogical content myself" (P2) because "I cannot accept being dominated by a tool" (P7). However, "some gamified learning tools require users to unlock the content one by one", which "makes me feel lost control" (P14). When the involved tools are with pre-set content, teachers showed that they felt lost control of the learning content provided by gamified learning tools. They commented that:

> (P14) "Some gamified learning tools contain a lot of information; however, they require users to start from the beginning and unlock levels one by one."
>
> (P7) "If the gamified learning tool is just an educational option to assist my teaching, I will use it. This is essential. I cannot accept being dominated by a tool." And
>
> (P2) "I wish to design the pedagogical content myself."

Teachers also reported that the use of gamified learning tools in the classroom brings challenges in discipline management. P1, P4 and P6 complained about the issue, while P6 specified it as:

> "Different from some countries, we have large-scale classes that contain around 60 students. There are so many students that we even need to keep them unexcited: once they get excited, the classroom management will be a disaster."

Teacher skills. Findings revealed that there is a lack of teacher knowledge or skills on gamification and its implementation among the participants. P13 from rural schools stated that he never heard of this concept, neither his colleagues. P6 declared that she does not have the latest knowledge about gamified learning tools, so she only uses the tools she used when she was a student, which would be outdated: "students of the new generation know more about gamified learning tools than us". P8 was concern about how to design a gamified class, what tools can be used and how to implement it successfully. P14 added that "we don't have teacher training on gamification. I feel nobody knows how to implement gamification properly."

Teacher attitude. Not all teachers are holding positive attitudes to the implementation of gamification in secondary schools. Some teachers, even those who frequently implement gamification in the classroom, stated that they are not in favour of implementing gamified learning because its philosophy could be against their teaching value established in the test-oriented educational system (e.g., China). P1 proposed that: "Rigorous learning in test-oriented systems is effort-demanding, and gamification is weakening the seriousness needed for rigorous learning." For the same reason,

P14 viewed gamified learning tools as a compromise to unmotivated students, which is to provide "distractive attractions" to trigger more behavioural inputs. With the belief that high-performance or self-regulated students do not need gamification to get motivated, P1 showed a negative attitude towards the use of gamified learning tools by saying "we need to cultivate students' ability to confront frustrations as well as to discipline themselves from unnecessary distractions. We cannot compromise when students wish to give up easily."

There are gamified tools that enable teachers to customise teaching content for specific contexts; however, they can also be rejected by teachers since the use of this type of tool is time-consuming and effort-demanding. Participants indicated that

> (P3) "Teachers are not willing to use gamification. The reason is simple: the implementation of gamified learning increases teachers' burden, as it requires teachers to think about when to use the tool, how to use the tool and how to achieve a balance between fun and academic achievement. Teachers need to be well-experienced in classroom management, content delivery and gamification design."

> (P7) "If I really want to achieve satisfactory results using gamified learning tools, I need much more time. I will have no time at all, even for sleeping."

It is notable that some teachers, even those who frequently implement gamification in the classroom, stated that the gamification philosophy could be against their values. An English teacher proposed that:

> (P1) "To be honest, I am disgusted. Rigorous learning in test-oriented systems is effort-demanding, and gamification is weakening the seriousness of rigorous learning."

Therefore:

> (P1) "We need to cultivate students' ability to confront frustrations as well as to discipline themselves from unnecessary distractions. We cannot compromise to students when they wish to give up easily."

What's more, participants complained that much more time and effort are needed for implementing a gamified class. None of the participants believes their schools provide correspondent incentives, so they tended to hold a negative attitude to this new approach.

Risk avoidance. Survey responses did not cover teachers' intention for risk avoidance in deciding whether to adopt gamified learning tools, which was an obvious theme in interview responses. Interview responses revealed that the teachers felt obligated to maximise the learning time efficiency for actual improvements on performance, especially those with decades of teaching experience. One participant commented that the secondary school students in China need to take six or nine subjects, so the time allocated for each subject is highly limited:

> (P8) "We will feel guilty if we require the students to spend too much time on one single subject (so we need to make good use of the time, therefore the time spared for gamified learning is limited)."

Almost all the experienced teachers with five or more years of teaching experience said the new teachers would like to try innovative technologies such as gamification; however, due to the lack of experience, in most cases, the new teachers are "only able to deliver classes with low lesson capacity, which leaves teaching goals unfulfilled" (P7). As a result, implementing gamification in secondary schools could be reviewed as a "risky" attempt.

Several teachers asserted that they "will not consider gamification" due to the worry of risking students' academic performance. They stated that:

> (P3) "Now the mainstream learning in China is still exam-oriented. Under this circumstance, gamified learning is not as effective as the traditional teacher-centred instruction. You must admit that repetitive practices play an important role in improving academic performance for *Gaokao* (the National Higher Education Entrance Examination in China)."

One participant confessed that he did not intend to use gamified learning tools because he is afraid of "being different from others", which "makes me feel insecure" (P2).

It is reported that teachers' personal innovativeness is also connected with risk avoidance. One participant (P2) indicated that each individual is different when comes to their openness to new technologies, as only a small portion of people would like to be the innovators actively try new technologies, while a large portion of participants only choose to adopt new technologies when they see other people are already using them.

The aforementioned phenomenon could be explained with the theory of subjective norms or social image. Subjective norms is one core construct in the TAM2 model, which refers to the belief that an important person or group of people will approve and support a particular behaviour (Venkatesh & Davis, 2000). Subjective norms are determined by the perceived social pressure from others which push an individual to behave in a certain way to comply with others' expectations (Fisbein & Ajzen, 1975; Ham et al., 2015). According to the interview responses, the subjective norms for the participants include helping students achieve a high score in exams, spending more time on rigorous learning and delivering classes with high lesson capacity. Secondary school teachers also wish to make learning fun; however, if the use of new approaches or new tools challenges the aforementioned subjective norms, they would feel at risk and therefore decrease the intention of using the new approaches or tools.

6.1.2. Tool factor- educational perspective

Resource availability. Though not covered in the survey responses, the majority of the interview responders mentioned the availability of gamified learning tools as the factor influencing their acceptance intention (12 out of 14, 86%).

Firstly, teachers reported that there is a lack of *mature* gamified learning tools. As reported by P11, "the blueprint of using gamification to engage learners is great; however, many designers fail to achieve what they wish to achieve".

Secondly, there is a lack of *suitable* gamified learning tools. P9 expressed that "not all the gamified learning tools are suitable to be used in the classroom", therefore "even if I had a thought of a gamified course, I might not find the tool; even if I found a tool, it might not suitable for my teaching". Since teachers are not expected to be equipped with technical knowledge to adapt the gamified learning tools, while there is a lack of gamified tools that allow teachers to make adaptations, the implementation of gamification in the classroom is greatly impeded (P10).

Thirdly, there is a lack of complex and multi-model tools: even though some tools are well-developed, they can only achieve limited functions to meet the needs of real-life test-oriented education. For example, in language learning, a large number of gamified learning apps work on word memorization, while few of them help reading comprehension, writing skills or literature appreciation. Take P6's responses as the example: "in language learning, a large number of mobile apps for word memorization are well-developed; however, it is very difficult to find one that serves for listening practices, sentence patterns and overall comprehension".

Last but not least, participants also reported the availability of gamified learning tools in rural areas (P12).

Compatibility. The next theme that emerged from survey responses was the compatibility of gamified learning tools (n=10). Three statements covered the basic meaning of compatibility, indicating that the use of gamified learning tools in schools might meet difficulties when coming to different platforms or devices. For example, the content in mobile apps might not able to be easily shown on the big screen.

Researchers also explore the wider meaning of compatibility of gamification, such as the compatibility of gamification to several elements of the curriculum, including the learning objectives, learning content, teaching-learning process, assessment and evaluation (Yıldırım & Demir, 2016).

The survey responses revealed that teachers have come across issues about the compatibility of gamified learning tools with the teaching schedule, learning content and teachers' existing knowledge. Teachers reported that they concern whether the gamified learning tools are compatible with the teaching schedule, as using gamification in the classroom might need one hour or more but the class duration is set as 40 minutes only. What's more, there might be "conflicts" when the pedagogical content provided by the learning tools might not align with the curriculum, which leads to "a waste of valuable classroom time" (P4). Though only mentioned by one participant, the compatibility of gamification to teachers' existing knowledge needs further attention.

What was not covered in the survey responses but in the interview responses is the compatibility of gamified learning tools with varied subjects and different types of knowledge. As responded by participants, "we cannot gamify any type of knowledge in any subject" (P6) because "each subject is definitely different from others" (P3). This issue becomes even more obvious when it comes to abstract knowledge, such as grammar and Mathematical formulas. P11 mentioned that he prefers to use gamified learning as an "ice-breaking activity" at the beginning of classes to attract

attention but rarely use it to teach key knowledge because the knowledge points are not suitable or not easy to be gamified.

Another participant proposed that the drop-out rate of gamification use will be very high when the content is test-oriented, because "test-oriented learning requires rigorous learning that involves a large number of practices. Once lots of practices are needed, gamification will be less motivating or even demotivating" (P9). In other words, gamification and test-oriented learning could be highly incompatible. An English teacher (P2) proposed to treat gamified learning with caution as "it is not suitable for knowledge-intensive classes or practice-intensive training", while some others (P3 and P7) stated that the use of gamification or gamified learning tools in the classroom is highly possible to "lower our teaching efficiency". Therefore, even though they are positive in using gamification "outside the school (in tutoring classes)" or "for early childhood education", they "will not consider using it in school time at all".

One participant (P7) commented that gamification could also be incompatible with public schools due to procedural issues. She stressed the difficulties in implementing gamification top-down from the Ministry of Education because:

> "The development of gamified learning tools is highly demanding for technical and financial supports, so when considering implementing gamification in schools, public school teachers need to apply a national or at least a provincial project that was assessed and followed up by the Ministry of Education. The project initiated by the Ministry of Education involves lots more considerations (such as content universality, the supply of supporting facilities, the validation of tool effectiveness, etc.), so the development of gamified learning tools will be largely limited."

She furtherly pointed out that the development of gamified learning tools can not only rely on the government. Gamification can only thrive when a large number of first-line teachers who directly leading the teaching-learning process take part in the development and implementation of gamified learning. She showed that private schools could be contributors to gamified learning in China since they could be both wealthy and more flexible in making decisions. Anyway, gamification seemed not compatible with public secondary schools in China.

Curriculum relatedness. As highlighted by De Grove et al. (2012), the integration of gamification features with pedagogical content is always a challenge for gamified learning, which was also valued by participants (n=41). Four participants mentioned their wishes of combining entertainment with general education, while all others (n=37) specified their expectations of integrating game or gamification features with the curriculum. The participants revealed a strong demand for high curriculum relatedness of the gamified learning tools, as a large number of the participants specified "good match" (P1) or even "strong link to each chapter" (P5) of the textbook.

The importance of curriculum relatedness was confirmed in interview responses as well. Teachers said that to achieve successful gamified learning, the designers should fully understand both the gamification features and the curriculum, summarise the commonalities, make connections and then make the knowledge 'gamifiable' (able to be gamified), but "few of the designers are equipped with adequate knowledge or skills". As stated by P6, in designing gamified learning tools, "those who know how to design do not fully understand the curriculum, while those who know the curriculum do not fully understand gamification design", therefore the curriculum relatedness of the product is not satisfactory.

The lack of curriculum relatedness impacts the final products. If design the tool from the gamification perspective, which involved lots of digital-game design techniques, the final products are highly possible to be ambiguous in

indicating the educational knowledge points, or they could be tools with 'pointsfication' (the superficial gamification that involves points, badges and leaderboards only). As a consequence, the use of gamification features is perceived as a weakening of pedagogical purposes (P4, P9 and P11) and teaching efficiency (P7).

Overall, all the responders expressed strong expectations on the curriculum relatedness of gamified learning tools; however, at the same time, they showed dissatisfaction with the development of gamified learning tools on this aspect.

PU-Process. As gamification is promising in engaging learners, participants were also in favour of its effectiveness in engaging students with fun (n=39). They expected the gamified learning tools to create a fun experience, trigger interests, enhance learning positivity and eventually engage the learners.

Revealing the same trend, up to 84.6% of the participants (11 out of 14) participants reported that the benefits of gamification in education is "obvious and well-known" via adding fun to the tedious classroom. It "attract(s) attention" (P6), "engage(s)" students (P1, P3 and P7), brings "authentic experience" (P4), and most importantly, brings "fun" (P2, P4, P5, P6, P8, P10, P11 and P12). P3 who has had a fully gamified gamification experience was highly positive about the implementation of gamified learning approach or the tools in the classroom. He said that the "*truly gamified gamification*" triggers participation, immersion and flow status, in which the users could be so involved that they could even feel a reduced awareness of self or time (Csikzentimihalyi, 1975; Hamari & Koivisto, 2014). P3 described that:

> "My positive attitude toward gamified learning originates from my own gamification experience 20 years ago. Under the supervision of our teachers, we spent one hour playing a board game designed to teach market operation rules. I learnt how to cooperate with my teammates to achieve a win-win situation. The point is, I learnt it not by reading books but by participating in the simulation situation. It triggered my inspiration and active thinking. The experience is so impressive that I can even recall every scene of the gamification process before sleep. Now 20 years have passed, and I still remember what I have learnt in that class."

It is notable that even though P3 was strongly interested in gamified learning, he admitted that gamified learning tools are not suitable for every subject. P3 said that "at least for my subject (Physics), I found it difficult to find or create a *truly* gamified learning tool which is *truly* engaging."

P1 highlighted that though teachers are expected to make learning fun, in reality, many teachers fail to do so. He confessed that:

> "In most cases, teachers are just reading the handout notes without creating a fun learning experience, especially when it comes to classes with large sizes."

P6 commented that:

> "A very significant drawback of our current teaching system is that the teaching nowadays is basically teacher-centred lectures: teacher talks, students listen, and nothing else. Gamified learning could possibly address this issue."

PU-Product. A large number of survey responses covered the theme effectiveness on outcomes (n=60) with the keywords "learning results", "academic performance" and "improvement on scores". P5 in the interview confirmed the importance by stating "effectiveness in improving academic performance is the key concern", and "none of the

other issues will be a problem once effectiveness is guaranteed". Other participants also claimed that if they are sure about the effectiveness, they will "definitely use it" (P1 and P7).

However, it is notable that interview responses revealed that the participants were not optimistic on this issue. The pessimism underlying interview responses raised the question "why gamified learning is promising in academic research but reported problematic by secondary school teachers". Participants stated:

> (P5) "Besides engaging learners, I wish the tool to achieve pedagogical goals. However, this blueprint has not been achieved since I started my career14 years ago".
>
> (P3) "When a gamified learning tool gets published, other people only need to care about whether they are fun enough. Teachers are not. We need to prove that they are of potential in bringing educational benefits. However, I cannot convince myself or others that the tools are really helpful in improving academic performance." And
>
> (P7) "Whether gamified or not, an educational tool needs to be connected closely with teaching aims. If it claims to be able to enhance the memorial rate, I expect it to achieve it. However, few products can achieve what they promised."

Cost-effectiveness. In general, participants in the current study held negative viewpoints in terms of the effectiveness of gamified learning in schools. To take a closer look at the responses, this author found a deeper theme: cost-effectiveness, which refers to "the ratio of effect to cost" (Michie, Van Stralen, & West, 2011, p. 23). As figured out by P14, "if given unlimited time and resources, I can assert that almost any educational approach or tool is helpful. The problem is we don't have unlimited time or resources, so we need to consider the cost".

Nine survey participants criticised the "efficiency" of using gamified learning tools without specifying detailed reasons. Interview responses added details.

Firstly, teachers doubted the cost-effectiveness of gamified learning tools in terms of students' time input:

> (P6) "Our students need to study for about 10 hours a day. They don't have enough time. However, to my knowledge, some gamified learning tools are sacrificing efficiency for fun." And
>
> (P3) "A *truly* gamified activity requires learners to experience it for multi times to gain the in-depth understanding of the key idea. It would be significant, profound and long-lasting; however, it is time-consuming, too.

P7 reported that the cost-effectiveness issue applies to teaching, too, so she would not use gamification in secondary schools even though "it is fine to implement it in tutoring classes or early-childhood stage", which do not occupy the valuable classroom time in formal education.

Also, one teacher with management experience complained about the cost-effectiveness of gamified learning in terms of financial input (P7). P7 stated that if she has the budget, she prefers to use it for "second class". Second class refers to the meaningful extracurricular activities that are organised for academic improvement, intellectual development, fitness, entertainment or public warfare (Zeng & Zheng, 2017). P7 explained:

> "If I want the student to be happy, I choose to pay for physical training, psychological counselling, or anything that presents me with tangible feedback. Compared with those, gamification is too costly regarding both human and financial resources.
>
> There are many alternatives to achieve the same results (on engaging students), such as changed tone, eye contact and exaggerated body movement. Compared with gamified learning tools, these are more reliable in bringing concrete feedback. Since there are various approaches to achieve the results, I have no reason to stick to the one that is not mature yet (gamified learning)."

Also, teachers complained about the cost-effectiveness of gamified learning in terms of financial input (P7) and learner engagement (P6). P7 stated that if she has the budget, she prefers to use it on "second classes", which refer to the meaningful extracurricular activities that are organised for academic improvement, intellectual development, fitness, entertainment or public warfare (Zeng & Zheng, 2017). She explained that the investment in the "second classes" such as physical training and psychological counselling provides more tangible feedback while costing less on human and financial resources. P6 stated that she has no reasons to stick to the not-mature-yet gamification since there are many alternatives to achieve the same results on engaging students, such as "changed tone, eye contact and body movement". Compared with gamification, "these are more reliable in bringing concrete feedback".

6.1.3. Tool factor- IT design perspective

Overall perceived ease of use. The survey responses revealed that secondary teachers value the overall perceived ease of use in deciding whether to adopt the gamified learning tools. Among the 368 survey statements, 64 pieces were reporting the teachers' expectation of having easy-to-use tools. Statements in this section were comparatively superficial, as the participants left simple comments such as "easy to use", "simple", "convenient" and "not complicated". Participant 6 explained that when using gamified learning tools in the classroom, there will be troubles in account registration and management; when involved varied tools, the different platforms will also be exclusive with each other, which also brings ease-of-use issues.

Notably, participants in the interview highlighted the ease of use in *implementing gamification as a whole*, rather than the ease of use of *technical issues* only. Previous scales evaluating the perceived ease of use (PEoU) mainly focus on the PEoU of technical issues, such as whether the tool is rigid and inflexible to interact with, whether the user finds it cumbersome to use the tool, and whether the user finds it easy to get the tool to do what he or she wants it to do (Davis, 1989; Venkatesh & Davis, 2000).

The interview responses showed teachers' needs for the ease of use of the tool from a broader perspective. The participants expected "a mature and actionable gamified curriculum", which enable the teachers to implement gamification "directly without too much preparation" (P3); they also expected a can-be-directly-used tool, which does not need teachers or students to learn how to use it, nor experience the complicated login-in or opening effects (P6).

Affordability (cost). Affordability indicates that a tool or an intervention is "affordable if within an acceptable budget it can be delivered to, or accessed by, all those for whom it would be relevant or of benefit" (Michie et al., 2011, p. 23).

In the survey, 16 responders enlisted cost as the keyword. The cost issue includes "free to use", "no advertisement" and "no need to top up for further information". Unexpectedly, only a limited number of the interview respondents were concerned about the cost of the gamified learning tools. When asked the barriers of implementing gamified learning tools in school, only two participants mentioned the affordability of the gamified learning tools by saying "the price of the tool will be one concern" (P10) and "I hope the tool is not costly" (P9), with the rest participants saying the cost is not a problem if the gamified learning tools are helpful in pedagogical activities.

Side effects. Side effects refer to the unintended consequences of the use of a tool or an intervention (Michie et al., 2011). Thirty-six survey responses mentioned the side effects of using gamified learning tools, including "limited function" (n=1), "digital addiction" (n=7), "distraction" (n=9), "privacy risk" (n=1), "eyesight damage" (n=8), "the involvement of undesirable or vulgar information" (n=2), "harmful radiation" (n=1) and the general negative influence to learning (n=7).

Besides, the interview responses revealed more in-depth side effects including the downplayed learning aims and weakened intrinsic motivation:

> (P5) "Students will be distracted from learning to game features. This is the most obvious disadvantage of using gamification in my classroom."
>
> (P9) "It distracts learners from learning."

(P1) “It decreases students’ endurance to independent learning.”

(P2) “It highlights points-earning rather than learning itself.”

(P9) “When contained a lot of information, it hurts students’ eye-sight.” And

(P3) “It could be somewhat addictive.”

Gamification design. Seven survey responses revealed teachers’ expectations of using gamification design in educational tools. The participants specified “competition”, “check-in function”, “challenge”, “points system”, “leaderboard” and other game elements or gamification mechanisms. P7 in the interview highlighted the use of gamification design rather than game elements only.

Interface aesthetics. In the survey, 16 participants expressed their expectations on interface aesthetics by commenting on the keywords “beautiful” and “pleasant”. Notably, no interview responses covered the theme interface aesthetics. The underlying reasons need further exploration.

6.1.4. External support

Support from parents and students. Parents play an unignorable role in hindering the development of gamification in education. Teachers in the current study indicated that parents are "highly possible" to limit students' access to digital devices or the internet since they worry about digital distractions or addictions. To force their children to focus on study, some parents, even those from wealthy families, refuse to buy smartphones or tablets.

Unexpectedly, the survey and interview responses revealed that students are also likely to refuse to use gamified tools for learning. The participant (P14) expressed that "we assumed that gamification will make students engaged. The reality says no" because the students in secondary schools are also under great pressure, so they tend to be utilitarian and wish to get quick benefits on academic performance. Participants commented:

> (P14) "We assumed that gamification will make students engaged. The reality says no."
>
> (P) "Students in secondary schools are also under great pressure, so some of them choose to be utilitarian: they are not patient to something without bringing instant benefits."
>
> (P11) "Play is to have fun, education is to reshape. Students would like to sacrifice entertainment for their goals."

Administrative support

School principals held negative viewpoints on gamified learning tools. P6 said that the leaders might be sceptical of new technologies, so they prefer to spend money on other things which provide them concrete feedback (e.g. psychological counselling, sports training). Schools are also not providing sufficient investments in equipment and technical support, which strongly limits teachers' implementation of gamification in classrooms.

Table 30. Teacher perceptions: Findings emerged from interview responses

Construct	#	Theme/Factor	Sub-theme and illustrative data
Teacher factor	1	Teacher control	"It is difficult to monitor or limit what students **use** the devices for. I want full control of the teaching process." (P1) "I would like to design the **content** myself" (P2)… "I cannot accept being dominated by a tool." (P7) "…the **classroom management** will be a disaster." （P6）
	2	Teacher knowledge and skills	Knowledge about gamification and skills in implementing it.
	3	Teacher attitude	"It **distracts** students from **learning itself** to game features." (P12) "It's **time-consuming and effort-demanding**." (P14)
	4	Risk avoidance	"Normally the new teachers who would like to use gamified learning tools only able **to deliver classes with low lesson capacity**, which leaves teaching goals unfulfilled." (P7) "Under this circumstance (test-oriented educational system), gamified learning is **not as effective** as the traditional teacher-centred instruction." (P3) "If I use gamification, I will **be different from others**. I am scared of it." (P4)
Tool factor-educational	5	Resource availability	(A lack of **mature** tools): "The blueprint of using gamification to engage learners is great; however, many designers fail to achieve what they wish to achieve." (P11) "Not all the gamified learning tools are **compatible (to the contexts)** to be used in the classroom", therefore "even if I had a thought of a gamified course, I might not find the tool; even if I found a tool, it might not suitable for my teaching." (P9) (A lack of **complex and multi-modal** tools): "In language learning, a large number of mobile apps for word memorization are well-developed; however, it is very difficult to find one that serves for listening practices, sentence patterns and overall comprehension." (P6) A lack of resources in rural areas: "I am working in a **rural school**. We don't have resources like this." (P12)
	6	Compatibility	"I need to follow a ***pre-set teaching schedule***, so I don't have time to implement gamification in the class." (P8) "We cannot gamify ***any type of knowledge*** in any subject." (P6) because "each ***subject*** is definitely different from others." (P3)

			"***Test-oriented learning*** requires rigorous learning that involves a large number of practices. Once lots of practices are needed, gamification will be less motivating or even demotivating." (P9) Compatibility to ***public schools***: "Private schools … could be both wealthy and more flexible in making decisions." (P7)
	7	Curriculum relatedness	"Normally the gamified learning tools are good in engaging students; however, most tools cannot help achieve **educational outcomes** we expected in the **curriculum**." (P4) "A good gamified learning tool needs to be **strongly related to each chapter** of the textbook." (P9)
	8	PU-Process	"What motivates me to use gamification is it really engages students." (P5)
	9	PU-Product	"Effectiveness in improving academic performance is the key concern." (P5)
	10	Cost-effectiveness	"Our students need to study for about 10 hours a day. They don't have enough **time**." (P3) "I choose to pay for physical training, psychological counselling, or anything that presents me tangible feedback. Compared with those, gamification is too costly regarding both **human and financial resources**." (P7) "I choose to **engage** students in other ways, such as changing tone." (P6)
Tool factor-IT perspective	11	Perceived ease of use	(Ease of use on **technical issues**): "It needs to be easy to use." (P7) (Ease of use in **implementing gamification as a whole**): "I wish to have a mature gamified curriculum that is ready-to-use in the current contexts." (P6)
	12	Cost	"The price of the tool **would be one concern**." (P10) "Price **would not be a concern** as long as the tools are helpful." (P1)
	13	Negative effects	Digital addiction, distraction, privacy violence, eyesight damage, the involvement of undesirable or vulgar information, harmful radiation, decreased endurance to independent learning.
	14	Gamification design	e.g. "competition", "challenge", "PK mode (player killer mode)"
	15	Interface aesthetics	e.g. pleasant interface
External support	16	Parental, students' and administrative support	"To force their children to focus on study, some **parents**, even those from wealthy families, refuse to buy smartphones or tablets." (P4)

	"**Students** in secondary schools are also under great pressure, so some of them choose to be utilitarian: they are not patient to something without bringing instant benefits." (P1) "We don't have sufficient **equipment or technical support**." (P1) "Principals could **be sceptical** to new technologies." (P6)

6.2. Discussion: Factors contributing to teachers' acceptance intention towards gamification

6.2.1. Teacher factor

Previous studies stressed that teachers' negative attitudes and limited teacher skills are factors hindering the implementation of GBL or digital gamification in classrooms (De Grove et al., 2012; Emel'yanenko et al., 2016; Koh et al., 2012; Tuparova et al., 2018). Survey and interview responses in the current study reported two new themes, including teacher control and risk avoidance.

The results revealed that teachers concerned about losing control in the classroom. GBL tools and gamified learning tools are providing students with an option of learning that enables students to do without a teacher in the educational process (Emel'yanenko et al., 2016; Sánchez-Mena & Martí-Parreño, 2017); accordingly, some authors believe that gamification changes the role of the teacher from leaders to partners, which weakens teachers' authority (Emel'yanenko et al., 2016; Tuparova et al., 2018). For teachers who do not have sufficient knowledge, the use of technological innovations can be a cause of anxiety, as well as GBL or gamified learning (Ketelhut & Schifter, 2011; Sánchez-Mena & Martí-Parreño, 2017; Smith, Stair, Blackburn, & Easley, 2018). Therefore, gamified learning tools should be categorised as those with and without real-life teacher-student interactions. For the tools which are expected to be supportive in real-life teacher-student interactions, there is a problem whether the gamification designers should address teachers' needs, or the teachers need to adjust their attitude to adapt to the new tools. The former one requires more studies to investigate teachers' needs and expectations, while the latter one needs professional development on gamification knowledge.

Risk avoidance is a new term, which explains the theory of subjective norms. Subjective norms is one core construct in the second version of the technology acceptance model (TAM-2), which refers to the belief that an important person or group of people will approve and support a particular behaviour. Subjective norms are determined by the perceived social pressure from others which push an individual to behave in a certain way to comply with others' expectations (Wong, 2016). According to the interview responses, the subjective norms for the participants include helping students achieve a high score in exams, spending more time on rigorous learning and delivering classes with high lesson capacity. Secondary school teachers also wish to make learning fun; however, if the use of new approaches or new tools challenges the aforementioned subjective norms, they would feel at risk and therefore decrease the intention of using the new approaches or tools.

> Gamified learning tools should be categorised as those with and without real-life teacher-student interactions.

6.2.2. Tool factor- IT perspective

The technical features of the tool from the information technology (IT) perspective were also highlighted. Secondary school teachers expect the gamified learning tools to be good-looking, which have been rarely covered in previous studies investigating the predictive factors of teachers' adoption intention of gamification. Notably, these two factors have been covered in literature in other fields; for example, in visual design, aesthetics is one core design principle that defines a design's pleasing qualities by involving factors such as balance, colour, movement, pattern, scale, shape and visual weight.

In gamification design, aesthetic is one of the three main constructs in the Mechanics-Dynamics-Aesthetics (MDA) framework for game design (Hunicke et al., 2004), which emphasises the providence of desirable emotional responses (e.g. challenge, competition, discovery, fantasy and sensation) in the game players when they interact with the game system. Interface aesthetics are frequently mentioned in the field of game design, but not in the studies on gamification. The underlying reasons need further exploration.

Perceived ease of use is an important factor in the technology acceptance model (TAM). As stressed by Zhao and Cziko (2001) and Nicolle and Lou (2008), users are interested in ease-to-use tools that will not cause disturbances while implementation. Notably, participants in the interview highlighted the ease of use in implementing gamification as a whole, rather than the ease of use of technical issues only. Previous scales evaluating the perceived ease of use (PEoU) mainly focus on the PEoU of technical issues, such as whether the tool is rigid and inflexible to interact with, whether the user finds it cumbersome to use the tool, and whether the user finds it easy to get the tool to do what he or she wants it to do (Davis, 1989; Venkatesh & Davis, 2000). In this study, teachers' needs for the ease of use of the tool come from a broader perspective. The participants expected "a mature and actionable gamified curriculum", which enable the teachers to implement gamification "directly without too much preparation" (P3). Therefore, when measure teachers' perceived ease of use (PEoU) of gamified learning tools, researchers need to consider not only the PEoU from the technical perspective but also the implementation perspective.

6.2.3. Tool factor- educational perspective

Analysing the tool factor from the educational perspective, both literature and findings highlighted resource availability, learning opportunities and curriculum integration. Survey results extended the concept of "resource availability" from the availability of gamified learning tools to the availability of *suitable* gamified learning tools which are compatible with the teaching content and convenient to use. The findings also extended the concept of "compatibility" from the compatibility of curriculum to the compatibility of teaching schedule, type of knowledge and type of schools (public or private).

Results showed that secondary school teachers in China were dubious about the capability of gamification in providing learning opportunities. Researchers indicate that teachers are concerned about standards, grades and outcome measures, so the gamified learning tools that cannot directly serve for enhanced academic performance are viewed as tools with low educational potential (Demirbilek & Tamer, 2010). There is also a lack of assessment measuring the capability of GBL or gamified learning tools in providing learning opportunities, as well as the problems of transferring skills learned in digital devices into real-life educational contexts (Emel'yanenko et al., 2016; Huizenga et al., 2017). This study confirmed the view that even though the teachers do "have the impression" that their students learned from GBL or gamification, they hesitate to assert the educational potential of the tools in their classes (Huizenga et al., 2017, p. 111).

Curriculum relatedness is another issue. If the tool is designed from the gamification perspective, which involved lots of digital-game design techniques, the final products are highly likely to be ambiguous in indicating the educational knowledge points, or they could be tools with 'pointsfication' (the superficial gamification that involves points, badges and leaderboards only). As a consequence, the use of gamification features is perceived as a weakening of pedagogical purposes and teaching efficiency.

Findings stressed participants' concerns on the effectiveness of gamification in education, which align with previous studies (Huizenga et al., 2017). The adoption of gamification as a new technology depends on whether it addresses teachers' priority needs, which is to deliver quality instruction and effectively achieve teaching goals (Kurt, 2012;

Zhao & Cziko, 2001). When it comes to gamified learning, researchers highlight its effectiveness in engaging or motivating students, too. However, in this study teachers felt that given the schooling context students may be less motivated by gamified learning as it did not seem to add value to their learning outcomes. The literature has rarely talked about effectiveness or cost-effectiveness; instead, they involved concepts such as learning opportunities and curriculum integration (Bourgonjon et al., 2013; De Grove et al., 2012; Ibrahim & Jaafar, 2011; Koh et al., 2012). One possible explanation is that the effectiveness of gamification for learning remains controversial (Silpasuwanchai et al., 2016), as a considerable number of researchers note that there are methodological limitations in most existing studies of gamification in education, including short experiment timeframe, small sample size, absence of control groups (Y. Chen et al., 2015; De-Marcos et al., 2014; Seaborn & Fels, 2015; Stanculescu et al., 2016); accordingly, "there is little empirical evidence to support their effectiveness" on engaging students or improving academic performances (Denny, 2013, p. 763). Cost-effectiveness is a concept furtherly explaining the effectiveness of gamified learning, which compares costs and outcomes of different behaviours, actions or interventions (Michie et al., 2011). The costs include time input, effort input and financial investment. Future studies should pay more attention to the measurement of gamification's effectiveness in learning and factors contributing to its cost-effectiveness.

6.2.4. External support

When comes to external support, related literature highlights institutional support (Ale & Chib, 2011; Beggs, 2000; Khlaif, 2018; Nicolle & Lou, 2008; H.-K. Wu et al., 2008). Administrative support includes financial support, infrastructure providence and technical assistance, which involves various factors such as devices, Wi-Fi, local network, power supply and technical assistance (Beggs, 2000; Khlaif, 2018). Align with the literature, findings revealed that the policy support from school principals is also playing a role in promoting the acceptance intention of gamified learning (Koh et al., 2012; Sánchez-Mena & Martí-Parreño, 2017).

CHAPTER 7

Findings (2)

07

Scale development

The TAI-GLT scale

This section reports the process of developing a scale (TAI-GLT) measuring teachers' acceptance intention to gamified learning tools. The scale was developed based on the Technology Acceptance Model (TAM) proposed by Davis (1989).

CHAPTER SEVEN: CONSTRUCTION AND VALIDATION OF THE TAI-GLT SCALE

Scale development and evaluation (TAI-GLT)

7.1. Theoretical framework: Teachers' acceptance intention to gamification

7.1.1. Previous game-related and gamification-related scales

In previous studies, game-related scales garnered more attention compared with gamification-related ones, which of course contain items measuring obviously game-like features. For example, the User Engagement Scale (UES) aiming to measure engagement during video game-play contains "aesthetics" (beautiful interface designs) (Wiebe, Lamb, Hardy, & Sharek, 2014); the Video Game Pursuit Scale (VGPu) by Sanchez and Langer (2020) that enlists the pursuits attracting users to play video games contains the factor "gaming behaviour"; similarly, the Game User Experience Satisfaction Scale (GUESS) by Phan, Keebler, and Chaparro (2016) contains game-like elements "audio aesthetics", "narratives" and "play engrossment".

There are scales focusing on gamification rather than video games. Tondello et al. (2016) proposed six types of users of gamification, which focused on the characteristics and preferences of varied users (Philanthropists being motivated by purpose, Socialisers being motivated by relatedness, Free Spirits being motivated by autonomy, Achievers being motivated by competence, Players being motivated by extrinsic rewards, and Disruptors being motivated by the triggering of change). Liu et al. (2019) developed the Festival Gamification Scale (FGS) based on self-determination theory, which contains five factors: relatedness, mastery, competence, fun, and narratives. Nevertheless, the user-type categorisation by Tondello et al. (2016) focused on users rather than gamified learning tools, and the FGS scale by Liu et al. (2019) has not been validated in educational contexts.

Eppmann, Bekk, and Klein (2018) constructed and validated a gameful experience scale (GAMEX) that measures gameful experience in gamification. Though containing the word "game", gameful experience does not indicate the use of video games; instead, it refers to "the positive emotional and involving qualities of using a gamified application" "in a non-game context" (Eppmann et al., 2018, p. 100). The 27-item GAMEX scale consists of six factors, including enjoyment, absorption, creative thinking, activation, absence of negative affect, and dominance.

It is notable that the GAMEX scale was developed based on three game-related scales rather than gamification-related ones: the immersion questionnaire (IQ) by Jennett et al. (2008), the engagement questionnaire (GEQ) by Brockmyer et al. (2009), and the game experience questionnaire (GExpQ) by IJsselsteijn, Kort, and Poels (n.d.) that did not publish on peer-reviewed academic journals. The aforementioned scales contain items that relate to video games, such as "I was in suspense about whether I would win or lose the game" (IQ, item 5), "I sometimes found myself to become so involved with the game that I wanted to speak to the game directly (IQ, item 7), "the game feels real" (GEQ, item 5), "I play without thinking about how to play" (GEQ, item 15), and "I really get into the game" (GEQ, item 18).

What's more, the items in both IQ and GEQ are strongly connected with the flow state theory, a popular concept in video game studies. Proposed by Csikszentmihalyi (1990), flow refers to the mental state in which a person is fully

immersed in a sense of deep enjoyment, which has characteristics such as complete concentration on the task and transformation of time (speeding up/slowing down). For example: "I was unaware of what was happening around me" (IQ, item 21), "if someone talks to me, I don't hear them" (GEQ, item 6), "to me it felt like only a very short amount of time had passed" (IQ, item 31), "when playing the game time appeared to go by very slowly" (IQ, item 32), "I lose track of time" (GEQ, item 1), and "time seems to kind of stand still or stop" (GEQ, item 8). However, the flow state theory emphasises the optimal experience originated from the balance between challenges and one's competence, which is comparatively difficult to achieve in educational contexts that cannot provide detection of one's ability, nor the correspondent customised difficulty-adapting tasks. Therefore, the flow-based survey items need further validation in gamified learning contexts. Generally speaking, there is a lack of scales measuring gamification-related issues, not to mention gamification for educational purposes.

The selection of generality or specificity is an issue in scale development (DeVellis, 2003). For instance, if gamification is regarded as one general educational technology, there are a large number of frameworks that could be used or adapted as the theoretical foundation. If research contexts are taken into consideration (e.g., the implementation of gamified learning tools in test-oriented educational systems), there will be a lack of appropriate theory related to the measurement problem at hand. Though it is promising to develop a scale specifically based on gamification-related theories, the current study selected the other way (regard gamification as one general educational technology) due to a limited number of supportive frameworks. Accordingly, the current study mainly focused on the acceptance willingness of the general technology while considering specificity by contextualising the factors and survey items measuring them, which will be detailed in the next section.

7.1.2. Framework conceptualisation process

After viewing the existing frameworks for technology acceptance intention, this author did not identify one that applies to gamified learning in secondary school contexts. To fill the gap, we merged the aforementioned frameworks into one with the following six steps:

Step 1: conceptualise a theoretical framework. After selected publications of technology acceptance intention, this author contextualised the model by removing frameworks that are not significantly applicable for the objective of the current study, such as those describing the acceptance intention of pure digital games. Contextualization is necessary to ensure that the selected models and the corresponding measurement items are culturally acceptable and appropriate (Younas & Porr, 2018). At this stage, this author took a wider perspective to view our topic, attempted to find theories that contribute to the current topic but are not involved in the first step. This step is to avoid the poor conceptualization of the model caused by a narrow literature review (Younas & Porr, 2018). What this author added was the determinant "perceived risk" proposed by Deng et al. (2018), which was to predict patients' acceptance intention toward mobile health services. The newly-added determinant was reviewed by experts for validity, as well as be confirmed in teachers' interview responses.

Step 2: establish the theoretical framework. This author merged all the selected frameworks into a comprehensive one. At this step this author came across with difficulties unifying the terms, such as in some frameworks one construct of PEoU was named "effort expectancy" (Ibrahim & Jaafar, 2011; Venkatesh et al., 2003), while another one named it as "complexity" (Rogers, 2010); similarly, Lin and Chen (2013) used "knowledge learning", while other researchers used other words to describe the identical or similar issue, such as improved student learning (Beggs, 2000), learning opportunity (De Grove et al., 2012; Ibrahim & Jaafar, 2011), multimedia instruction (Cigdem & Ozturk, 2016), and useful-product (Bourgonjon et al., 2009). At this step, this author grouped terms with similar meanings into groups and named each group as a unit of the classification. When there was a term with a new meaning,

this author created one new unit, Then, this author moved back and forth between the term pool and the ongoing classification to identify whether each unit represents a particular communicative purpose. It is possible to have units with similar meanings with other ones, so this author also moved back and forth to merge the units with similar meanings into groups. The theoretical framework was confirmed when there were no new functional units added or removed.

Step 3: validate the theoretic framework. This author validated the formation of the new framework using the investigator triangulation method, which is to involve multiple observers to record and describe the data, cross-check the observations and agree about the final results (Johnson, 1997). This author employed two researchers to go through the framework-establishment process. The researchers were briefed on the methods and principles in merging frameworks into one, then work independently according to his/her understanding of the context. This was a dynamic process involving significantly varied categorization, filter standards and term selection. After obtained the final result, this author organised a meeting to discuss the different outputs and eventually reached an agreement on the final framework (see Figure 13).

Step 4: generate an item pool and eliminate redundancy. The item pool can be generated in a variety of ways such as expert interviews and a literature review of related research or questionnaires (Robertson, 2017; Younas & Porr, 2018). This study combined the two by analysing the themes of the interview responses, fitting the themes in the selected theoretical framework, finding related surveys, and selecting or removing survey items based on the interview responses. At this step, the interview responses as the benchmark enabled this author to select or dismiss survey items with reasonable justification. The interview responses also helped form the survey items to measure the variables without existing surveys, including the variables "perceived usefulness- process" (PU-Process), "complexity", and "perceived risks".

Step 5: Have the initial item pool reviewed by experts. Two researchers in the gamification field were invited to assess whether the items covered all facets of the construct of interest for content validity and whether the items look like measuring what they are supposed to measure for face validity. The experts also checked the items including whether the items are vague, whether the wording is adequate, whether the items pose two questions at the same time, whether the items use adverbs excessively, whether the items are easy to read for the audience and whether the items are conceptual relevant in terms of gamified learning (da Silva Brito et al., 2018; Jackson & Marsh, 1996; Younas & Porr, 2018).

Step 6: Translate the scale and collect data. Since the items are to be administered in Chinese other than the development language, a back-translation process was ensured. The questionnaire was available to the public via a web link, so that potential participants could access it and complete it in a convenient place and time.

Step 7: Evaluate and refine the survey. Once the data has been collected and cleaned, it is ready to be analysed. Statistical techniques enable researchers to establish the stability, validity and reliability of the survey by eliminating items that do not contribute additional information. The statistical techniques used in the current study are internal consistency analysis, correlation analysis, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), as will be specified in the section of Findings.

Besides the aforementioned four main steps, this author also involved other steps including determining the format of the measurement as a 7-point Likert survey, including validation items, conduct a pilot study, removing invalid surveys and optimising scale length, as suggested by DeVellis (2003) and Robertson (2017).

To guarantee the reliability of the qualitative analysis, this author employed the coding reliability approach, which is to use multiple coders to independently code the data and compare the results of different coders. This author used Cohen's Kappa as the measurement of the inter-coder agreement. According to McHugh (2012), the Cohen's Kappa value 0.01-0.20 indicating none to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement. The result turned out to be .958 with the *p*-value lower than .001, which indicates good inter-coder reliability.

The final survey after the aforementioned steps contains 31 items, with three measuring perceived enjoyment, 11 for perceived usefulness, seven for perceived ease of use, four for facilitating conditions, three for perceived risks, and three for control variables. The specific survey items are shown in Appendix A.

7.1.3. Determinant inclusion and exclusion criteria

As detailed in the previous section, this author selected and summarised previous TAM-related frameworks. At this stage, concerning contextualisation, frameworks that are not significantly applicable for the objective of the current study have been removed, such as those describing the acceptance willingness of fully-developed digital games.

Notably, this study had special emphasis on taking a wider perspective to review the current topic, which aimed to connect theories that are substantively contributing to the current topic even though not seemed closed-related from the names. This step is to avoid the poor conceptualisation of the model caused by a narrow literature review (Younas & Porr, 2018).

During this process, one determinant *perceived risk* in a model proposed by Deng et al. (2018) was added in the framework, which was to predict patients' acceptance willingness toward mobile health services. The actual connotation of perceived risks may vary from the mobile health service to educational gamification; however, this term provides meaningful insight into the possible factors that may hinder the adoption of new technology. A similar concept can be found in the APEASE framework proposed by Michie et al. (2011), which is to assess the "effectiveness" of an intervention. *Compatibility* and *complexity* were also added as two sub-constructs under the construct PEoU, which originate from the diffusion of innovation theory proposed by Rogers (2010). The term "cost-effectiveness" was also added. The newly-added constructs or sub-constructs, as well as their connotations in the current research context, were reviewed by experts for validity.

The next step was to merge the related frameworks into one while eliminating factors that are not significantly relevant. The retention of terminology was an issue, as different frameworks use varied terms to indicate the same or similar objectives. For example, the term in most TAM-extension frameworks "attitude" is equivalent to "trust" in the framework by Deng et al. (2018) or "perceived satisfaction" in the framework by Cigdem and Ozturk (2016). Terminology consistency was guaranteed via expert interviews.

A knotty issue is that, in assessing whether gamified learning tools are useful, varied criteria can be applied from different perspectives; consequently, it is not rational to conclude the benefits brought by gamified learning tools into one term only (perceived usefulness). To be specific, a gamified learning tool can be useful in boosting learning inputs, in triggering learning motivation, in promoting academic achievements, etc.; however, those functions are not necessarily connected. If the terms are summarised into one, the final survey items would indicate more than one

object, which brings potential challenges to variance and reliability. Previous surveys assessing PU covered varied benefits that make educational technology perceived useful, including advantages over traditional teaching (Beggs, 2000), increased student interest (Beggs, 2000), improved performance expectancy (Venkatesh et al., 2003), enhanced student learning (Beggs, 2000), increased learning opportunity (De Grove et al., 2012; Ibrahim & Jaafar, 2011), and knowledge learning (Lin & Chen, 2013).

Though varied benefits have been claimed, two of the most common goals of educational gamification are improving students' learning engagement (Landers & Armstrong, 2017) and enhancing students' academic performance (Lin & Chen, 2013). Engagement plays an important role in learning, which "has been proved a positively correlated with outcomes of student success" such as satisfaction, academic achievement and self-efficacy (Ibanez et al., 2014, p. 292). Bourgonjon et al. (2009) criticised that it is too narrow to focus on academic results only, because "education means more than mere outcome" (p. 653). Accordingly, Bourgonjon et al. (2009) divided the perceived usefulness into "perceived usefulness-process" (PU-Process) and "perceived usefulness-product" (PU-Product). To minimise difficulties in understanding, PU-Process and PU-Product were reworded as PU-Process and PU-Product.

Factors that do not fall into any relevant category were excluded, such as the time needed to learn (Beggs, 2000). Factors proposed for specific contexts were also excluded, such as expertise (Y. Sun & Jeyaraj, 2013), self-taught computer literacy (Cheng, Lou, Kuo, & Shih, 2013), control variables (education and chronic diseases) (Deng et al., 2018). The factor "interaction" was excluded also. As an important concept in gamification, interaction highlights human-device interactions (Cigdem & Ozturk, 2016; Huang & Liaw, 2018) or social interactions (Lin & Chen, 2013). Though both types of interactions apply to the current research settings, they are comparatively not stressed for two reasons: the human-device interactions can fall into the categories of complexity or the general perceived ease of use, which is available in the established framework; the social interactions are popular in student-dominating gamification tools which are comparatively less used in teacher-leading classroom time.

One possible controversial decision is to exclude self-efficacy from the theoretical framework. Self-efficacy, "a learner's belief that he or she is capable of performing a task and reaching a goal" (Huang & Liaw, 2018, p. 24), is a common determinant in TAM-extension frameworks (Cigdem & Ozturk, 2016; Huang & Liaw, 2018; Ibrahim & Jaafar, 2011; Lin & Chen, 2013; Park et al., 2012; Zhang et al., 2017). Self-efficacy is a factor focusing on an individual's inner strength, which furtherly influences an individual's learning attitude or learning motivation (Huang & Liaw, 2018). Since the emphasis of the current study is the tools rather than individuals (e.g., the perceived risks brought by the tools, the perceived usefulness provided by the tools, etc.), self-efficacy has been excluded.

During the process, the determinant "cost-effectiveness" has been considered and eventually excluded. Cost-effectiveness, referring to "the ratio of effect to cost" in completing activities (Michie et al., 2011, p. 23), is a factor in the APEASE framework that assesses the effectiveness of a behaviour change intervention (Michie et al., 2011). Cost-effectiveness brings new insights for understanding the perceived usefulness of gamified learning tools since a tool can be regarded as both 'useful' if it provides educational benefits and "useless" if the benefits it provides is less than what can be provided by other tools or approaches. However, the APEASE framework was designed for health-related interventions, and its applicability in educational contexts has not been validated.

7.1.4. The theoretical framework

Following the aforementioned procedures, the current study established the theoretical framework shown in Figure 13 and Table 31.

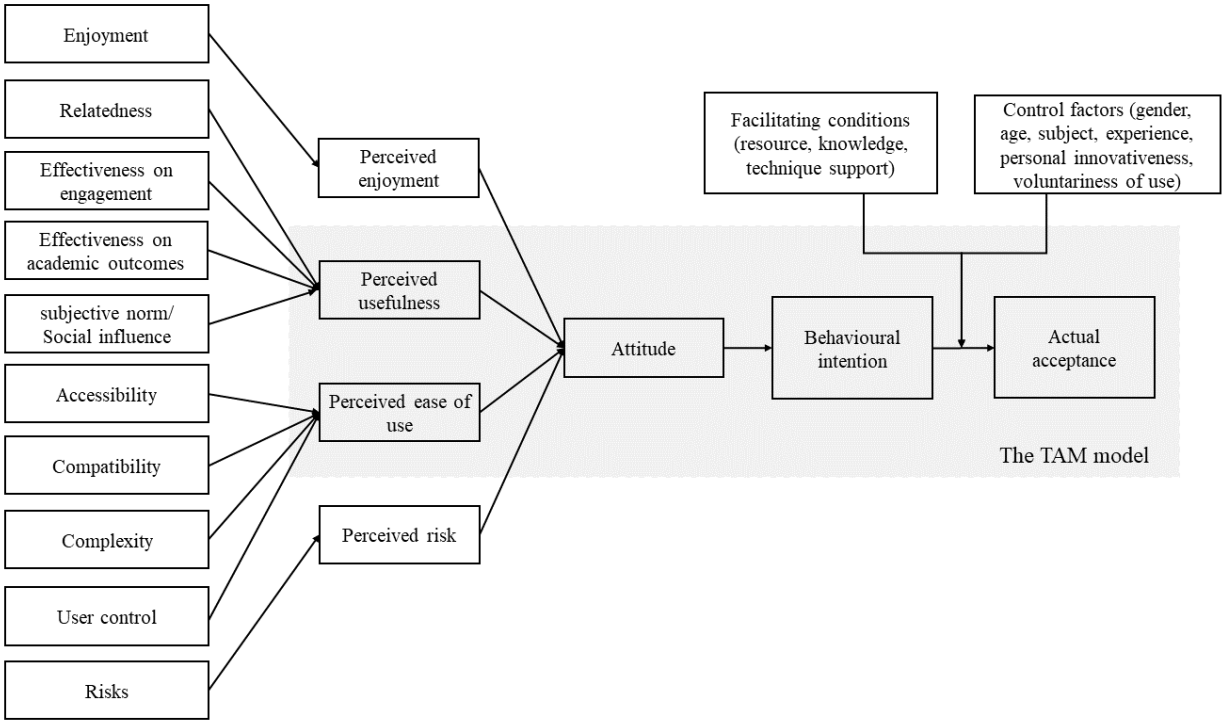

Figure 13. The theatrical framework: determinants of teachers' acceptance intention to gamified learning tools

Table 31. Constructs and sub-constructs of the theoretical framework

Construct	Sub-construct	Synonyms and the source
Perceived enjoyment	n/a	Enjoyment (Herzig et al., 2012; Ibrahim & Jaafar, 2011), flow (Herzig et al., 2012)
PU	Overall PU	Perceived usefulness (Huang & Liaw, 2018; Park, 2009; Park et al., 2012; Y. Sun & Jeyaraj, 2013; Zhang et al., 2017), usefulness (De Grove et al., 2012)
	Relatedness	Major relevance (Park et al., 2012), curriculum-relatedness (De Grove et al., 2012)
	Usefulness-engagement	Advantage over traditional teaching (Beggs, 2000), increased student interest (Beggs, 2000), performance expectancy (Ibrahim & Jaafar, 2011; Venkatesh et al., 2003), useful-process (Bourgonjon et al., 2009)
	Usefulness-academic outcome	Improved student learning (Beggs, 2000), learning opportunity (De Grove et al., 2012; Ibrahim & Jaafar, 2011), knowledge learning (Lin & Chen, 2013), multimedia instruction (Cigdem & Ozturk, 2016), useful-product (Bourgonjon et al., 2009)
	Subjective norm/ social influence	Social influence (Ibrahim & Jaafar, 2011; Y. Sun & Jeyaraj, 2013; Venkatesh et al., 2003), subjective norm (Park, 2009; Park et al., 2012; Sugar et al., 2004)
PEoU	Overall PEoU	Ease of use (Beggs, 2000; De Grove et al., 2012), perceived ease of use (Huang & Liaw, 2018; Park, 2009; Park et al., 2012; Y. Sun & Jeyaraj, 2013; Zhang et al., 2017)
	Accessibility	Equipment availability (Beggs, 2000), materials in discipline (Beggs, 2000), system accessibility (Park, 2009; Park et al., 2012)
	Compatibility	Compatibility with materials (Beggs, 2000), work compatibility (Y. Sun & Jeyaraj, 2013), compatibility (Rogers, 2010)
	Complexity	Personal comfort (Beggs, 2000), effort expectancy (Ibrahim & Jaafar, 2011; Venkatesh et al., 2003), complexity (Rogers, 2010)
	User control	Perceived learner control (Zhang et al., 2017), perceived behavioural control (Sugar et al., 2004)
Perceived risks	(A lack of) support	Training (Beggs, 2000), administrative support (Beggs, 2000), support (Selim, 2007), Colleague use (Beggs, 2000)
	The general perceived risks	Perceived risks (Deng et al., 2018)
Facilitating conditions	Necessary resources	(Y. Sun & Jeyaraj, 2013; Wong, 2016)
	Technique support	(Y. Sun & Jeyaraj, 2013; Wong, 2016)
	Necessary knowledge	(Y. Sun & Jeyaraj, 2013; Wong, 2016)
	Fitness to the workflow	(Wong, 2016)
Control variables	Gender	(Cheng et al., 2013; Deng et al., 2018; Ibrahim & Jaafar, 2011; Venkatesh et al., 2003),
	Age	(Deng et al., 2018; Venkatesh et al., 2003)
	Experience	Gaming experience (Ibrahim & Jaafar, 2011), experience with DGBL (digital game-based learning) gaming experience (Cheng et al., 2013), experience (De Grove et al., 2012; Venkatesh et al., 2003)
	Personal innovativeness	Personal innovativeness in IT (Ibrahim & Jaafar, 2011; Zhang et al., 2017)
	Voluntariness	(Venkatesh et al., 2003)

7.2. Expert comment on the framework

After drafting the first version of the theoretical framework, two experts have been consulted for comments and suggestions. The two experts have been given a briefing on the definition of gamification and gamified learning, the current research context, the research scope, the structure of the theoretical framework, and the definition of each construct in the framework.

PU-Process and PU-Product. Both of the two experts confirmed the significance of dividing perceived usefulness of gamified learning tools into two aspects: engagement and academic outcomes. This is especially important in a test-oriented educational system, in which effectiveness is widely regarded strongly-connecting with the improvement of academic performance. Enhanced academic performance is "the result of the long-term contribution of *many* facilitating factors and their dynamic interactions with other factors (both facilitating and debilitating ones)"; therefore, if the researchers only change one variable for a short time without strictly controlling other variables, "it might not be proper to assert whether the selected variable contributes to students' academic performances" (Expert 1).

Holding the same viewpoint, the second expert gave examples of the contributing factors, such as demographic differences (e.g., age, race, gender, etc.) (Rodgers & Rose, 2001; Stewart, 2008), personality (Chamorro-Premuzic & Furnham, 2003), affective factors (e.g., depression, anxiety, motivation, stress, satisfaction, etc.) (McKenzie & Schweitzer, 2001), student competence or student effort (e.g., study skills, academic score when entering the school, the use of varied learning strategies) (Abdulghani et al., 2014; Fenollar, Román, & Cuestas, 2007; McKenzie & Schweitzer, 2001; Stewart, 2008), achievement goals (mastery goals, performance-approach goals, performance-avoidance goals, work avoidance) (Fenollar et al., 2007), self-efficacy (Fenollar et al., 2007; Honicke & Broadbent, 2016), family background or parental involvement (e.g., mothers' education, parental support, parent monitoring, socio-economic background, varied family structure) (Abdulghani et al., 2014; Considine & Zappalà, 2002; Okioga, 2013; Rodgers & Rose, 2001; Stewart, 2008), learning environment (e.g., class size, school size, school location) (Fenollar et al., 2007; Stewart, 2008), and non-academic hindrances (e.g., sleep deprivation, language barriers, homesickness) (Abdulghani et al., 2014).

Under this circumstance, a gamified learning tool is possible to be highly engaging but not bringing instant improvements to academic achievement, so the assessment of its effectiveness is an issue.

Expert 1 emphasised that it is assumed that teachers in test-oriented countries take academic achievements as the absolute priority, but the reality might say no: "a considerable number of teachers care about students' experience in learning". The motivation varies from teacher to teacher. Some teachers are teaching high-performing students, so they feel free to use the saved time for extra arrangements such as gamified learning activities; some teachers work in international schools that advocate international teaching methodologies, so they are encouraged to pay more attention to innovative approaches or technologies; some teachers work in the public schools mainly delivering lectures in the traditional teacher-centred way, but they wish to use educational gamification to; there are also teachers teaching students with "extremely low academic performance", then some of the teachers wish to use gamification to trigger students to give more input without expecting academic improvements. At the same time, whatever type of the school is, the well-skilled teachers are also in greater concern of students' learning experience compared with the non-skilled ones. It is because those well-skilled teachers mastered how to deal with test-oriented teaching, so they have the extra time, effort and mental capacity for higher pursuits.

Expert 2 commented that the test-oriented educational system does bring extrinsic pressure, but once the schools or teachers put the pressure aside, they "would be happy to create and experience happy education with the students". If being provided facilitating conditions, such as encouraging policies and essential resources, their willingness "is highly possible to transfer to actual behaviour".

To conclude, the effectiveness of gamified learning tools should be assessed separately from the perspective of engagement and academic performance. Though it seems that teachers in test-oriented educational systems take exam scores as the priority, the fact is that they also value students' learning experience and the positive benefits brought by improvement learning engagement.

Perceived enjoyment and PU-Process. Expert 2 criticised the theoretical framework proposed by this study. Expert 2 pointed out that the concepts of perceived enjoyment and the effectiveness of engagement are overlapping in the current context. To be specific, the perceived enjoyment may result in improved engagement, so the two concepts are possible to have a causal relationship. Whether they are measuring the same construct needs to be confirmed in exploratory factor analysis.

Control variables and facilitating conditions. Expert 1 suggested adding subject as one control variable since the effectiveness and user experience of gamified learning tools are largely dependent on varied subjects. The two experts commented that though gender and age have been widely used as the variables affecting the effectiveness of educational interventions, they are non-grouped factors (e.g. height and weight) rather than ordinal factors (ranked categories such as satisfaction level and level of agreement), so they are not suitable to be assessed by Likert scales. The same situation applies to teaching subjects, too, as subject is a nominal factor like marital status and eye colour. As to facilitating conditions, previous scales detailed the specific measuring items (necessary resources, technique support, necessary knowledge, and fitness to the workflow), but more can be considered in future studies such as policy support and peer encouragement (Expert 1 and Expert 2).

To sum up, the two experts confirm, criticised and commented based on the proposed framework, as simplified in Table 32.

Table 32. Summaries of expert comments

Status	Expert comment
Confirmed	It is important to divide the effectiveness of gamified learning tools into two aspects: engagement and academic outcomes.
Criticised	A gamified learning tool is possible to be highly engaging but not bringing instant improvements to academic achievement, which brings mixed or even contradictory research findings.
	Gender and age should be excluded from the scale.
	"Teaching subject" should be included as a control variable.
Issues to be tested or explored	The research findings might not be limited to test-oriented educational contexts.
	The connotation of PEoU may be varied in the current context: the PEoU can indicate the PEoU of *the technological or the devices*, or the PEoU of the *implementation process*.
	The concepts of perceived enjoyment and the effectiveness of engagement are overlapping in the current context.
	More factors should be considered as facilitating conditions.

7.3. Scale development of the TAI-GLT scale

Based on the theoretical framework in Figure 13, survey items have been selected from previous scales. Items for the sub-construct "PU-Process" and the construct "perceived risks" were not found in existing surveys, so they were drafted based on literature and interview responses. The specific survey items are as shown in Table 33.

Though attitude and behavioural intention are two factors in the framework, no survey items are measuring them because the whole scale is designed to measure them but the scale is not validated yet.

Table 33. Items of the scale measuring teachers' acceptance intention

Construct	Code	Sub-construct	Item	Source
Perceived enjoyment (ENJ)	ENJ1	Flow	Using the tool is pleasurable to students.	(Davidson, 2018)
	ENJ2	Flow	Students can forget about time passing while learning with the gamified learning tools.	(Fu, Su, & Yu, 2009)
	ENJ3	Flow	Students can become involved in learning with the gamified learning tools.	(Fu et al., 2009)
Perceived usefulness (PU)	PU1	Overall PU	My job would be difficult to perform without the tool.	(Davis, 1989)
	PU2	Overall PU	Using the tool improves my performance.	(Davis, 1989)
	PU3	Overall PU	Overall, I find the gamified learning tool useful in my teaching.	(Bourgonjon et al., 2013)
	PU4	Relatedness	In my teaching, usage of the tool is relevant.	(Venkatesh & Davis, 2000)
	PU5	Relatedness	The tool would fit the current curriculum.	(Adukaite et al., 2017)
	PU6	PU-Product	Gamified learning tools offer opportunities to experiment with knowledge.	(Bourgonjon et al., 2013)
	PU7	PU-Product	The quality of the output the students get from the gamified learning tool is high.	(Venkatesh & Davis, 2000)
	PU8	PU-Process	Students are happy when using gamified learning tools in schools.	From interview responses
	PU9	PU-Process	Teachers enjoy using gamified learning tools in schools.	From interview responses
	PU10	Social influence	People who influence my behaviour think that I should use gamified learning tools.	(Venkatesh & Davis, 2000)
	PU11	Social influence	People who are important to me think that I should use gamified learning tools.	(Venkatesh & Davis, 2000)
Perceived ease of use (PEoU)	PEoU1	General PEoU	The gamified learning tool is rigid and inflexible to interact with.	(Davis, 1989)
	PEoU2	General PEoU	The gamified learning tool often behaviours in unexpected ways.	(Davis, 1989)
	PEoU3	General PEoU	I find it cumbersome to use the gamified learning tool.	(Davis, 1989)
	PEoU4	Accessibility	It is difficult to get access to a suitable gamified learning tool.	(Park et al., 2012)
	PEoU5	Compatibility	The gamified learning tools have bad compatibility with the current devices.	(Park et al., 2012)
	PEoU6	Complexity	The gamified learning tools are complex to use.	From interview responses
	PEoU7	User control	I feel lost control when I'm using the gamified learning tools in teaching.	(Zhang et al., 2017)
Perceived risks (RISK)	RISK1	(A lack of) support	The principals, parents and students are possible to refuse using the gamified learning tool.	From literature
	RISK2	Controllability	Using the gamified learning tool in school is easy to lose control.	From interview responses
	RISK3	Side effects	There are side effects in using the gamified learning tool (e.g. hurt eyesight, distract learners, etc.).	From interview responses
Facilitating conditions (FC)	FC1	Necessary resource	I have the resources necessary to use gamified learning tools.	(Wong, 2016)
	FC2	Necessary knowledge	I have the knowledge necessary to use gamified learning tools.	(Wong, 2016)
	FC3	Workflow	Gamified learning tools fit well into my workflow	(Wong, 2016)

	FC4	Technical support	A specific person or group (e.g. technical support team) is available for assistance with difficulties using gamified learning tools.	(Wong, 2016)
Control variable (CV)	CV1	Experience	I have used gamified learning tools before.	(Bourgonjon et al., 2013; De Grove et al., 2012)
	CV2	Voluntariness	My use of the gamified learning tool is voluntary.	(Venkatesh & Davis, 2000)
	CV3	Personal innovativeness	Among my peers, I am usually the first to try out new information technologies.	(Zhang et al., 2017)

Note: The items in italics are those deleted in scale evaluation.

7.4. Scale validation of the TAI-GLT scale

7.4.1. Instrument validation

After arriving at a preliminary version of the scale, items should be assessed for content validity and face validity (DeVellis, 2003; Robertson, 2017). Two experts in the field of gamification have been consulted for the following three issues: whether the scale contains a sufficient and appropriate sample of items to represent the constructs of interests, whether all facets of the construct of interest are being measured, and whether the scale looks like measuring what it claimed to measure.

The internal consistency of the scale was examined based on a reliability analysis (Cronbach's alpha) on the entire dataset (N = 516). Findings showed that the survey measuring most of the constructs demonstrated a high level of internal reliability, as the Cronbach's alpha values exceed the common threshold of 0.70 (Hair, Black, Babin, Anderson, & Tatham, 1998; Taber, 2018): perceived enjoyment (α = .870), perceived usefulness (α = .916), perceived ease of use (α = .916), perceived risk (α = .825), attitude (α = .836), behavioural intention (α = .787) and facilitating conditions (α = .747). Control variables is with a comparatively low internal reliability (α = .586), so the correspondent three items have been removed.

7.4.2. Exploratory factor analysis (EFA) results

The Bartlett's test was significant ($\chi 2(276) = 8233.012$, $p < .001$) and the overall Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (MSA) was good at .940 (cut-off criterion: .50. KMO higher than .80 is excellent), which verified that the collected data were sufficient to employ factor analyses.

The correlations among the 28 items have been explored. The results show that facilitating conditions, perceived enjoyment and the perceived usefulness were strongly and positively correlated as a group (greater than .30, significant at the 0.01 level), the perceived ease of use and the perceived risks were correlated as a group (greater than .30, significant at the 0.01 level), and the two groups were with insignificant correlations. The results indicate that the 28 items consist of at least two distinct variables.

After confirming that the data set is suitable for factor analysis, an EFA was conducted using Principal Components Analysis (PCA) and Varimax rotation (see Table 34). Following suggestions of Costello and Osborne (2005), five items have been removed due to low communalities (lower than .50): the four items for facilitating conditions (.401, .393, .289 and .407 respectively) and item PU1 (.153). The EFA analysis extracted two factors with eigenvalues greater than one, which explained 60.11% of the total variance, with the first factor accounting for 34.07% and the second factor 26.04%. It is notable that the factor extraction results indicated that the perceived usefulness and perceived enjoyment were statistically regarded as one factor, as well as the constructs perceived ease of use and perceived risks.

The Average Variance Extracted (AVE) values for each factor was checked to assess the construct validity of the scale (see Table 34). The results indicated that AVE values for PU and PEoU were .598 and .600, which is above the cut-off criterion 0.5. Both of the two factors demonstrated good internal consistency as by their Cronbach's alpha *(α) and* composite reliability. As shown in Table 34, the Cronbach's alpha for the two constructs were .943 and .926, and the composite reliability for the two constructs were .951 and .937, higher than the cut-off criteria .70.

However, it is notable that a high value of alpha (> 0.90) “may suggest redundancies” because the items are “testing the same question but in a different guise” (Tavakol & Dennick, 2011, p. 54). Though the acceptable values of alpha vary from .70 to .95 in previous studies, the recommend alpha value is 0.90 at the maximum (Tavakol & Dennick, 2011). The reduction of redundancy can be achieved by computing the correlation of each test item with the total score test and remove those with low correlations (Tavakol & Dennick, 2011), which was detailed in the next section.

Besides the aforementioned four main steps, this study also involved other steps such as determining the format of the measurement as a 7-point Likert survey, including validation items, conducting a pilot study, and removing invalid surveys, as suggested by DeVellis (2003) and Robertson (2017).

Table 34. Descriptive data and EFA results for the TAI-GLT scale

Scale item	Mean	SD	*h2*	Factor 1	Factor 2
ENJ1	5.44	1.389	0.656	**0.810**	0.019
ENJ2	5.34	1.408	0.570	**0.755**	0.005
ENJ3	5.43	1.416	0.649	**0.804**	0.050
PU2	5.02	1.263	0.607	**0.779**	0.009
PU3	5.22	1.340	0.705	**0.836**	-0.077
PU4	5.07	1.331	0.633	**0.794**	-0.057
PU5	4.87	1.414	0.472	**0.686**	-0.040
PU6	5.21	1.360	0.662	**0.807**	-0.101
PU7	5.33	1.322	0.673	**0.819**	-0.048
PU8	5.43	1.433	0.617	**0.785**	0.022
PU9	4.92	1.394	0.500	**0.707**	-0.011
PU10	4.90	1.431	0.530	**0.727**	0.045
PU11	4.97	1.390	0.526	**0.725**	0.009
PEoU1	4.46	1.542	0.575	0.075	**0.754**
PEoU2	4.25	1.626	0.666	0.044	**0.815**
PEoU3	4.03	1.690	0.652	-0.053	**0.806**
PEoU4	4.23	1.649	0.630	0.004	**0.794**
PEoU5	4.21	1.632	0.631	-0.018	**0.794**
PEoU6	4.04	1.697	0.661	-0.044	**0.812**
PEoU7	4.01	1.653	0.630	-0.127	**0.784**
RISK1	4.45	1.540	0.518	0.048	**0.718**
RISK2	4.41	1.645	0.561	-0.022	**0.749**
RISK3	4.61	1.529	0.503	-0.041	**0.708**
Total variance explained: 60.11%				34.07%	26.04%
Cronbach's alpha: .887				.943	.926
Composite reliability (CR)				.951	.937
Average variance extracted (AVE)				.598	.600

Note: h^2: communalities. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

7.4.3. Shortened scale version one

According to Worthington and Whittaker (2006), when a survey contains more than the desired number of items, designers can choose to delete items that have the lowest factor loadings, the highest cross-loadings, low conceptual consistency with other items in the factor, and the ones contribute the least to the internal consistency of the scale scores.

Since the first-round-validated scale was with satisfactory internal consistency and factor loadings, the shortening of the scale aims to reduce redundancy; therefore, the scale is not with the problem of cross-loadings, conceptual consistency and low internal consistency. Accordingly, this author mainly paid attention to items with low factor loadings. What's more, the value of alpha is also affected by the number of survey items (a low number of questions leads to a low value of alpha), so the alpha value will naturally drop to the satisfactory level (a maximum alpha value of 0.90) when enough items are deleted (Tavakol & Dennick, 2011).

Table 35 presents the process of cutting survey items from 13 to five for the items measuring teachers' the perceived usefulness (PU) of gamified learning tools in secondary schools. By removing the items with low factor loadings, this study reduced the number of items from 13 to eight, with the Cronbach's alpha dropping from 0.943 to 0.923. Further reduction was achieved by both referring to the factor loadings and the function 'Cronbach's alpha if item deleted' that generated by the software IBM SPSS. Eventually, this study obtained five survey items with the value of the internal consistency of 0.899.

Following the same procedure, there were six items retained out of 10 items measuring teachers' perceived ease of use (PEoU) of gamified learning tools in secondary schools, as shown in Table 36. The specific survey items for the PU and PEoU scales are detailed in the Appendices.

Table 35. The TAI-GLT scale (PU section) with different number of items (13, 8 and 5)

#	Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
1	Number of items: 13. Cronbach's alpha: 0.943				
	PU12 (ENJ1)	61.69	161.542	0.765	0.937
	PU13(ENJ2)	61.78	163.258	0.701	0.939
	PU14(ENJ3)	61.70	161.216	0.758	0.938
	PU2	62.11	165.017	0.735	0.938
	PU3	61.91	161.436	0.800	0.936
	PU4	62.06	163.084	0.754	0.938
	PU5	62.26	165.365	0.635	0.941
	PU6	61.92	162.031	0.768	0.937
	PU7	61.80	162.479	0.779	0.937
	PU8	61.70	161.546	0.737	0.938
	PU9	62.21	164.889	0.660	0.941
	PU10	62.23	163.536	0.680	0.940
	PU11	62.16	164.329	0.679	0.940
2	Number of items: 8. Cronbach's alpha: 0.923				
	PU12 (ENJ1)	36.54	60.396	0.759	0.911
	PU14(ENJ3)	36.55	60.178	0.752	0.912
	PU3	36.76	60.397	0.792	0.909
	PU4	36.91	61.558	0.736	0.913
	PU5	37.11	62.988	0.611	0.923
	PU6	36.77	60.653	0.764	0.911
	PU7	36.65	60.910	0.777	0.910
	PU8	36.55	60.240	0.737	0.913
3	Number of items: 5. Cronbach's alpha: 0.899				
	PU14 (ENJ3)	21.18	21.722	0.754	0.876
	PU3	21.39	22.265	0.761	0.875
	PU6	21.40	22.381	0.734	0.880
	PU7	21.28	22.443	0.758	0.876

PU8	21.18	21.679	0.745	0.878

Table 36. The TAI-GLT scale (PEoU section) with different number of items (10, 7 and 6)

#	Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
1		Number of items: 10. Cronbach's alpha: 0.926			
	PEoU1	38.25	130.976	0.687	0.920
	PEoU2	38.46	127.177	0.757	0.916
	PEoU3	38.68	126.280	0.750	0.916
	PEoU4	38.48	127.586	0.733	0.917
	PEoU5	38.50	127.742	0.737	0.917
	PEoU6	38.66	125.874	0.757	0.916
	PEoU7	38.70	127.655	0.729	0.917
	PEoU8 (RISK1)	38.26	132.121	0.653	0.921
	PEoU9 (RISK2)	38.29	129.210	0.687	0.920
	PEoU10 (RISK3)	38.10	132.507	0.646	0.922
2		Number of items: 7. Cronbach's alpha: 0.916			
	PEoU1	24.77	67.898	0.685	0.908
	PEoU2	24.98	64.924	0.768	0.900
	PEoU3	25.20	63.679	0.785	0.898
	PEoU4	25.00	65.410	0.734	0.904
	PEoU5	25.02	65.106	0.756	0.901
	PEoU6	25.18	64.224	0.757	0.901
	PEoU7	25.22	66.090	0.702	0.907
3		Number of items: 6. Cronbach's alpha: 0.898			
	PEoU1	20.74	47.202	0.666	0.889
	PEoU2	20.95	44.574	0.758	0.875
	PEoU4	20.97	44.855	0.729	0.879
	PEoU5	20.99	44.650	0.750	0.876
	PEoU6	21.16	43.797	0.757	0.875
	PEoU7	21.19	45.653	0.685	0.886

7.5. Discussion: Revision of the scale TAI-GLT

7.5.1. Structure of the framework

Data of the current study revealed surprising results: the structure of the proposed framework in Figure 14 is not as promising as expected from the statistics perspective: firstly, the face validity (whether the items are measuring what they claimed to measure) is not satisfactory as the two constructs perceived risk and PEoU statistically regarded as one factor rather than two distinct ones. The same thing happens to the perceived enjoyment and PU; secondly, the data demonstrated that two assumed-important constructs were not as significant as expected, including facilitating conditions and control variables.

The face validity problems confirmed experts' comments that "perceived enjoyment" and "PU-Process" are overlapping concepts. Previous TAM-related frameworks involved enjoyment, such as the Technology acceptance model 3 (TAM3) by Venkatesh and Bala (2008), the educational games acceptance model by Ibrahim and Jaafar (2011), and the gamification effects on user acceptance model by Herzig et al. (2012). Findings of the current study indicated that the measurement of perceived enjoyment can be replayed by the measurement of one type of learning engagement, such as emotional engagement in the classroom (which is also named affective engagement).

The face validity problems also confirmed experts' comment that PEoU in the current context (using gamified learning tools in secondary schools in formal education) means more than the cumbersomeness or smoothness of experiencing the digital devices; instead, the implementation process of educational gamification as a whole is also of great concern. Survey items covered three aspects of the perceived risks: a lack of support, the possibility of losing classroom control, and side effects (e.g. hurt eyesight, distract learners, etc.), all of which require teachers to investigate more time and effort. Once the implementation of gamification is time-consuming and effort-demanding, the process becomes "not easy", which thereby decrease the perceived "ease" of use of the gamification approach or gamified learning tools.

Facilitating conditions emphasise necessary resources, necessary knowledge, technical support, and compatibility with the workflow, as proposed by Wong (2016). The four items to measure facilitating conditions were excluded mainly because they were separated into two factors in the EFA with each factor containing two items, while each factor should contain at least three items for further analysis such as structural equation modelling (SEM) (MacCallum, Widaman, Zhang, & Hong, 1999; Raubenheimer, 2004). Therefore, it is important that the removal of the four items does not indicate the improperness of involving the construct "facilitating condition" in the TAM-extension model. Future studies should construct and validate a scale measuring facilitating conditions with more initial items. Other situations besides the aforementioned four aspects can also be considered, such as easy access to the right tool, policy support, peer influence, etc.

Items measuring control variables presented high variance, which makes them not suitable for being put into one scale. However, the variables can be used to assess how people of different groups perceive the gamified learning tools. In previous studies, most of the involved variables are age and gender, such as the study of Deng et al. (2018) and Venkatesh et al. (2003). More variables can be considered, such as experience, voluntariness and personal innovativeness. As suggested by Expert 1 in the expert interview, subject can be considered because gamifying an educational activity is highly relevant to the content, while the content highly depends on the nature of the teaching subject.

7.5.2. Excluded items from the TAI-GLT scale

The first item being removed from the PU scale is PU1 that adapted from the original PU scale proposed by (Davis, 1989): “my job would be difficult to perform without the tool”. on contrary, interview responses revealed that the participants do not regard gamification as a facilitating tool to their teaching efficiency, as the gamified learning tools are both effort-demanding and time-demanding on preparation for teachers (P1 and P14) while being not able to guarantee quick benefits on academic performance for students (P1, P3, P5, P7, and P14). At the same time, digital gamified learning tools as the mediums of new technology bring challenges on acceptance willingness and device operation, as one participant aged over 46 reported that he as an “elder man working in a rural school” does not know, nor interested in, new technologies. He furtherly suggested to “interview young teachers working in urban schools: they might know more about educational innovations” (P13). However, almost all of the participants admitted that gamified learning tools are of great help or of great potential in making students engaged. In short, at least in the current study, gamified learning and the existing learning tools aim to trigger students’ learning engagement at the sacrifice of teachers’ convenience. Therefore, from secondary teachers’ perspectives, gamified learning tools are icing on the cake (improving students’ learning experience) rather than providing timely assistance (enhancing academic performance).

Consequently, the perceived usefulness in the current research context was interpreted as the usefulness for students from the learning experience perspective rather than the usefulness for teachers from the implementation perspective. On the same principle, the factor analysis results indicated that several items should be removed, including the one related to teacher’s performance (PU2 “Using the tool improves my performance”), teachers’ teaching experience (PU9 “Teachers enjoys using gamified learning tools in schools”), as well as the two related with subjective norms (PU10 “People who influence my behaviour think that I should use gamified learning tools” and PU11 “People who are important to me think that I should use gamified learning tools”).

PU4 and PU5 were included in the 8-item scale and excluded from the 5-item scale due to redundancy concerns (PU4 “In my teaching, usage of the tool is relevant” and PU5 “The tool would fit the current curriculum”), which were designed to test the sub-theme curriculum relatedness (Adukaite et al., 2017; Venkatesh & Davis, 2000). Curriculum relatedness is comparatively not as straightforward as effectiveness when connecting with the keyword usefulness; however, “useful” can be interpreted twofold: it can be interpreted as “competent” or “effective” in achieving goals, which is in strong relation with “effectiveness” while being interpreted as “serviceable” or “functional” for a practical purpose, which is related with “being helpful”. Previous literature indicates that the integration of game features and pedagogical needs is a huge problem for educational gamification, while the interview responses show that the secondary school teachers are highly possible to quit using the gamified learning tools if they are not directly related to the current curriculum. Therefore, though PU4 and PU5 were removed from the current study, the rationale of excluding the sub-construct “relatedness” needs further validation.

All of the proposed items measuring PEoU have been validated, except the four being excluded for redundancy concerns. PEoU3 (“I find it cumbersome to use the gamified learning tool”) is similar to PEoU1 and PEoU2 (“The gamified learning tool is rigid and inflexible to interact with” and “The gamified learning tool often behaviours in unexpected ways”); the three items related with the perceived risks (a lack of support, a lack of classroom control, and risks of side effects) can be generalised as the risks of losing control and the challenges of complexity, which are measured by PEoU7 and PEoU6 (“I feel lost control when I'm using the gamified learning tools in teaching” and “The gamified learning tools are complex to use”). Again, the concept “use” interpreted by the participants is not

about the interaction with the digital devices but also the implementation of the gamification approach as a complete activity.

7.5.3. Conclusion and limitations

Based on the technology acceptance model (TAM), this author constructed and validated a two-dimensional scale measuring teachers' acceptance intention of gamified learning tools.

To sum up, the proposed framework was confirmed not fully valid in the current research contexts, as only two of the proposed six constructs were validated from the statistics perspective. The constructs "perceived enjoyment" and "perceived risks" should be removed; the construct "facilitating conditions" should be extended and revalidated in future studies; and the construct control variables (gender, age, subject, experience, personal innovativeness, and voluntariness) should be excluded from the scale.

Moreover, though the TAI-GLT scale has been validated, the number of items measuring the sub-constructs is limited: according to MacCallum et al. (1999) and Raubenheimer (2004), there should be at least three items for each construct for further studies. This is also why in shortening the scale some assuming-important items were excluded from the statistical analysis, such as the items measuring the relatedness of gamified learning tools. To fix this problem, future studies could consider designing at least three items for each sub-construct.

CHAPTER 8

Findings (3)

08

Scale development

PU-GLT and PEoU-GLT

Gamified learning tools (GLTs) in the current context refer to educational websites, software or mobile apps that employ game design elements for improved learning engagement or performance.

Though theoretically promising, gamified learning tools meet challenges in user acceptance. It is proposed that technology acceptance consists of two main factors: the perceived usefulness (PU) and perceived ease of use (PEoU); however, little is known about what makes a gamified learning tool perceived useful or perceived easy to use.

This study aims to explore the determinants of PU and PEoU of a gamified learning tool for foreign language learning, followed by the construction and evaluation of the correspondent scales. After an interview study and a survey study that involved 361 and 502 secondary school teachers respectively, this author constructed and evaluated two scales: a 15-item survey measuring the three determinants of PU-GLT (curriculum relatedness, PU-Product, PU-Process, and social influence) and an 11-item survey measuring the four determinants of PEoU-GLT (effort expectancy, support, and affordability). Applications and limitations were discussed.

CHAPTER EIGHT: CONSTRUCTION AND VALIDATION OF THE PU-GLT AND PEOU-GLT SCALES

Scale development and evaluation of PU-GLT and PEoU-GLT

8.1. Theoretical framework: Determinants of PU and PEoU

Though theoretically promising, gamification as an innovative technology meets challenges in user acceptance and actual implementation. If not being accepted, the full potential of educational technologies will be limited, so more studies are required to improve users' acceptance intention of certain technologies (Rahmi et al., 2018). The Technology Acceptance Model (TAM) proposed by Davis et al. (1989) is a well-known framework in the field, which contains two main constructs: the perceived usefulness (PU) and the perceived ease of use (PEoU).

However, there is a dearth of research on what precisely makes a learning tool perceived useful or perceived easy to use (Aljukhadar & Senecal, 2015; Alsabawy et al., 2016); namely, little attention has been paid to the determinants of PU and PEoU, neither their antecedents (Aziz & Macredie, 2005; Reinicke & Marakas, 2005; Venkatesh & Davis, 1996). With this gap existing, it is understudied that what contributes to users' resistance in using a specific educational technology, as well as how to address the hindrances. The lack of study is even more obvious when comes to user acceptance of gamification, as it is a new research area.

It is of great significance to better understand PU and PEoU, the two main determinants of technology acceptance, in educational gamification contexts (Venkatesh & Davis, 1996). Besides, Baaren et al. (2008) highlighted that "it is important to realise that the meaning of perceived usefulness … has an expiry date", so even though the determinants of PU and PEoU of technology have been investigated (which is not in gamification studies), latest research examining the determiants is also important.

This study aims to explore the determinants of teachers' PU and PEoU in the context of using gamified learning tools (GLT) for foreign language learning (FLL) in secondary schools. The current research is a scale development study that employed the exploratory sequential mixed method. This study helps understand the contributing factors of acceptance intention, provide actionable recommendations for gamification design for FLL, and eventually contribute to improved acceptance intention of gamification for FLL.

This study focused on gamified learning tools that will be initialled by teachers in the learning process: teachers play a critical role in introducing pedagogical innovations in the classroom, especially when the innovations are related to new technologies (José Martí-Parreño et al., 2016). Besides, this study selected foreign language learning (FLL) as the subject. Gamification is a field highly dependent on content, so the selection of subject is an issue. Foreign language is a special subject among the ones in secondary schools: it involves repetitive drills that are not necessarily engaging, and it requires students to be equipped with prior knowledge such as vocabulary; therefore, once the students were left behind, they are highly possible to be frustrated in catching up (Purgina et al., 2020). Therefore, the motivation maintenance and engagement triggering in FLL is of great importance (Purgina et al., 2020).

This study aims to answer the following three research questions (RQs):

RQ 1: What are the determinants of teachers' perceived usefulness (PU) of gamified learning tools (GLT) for foreign language learning in formal education? What items should be contained in the correspondent measuring scale PU-GLT?

RQ 2: What are the determinants of teachers' perceived ease of use (PEoU) of gamified learning tools (GLT) for foreign language learning in formal education? What items should be contained in the correspondent measuring scale PEoU-GLT?

RQ 3: How reliable and valid are the scales PU-GLT and PEoU-GLT?

8.1.1. Framework establishment process

Based on previous literature, this author conceptualised two frameworks that indicate the determinants of PU and PEoU for gamified learning tools (GLTs) in foreign language learning contexts, which were abbreviated as PU-GLT and PEoU-GLT.

Data elimination was one main challenge during the qualitative data analysis. Since the current study focuses on teachers' acceptance intention to gamified learning *tools*, the determinants focusing on *individual characteristics* were excluded (e.g. experience, personal innovativeness, self-efficacy, gender, etc.); since this study highlights the initial role of teachers, the functions mainly valued in independent learning were excluded, too (e.g. social interaction, hedonic experience, etc.); "ease of use" and "complexity" were removed, since in the original frameworks they were nominated as determinants of PU, which was not the research focus of the current study; comparatively less-discussed concepts were removed, too, such as critical mass (Bourgonjon et al., 2013) and tutor attribute (Teo, 2011b).

Terminology unification was another challenge. Besides the simple ones, such as summarising "compatibility with discipline", "curriculum relatedness" and "major relevance" as "compatibility", there were challenges too. Grounded theory analysis techniques were used for challenging terminology unification, such as summarising "time needed to learn", "training", and "ease of understanding" as "effort expectancy". This was a dynamic process that involved multiple rounds of modification. The results were confirmed by consulting two experts in the field.

8.1.2. The framework establishment for PU-GLT

In organising the determinants of PU-GLT, seven themes emerged: compatibility, curriculum relatedness, learning opportunities, PU-Process, PU-Product, social influence, and complexity, as shown in Table 37 and Table 38.

Compatibility can be interpreted as a theme that contains various connotations, such as compatibility with discipline, compatibility with workflow, compatibility with teacher knowledge, compatibility with infrastructures, compatibility with devices, etc. Therefore, compatibility can be regarded as the determinant of both PU and PEoU (Raza et al., 2017). Previous studies showed a tendency of laying focus on the compatibility of educational technologies with PU (Beggs, 2000; Rahmi et al., 2018), so as the current research. What compatibility means in the contexts of gamified learning tools for FLL needs further investigation, as will be detailed in the interview response analysis in the current study.

This author specified "curriculum relatedness" as an obvious theme, even though it can be categorised as a sub-theme "compatibility with the curriculum". Digital games have been reported ill-integrated with the curriculum content, which is outlined as one main negative influence on teachers' acceptance of educational video games (Koh et al., 2012; McFarlane et al., 2002; Prensky, 2001; Proctor & Marks, 2013; Tuparova et al., 2018; Van Rosmalen et al., 2015). Koh et al. (2012) stressed that this finding is limited to digital games. The integration issue applies to gamification as well, as highlighted by Tuparova et al. (2018), De Grove et al. (2012), and Sánchez-Mena and Martí-Parreño (2017).

An eye-catching framework modelling split PU into two individual concepts: the perceived usefulness in supporting the learning process the perceived usefulness in helping achieve outcomes. When talking about usefulness, people tend to connect it with academic performance improvement, as a large number of previous studies chose to assess the effectiveness or usefulness of a gamified learning tool by comparing participants' pre-and-post assessment results (Homer, Hew, & Tan, 2018; Hong et al., 2020; Kurniawan et al., 2019; Y. W. Lam, K. F. Hew, & K. F. Chiu, 2018; Loewen et al., 2019; Rachels & Rockinson-Szapkiw, 2018; Reynolds & Taylor, 2020). Bourgonjon et al. (2009) criticised that it is too narrow to focus on academic outputs only when investigating the PU of video games in the classroom, because "education means more than mere outcome" (p. 653). Accordingly, Bourgonjon et al. (2009) divided the perceived usefulness into "perceived usefulness-product" (PU-Product) and "perceived usefulness-process" (PU-Process). PU-Product refers to the perceived usefulness in helping achieve expected outcomes, and PU-Process refers to the perceived usefulness in supporting the learning process (Bourgonjon et al., 2009). To minimise difficulties in understanding, this author reworded these two themes as PU-Product and PU-Process.

Learning opportunities was removed in framework conceptualisation because this author reviewed it as a sub-theme of PU-Product. Complexity was removed, too, because this author preferred to regard it as one determinant of PEoU rather than PU. Complexity was nominated because PEoU can be regarded as the determinant of PU, such as in the original TAM model (Davis et al., 1989) and in the study of Baptista and Oliveira (2019).

This author combined the concepts of subjective norm and social image as social influence. Subjective norm refers to the belief that an important person or group of people will approve and support a particular behaviour (Venkatesh & Davis, 2000). Individuals often respond to social normative influences "to establish or maintain a favourable image within a reference group" (Venkatesh & Davis, 2000, p. 189), so social image is defined as the degree to which using technology is perceived to enhance one's status. Though these two concepts were individual determinants of perceived usefulness in the TAM2 model by Venkatesh and Davis (2000), the authors grouped them as "social forces" in the "social influence processes" in the technology acceptance decision (187). The term "social influence" was adopted in the current study.

8.1.3. The framework establishment for PEoU-GLT

In organising the determinants of PEoU-GLT, four themes emerged: availability, effort expectancy, teacher control, and support (see Table 37 and Table 39). The theme "affordability" of gamified learning tools for FLL emerged from the interview, so affordability was added too in the PEoU-GLT framework.

Effort expectancy is an obvious theme in interview responses. The availability of gamified learning tools can be regarded as a sub-theme of effort expectancy because the lack of available tools indicates more effort input in searching or designing a suitable one. However, interview responses revealed that the availability of suitable gamified learning tools is a huge challenge, so it was listed as an individual determinant.

The theme "affordability" of gamified learning tools emerged, so it was added too in the PEoU-GLT framework. Interview responses showed that "there are many free apps for gamified language learning" (P14), however, most of them are for self-directed learning. In the current research context, the teacher-leading gamification use for formal education, the supporting tools "can be costly" (P14).

8.1.4. The framework (PU-GLT and PEoU-GLT)

Eventually, this author obtained two theoretical frameworks for the PU-GLT and PEoU-GLT scales respectively, as illustrated in Figure 14.

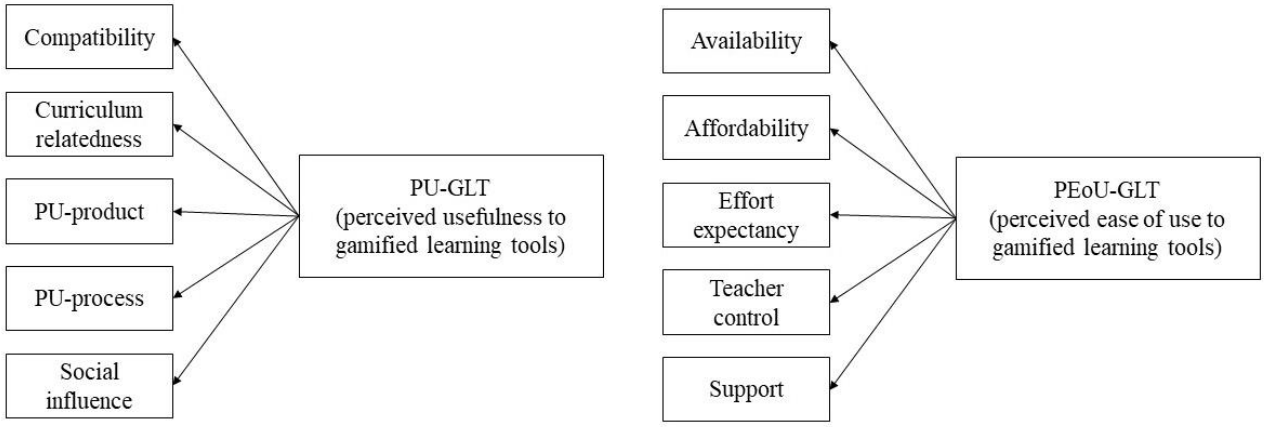

Figure 14. The theoretical framework: determinants of PU-GLT and PEoU-GLT

Table 37. Factors influencing teachers' acceptance intention to new technologies

Domain	Determinant	Synonymous term and reference Origin
PU	Compatibility	Compatibility (Rahmi et al., 2018; Y. Sun & Jeyaraj, 2013), compatibility with discipline (Beggs, 2000).
	Compatibility- curriculum relatedness	Curriculum relatedness (De Grove et al., 2012), major relevance (Park et al., 2012)
	Learning opportunities	Learning opportunities (Baptista & Oliveira, 2019; Bourgonjon et al., 2013; De Grove et al., 2012), multimedia instruction (Cigdem & Ozturk, 2016), course delivery (Teo, 2011b), learning environment (Teo, 2011b), knowledge learning (Lin & Chen, 2013)
	Perceived usefulness- process	Increased student interest (Beggs, 2000), flow (Herzig et al., 2012), affective needs (Hashim et al., 2015), enjoyment (Rahmi et al., 2018), user satisfaction (Aziz & Macredie, 2005), hedonic experience (Lin & Chen, 2013)
	Perceived usefulness- product	Improved students learning (Beggs, 2000), advantage over traditional teaching (Beggs, 2000), cognitive needs (Hashim et al., 2015), effectiveness (Aziz & Macredie, 2005), efficiency (Aziz & Macredie, 2005), performance expectancy (Venkatesh et al., 2003)
	Social influence/ Subjective norm	Social influence (Malaquias et al., 2018; Venkatesh et al., 2003), social needs (Hashim et al., 2015), socialness (Baptista & Oliveira, 2019), subjective norm (Bourgonjon et al., 2013; Park, 2009; Park et al., 2012; Rahmi et al., 2018; Teo, 2011a), colleague use (Beggs, 2000)
	Complexity/General PEoU	Ease of use (Baptista & Oliveira, 2019; Beggs, 2000), perceived ease of use (Teo, 2011a), complexity (Bourgonjon et al., 2013)
PEoU	Availability	equipment availability (Beggs, 2000), materials in discipline (Beggs, 2000), system accessibility (Park, 2009; Park et al., 2012), ease of finding (Brown, 2002), objective usability (Venkatesh & Davis, 1996)
	(Affordability)	n/a
	Effort expectancy	Time needed to learn (Beggs, 2000), computability with material (Beggs, 2000), training (Beggs, 2000), ease of understanding (Brown, 2002), effort expectancy (Venkatesh et al., 2003)

	Teacher control	Perceived learner control (Zhang et al., 2017), personal comfort (Beggs, 2000),
	Support	Administrative support (Beggs, 2000), facilitating conditions (Teo, 2011b; Venkatesh et al., 2003; Wong, 2016)
Individual factor	Experience	Experience (Bourgonjon et al., 2013; Bourgonjon et al., 2009; De Grove et al., 2012; Venkatesh et al., 2003), direct experience (Venkatesh & Davis, 1996)
	Personal innovativeness	Personal innovativeness in IT (Malaquias et al., 2018; Zhang et al., 2017), personal innovativeness (Bourgonjon et al., 2013)
	Self-efficacy	LMS self-efficacy (Cigdem & Ozturk, 2016), e-learning self-efficacy (Park, 2009; Zhang et al., 2017), computer self-efficacy (Venkatesh & Davis, 1996), mobile learning self-efficacy (Park et al., 2012), self-efficacy (Brandyberry, Li, & Lin, 2010; Brown, 2002; Lin & Chen, 2013; Rahmi et al., 2018), anxiety (Rahmi et al., 2018), computer anxiety (Brown, 2002)
Tool factor	Enjoyment (tool)	Enjoyment (Herzig et al., 2012; Rahmi et al., 2018), flow (Herzig et al., 2012)
	Interactivity	Interactivity (Cigdem & Ozturk, 2016), interaction (Rahmi et al., 2018), social interaction (Lin & Chen, 2013)
Others	Complexity/general PEoU	Ease of use (Baptista & Oliveira, 2019; Beggs, 2000), perceived ease of use (Teo, 2011a), complexity (Bourgonjon et al., 2013)
		Gender (Venkatesh et al., 2003), age (Venkatesh et al., 2003), voluntariness of use (Venkatesh et al., 2003)

Notably, previous studies were not consistent in determining whether certain determinants contribute to PU or PEoU. In a meta-analysis study that involved 203 studies for the factors affecting PU and PEoU in adopting e-learning systems, the factor "subjective norm" was the determinant of PU in 33 studies and the determinant of PEoU in 12 studies; similarly, the factor "enjoyment" was the determinant of PU in 12 studies and the determinant of PEoU in 13 studies (Rahmi et al., 2018). This phenomenon applies to the current study too, as the determinant "compatibility" emerged for both PU-GLT and PEoU-GLT, though with different connotations.

8.2. Scale construction: The PU-GLT and PEoU-GLT

Though this is a scale development study, existing validated scales were still employed when possible for guaranteed reliability, as suggested by Robertson (2017). However, due to the exploratory nature of the current study and the innovativeness of the current topic, only a limited number of survey items were adopted from previous scales.

In the PU-GLT scale, the first two survey items measuring curriculum relatedness were adopted from the scale by De Grove et al. (2012). Since it is suggested to prepare more than three survey items for each construct, two more items based on teacher interviews were added. The first three items measuring PU-Product were modified from the scale by Adukaite et al. (2017). Social influence is a combined theme of subjective norm and social image, which was measured by items modified from the TAM2 Measurement Scales by Venkatesh and Davis (2000). All other items were generated and modified from interview responses, as detailed in Table 38 and Table 39.

Almost all the survey items measuring PEoU-GLT were established based on interview responses since previous scales mainly focused on the PEoU of technology from the technological perspective while in the current context PEoU also includes the perceived ease of use of technology in the implementation process. Survey items were drafted, as detailed in the Appendices.

Table 38. Items of the PU-GLT scale

Construct	#	Survey item
Compatibility	1	The gamified FLL tool is suitable for my class in consideration of the teacher-student ratio.
	2	It is easy to find time to use the gamified FLL tool in my class.
	3	The gamified FLL tool is suitable for the type of knowledge I teach (e.g., memorisation, application, analysis, etc.).
	4	I know how to use gamified learning tools in my class.
Curriculum relatedness	1	The gamified FLL tool fits the curriculum.
	2	It is clear how the gamified FLL tool can be used to fit the curriculum.
	3	The content provided by the gamified FLL tool is relevant to the curriculum.
	4	The content provided by the gamified FLL tool is relevant to the assessment.
Usefulness-product (PU-Product)	1	The gamified FLL tool better visualises the knowledge points.
	2	Using the gamified FLL tool helps get higher scores in the assessment.
	3	Using the gamified FLL tool helps get better academic achievements.
	4	The gamified FLL tool fosters students' awareness about the learning objectives.
	5	The gamified FLL tool enhances students' knowledge about the learning objectives.
	6	The gamified FLL tool develops a better understanding of the learning objectives.
Usefulness-process (PU-Process)	1	The gamified FLL tool is fun.
	2	The gamified FLL tool is engaging.
	3	The gamified FLL tool increases students' interests.
	4	Students are happy when using the gamified FLL tool in schools.
	5	Teachers enjoy using the gamified FLL tool in schools.
Social influence	1	People who influence my behaviour think that I should use gamified learning tools.
	2	People who are important to me think that I should use gamified learning tools.
	3	People in my organization who use gamified learning tools have more prestige than those who do not.
	4	People in my organization who use gamified learning tools have a high profile.
	5	Using gamified learning tools is a status symbol in my organization.

Table 39. Items of the PEoU-GLT scale

Construct	#	Survey item
Effort expectancy	1	There are good gamified learning tools for FLL. I don't need to design or modify too much.
	2	There is a systematic gamified learning curriculum for FLL. I can just follow it.
	3	I am in control of classroom discipline when using gamified learning tools for FLL.
	4	We have the infrastructures to support gamified FLL classes (e.g., Wi-Fi and tablets)
	5	Students have digital devices and are able to use them in school.
Support	1	It is easy to find time to use the gamified FLL tool for FLL.
	2	If I would like to use gamified learning tools in school, I would not come across difficulties.
	3	We have technicians to help solve problems using gamified learning tools.
	4	My colleagues appreciate my attempts in using gamified learning tools.
	5	Parents are supportive of students using gamified learning tools in school.
	6	School policies are supportive of using gamified learning tools in school.
Affordability	1	Cost is not an obstacle for students or parents in using gamified learning tools.
	2	Cost is not an obstacle for the school in using gamified learning tools.
	3	The gamified learning tools are affordable.
Teacher control	1	I know how to select suitable gamified learning tools.
	2	I know how to use gamified learning tools in school.
	3	I am in control of the learning content when using gamified learning tools.
	4	I am in control of the learning pace in using gamified learning tools.
Availability	1	If I want to use gamified learning tools, I can easily find some.
	2	There are mature, useable and effective gamified learning tools.
	3	The gamified FLL tool can cover multiple aspects of foreign language learning.

8.3. Scale evaluation

8.3.1. Suitability assessment of the data

Before the data analysis, scale developers are suggested to conduct an initial examination of items' performance to confirm whether the data is suitable for factor analysis (DeVellis, 2003; Fabrigar et al., 1999). According to Pallant (2016), there are two main concerns affecting the results: sample size and the strength of the inter-correlations among items.

The adequate sample size for factor analysis has been discussed in the section "4.6. Step 6: Administer items to a development sample", which is 200 or more in the current study.

In regard to the strength of the inter-correlations among items, Pallant (2016) suggested using two statistical measures generated by IBM SPSSS to help assess the factorability of the data: Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. The dataset was confirmed suitable and sufficient for factor

analyses since Bartlett's test was significant ($\chi 2(190) = 2725.408$, $p < .001$) and the overall Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (MSA) was good at .901 (cut-off criterion: .50. KMO higher than .80 is excellent).

Besides, Tabachnick and Fidell (2013) recommended a benchmark of .30 as to the coefficients in the correlation matrix, which reveals the strength of the inter-correlations among items.

8.3.2. The evaluation of the PU-GLT scale

After confirming that the data set is suitable for factor analysis, an exploratory factor analysis (EFA) was conducted using Principal Components Analysis (PCA) and Varimax rotation (see Table 40). Following suggestions of Costello and Osborne (2005), five items have been removed due to low communalities (lower than or close to .50): PU_product-3 (.449), PU_process_4 (.499), PU_process_5 (.464), PU_product_1 (.482), and PU_product_4 (.501). The items presented cross-loading problem were also excluded, leaving a scale with 15 items.

EFA analysis extracted four factors with eigenvalues greater than one, which explained 69.19% of the total variance. Social influence as the domain factor explains 41.77% of the perceived usefulness of gamified learning tools, followed by curriculum relatedness (12.28%), PU-Process (9.16%), and PU-Product (5.97%).

In previous scale-development studies, the two common validity test measures are convergent validity and discriminant validity.

To assess convergent validity, three criteria were taken into consideration including Cronbach's alpha, composite reliability (CR), and the average variance extracted (AVE). Findings showed that the survey measuring all the four constructs demonstrated a high level of internal reliability, as the Cronbach's alpha values exceed the common threshold of 0.70 (Hair et al., 1998; Taber, 2018): curriculum relatedness (n = 4, α = .855), PU-Product (n = 3, α = .789), PU-Process (n = 3, α = .767), and social influence (n = 5, α = .855). The overall reliability of the original PU-GLT scale was .899. AVE values for all the factors were above the cut-off criterion 0.5. The CR values for all the four constructs were greater than the benchmark .70 (Nunnally, 1978), which furtherly confirmed the convergent validity of the current scale.

Two considerations were involved in determining discriminant validity: the square root of AVE and cross-loadings. The cross-loading of each item should be greater than the loading on other constructs. Besides, the square root of the AVE of each construct should be larger than the correlation of the specific construct with any of the other constructs (Fornell & Larcker, 1981). The two criteria were met, as shown in Table 41 and Table 42.

8.3.3. The evaluation of the PEoU-GLT scale

The evaluation of the scale for PEoU-GLT followed the same procedure. The data measuring PEoU-GLT showed that the Bartlett's test was significant ($\chi 2(55) = 798.043$, $p < .001$) and the overall KMO was good (.813).

The EFA analysis extracted three factors with eigenvalues greater than 1, which explained 62.23% of the total variance. The overall reliability of PEoU-GLT was satisfactory (α = .802) though the Cronbach's alpha for the construct "support" is comparatively low (α = .637). For the three constructs, the composite reliability values were larger than the threshold .70, and the AVE values were higher than the threshold .50 (see Table 43). The square root of the AVE of each construct was larger than the correlation with other constructs (see Table 44 and Table 45).

Table 40. EFA analysis results of the PU-GLT scale

Item	Communalities	Factor extracted			
		1	2	3	4
Curriculum_relatedness_1	0.724		0.766		
Curriculum_relatedness_2	0.626		0.662		
Curriculum_relatedness_3	0.732		0.771		
Curriculum_relatedness_4	0.754		0.812		
PU_product_3	0.679				0.710
PU_product_5	0.767				0.748
PU_product_6	0.731				0.722
PU_process_1	0.697			0.793	
PU_process_2	0.731			0.844	
PU_process_3	0.592			0.684	
Social_influence_1	0.673	0.738			
Social_influence_2	0.606	0.695			
Social_influence_3	0.738	0.830			
Social_influence_4	0.602	0.714			
Social_influence_5	0.726	0.778			
Total variance explained: 69.19%		41.772%	12.278%	9.164%	5.972%
Cronbach's alpha: .899		.855	.855	.767	.789
Composite reliability (CR):		.867	.840	.819	.771
Average variance extracted (AVE):		.566	.570	.603	.528

Table 41. The square root of AVE (in bold) and correlations between constructs in PU-GLT

Latent construct	#	AVE	Correlation/Square root of AVE			
			1	2	3	4
Curriculum relatedness	1	0.566	**0.752**			
PU-Product	2	0.57	.632**	**0.755**		
PU-Process	3	0.603	.448**	.394**	**0.777**	
Social influence	4	0.528	.503**	.530**	.315**	**0.727**

Table 42. Correlation matrix of items in the PU-GLT scale

Item	#	1	2	3	4	5	6	7	
Curr_Rel_1	1	**1**							
Curr_Rel_2	2	**.574****	**1**						
Curr_Rel_3	3	**.678****	**.559****	**1**					
Curr_Rel_4	4	**.599****	**.519****	**.641****	**1**				
PU Product_3	5	.422**	.364**	.401**	.296**	**1**			
PU Product_5	6	.506**	.486**	.477**	.454**	**.494****	**1**		
PU Product_6	7	.472**	.429**	.498**	.486**	**.473****	**.691****	**1**	
PU process_1	8	.415**	.348**	.345**	.226**	.337**	.256**	.291**	
PU process_2	9	.230**	.364**	.300**	0.113	.331**	.164**	.165**	
PU process_3	10	.340**	.403**	.360**	.287**	.335**	.269**	.327**	
SN_1	11	.405**	.281**	.388**	.388**	.316**	.355**	.312**	
SN_2	12	.311**	.407**	.318**	.336**	.340**	.360**	.391**	
SN_3	13	.327**	.181**	.386**	.348**	.311**	.353**	.381**	
SN_4	14	.309**	.298**	.332**	.322**	.391**	.363**	.309**	
SN_5	15	.324**	.264**	.333**	.419**	.262**	.435**	.421**	

		8	9	10	11	12	13	14	15
Curr_Rel_1	1								
Curr_Rel_2	2								
Curr_Rel_3	3								
Curr_Rel_4	4								
PU Product_3	5								
PU Product_5	6								
PU_Product_6	7								
PU process_1	8	**1**							
PU process_2	9	**.572****	**1**						
PU process_3	10	**.539****	**.460****	**1**					
SN_1	11	.316**	.180**	.323**	**1**				
SN_2	12	.288**	.219**	.344**	**.651****	**1**			
SN_3	13	.154*	0.116	.258**	**.570****	**.476****	**1**		
SN_4	14	.203**	.132*	.312**	**.441****	**.458****	**.612****	**1**	
SN_5	15	0.114	-0.044	.250**	**.483****	**.499****	**.689****	**.547****	**1**

Table 43. EFA analysis results of the PEoU-GLT scale

Item	Communalities	Factor extracted		
		1	2	3
Effort_expectancy_1	0.607	0.775		
Effort_expectancy_2	0.658	0.758		
Effort_expectancy_3	0.561	0.719		
Effort_expectancy_4	0.692	0.825		
Effort_expectancy_5	0.617	0.759		
Support_1	0.652			0.779
Support_2	0.534			0.666
Support_4	0.572			0.730
Affordability_1	0.622		0.766	
Affordability_2	0.645		0.757	
Affordability_3	0.684		0.821	
Total variance explained: 62.229%		33.663%	18.936%	9.629%
Cronbach's alpha: .802		.838	.727	.637
Composite reliability (CR)		.878	.825	.770
Average variance extracted (AVE)		.590	.611	.528

Note: Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Table 44. The square root of AVE (in bold) and correlations between constructs in PEoU-GLT

Latent construct	#	AVE	Latent construct		
			1	2	3
Effort expectancy	1	.590	**.768**		
Support	2	.611	.337**	**.782**	
Affordability	3	.528	.168**	.429**	**.727**

Table 45. Correlation Matrix of the PEoU-GLT survey items

Item	#	1	2	3	4	5	6	7	8	9	10	11
Effort1	1	**1**										
Effort2	2	**.587****	**1**									
Effort3	3	**.420****	**.508****	**1**								
Effort4	4	**.537****	**.477****	**.532****	**1**							
Effort5	5	**.459****	**.508****	**.487****	**.581****	1						
Support1	6	0.106	.249**	.271**	.137*	**.227****	**1**					
Support2	7	.281**	.380**	.172**	.177**	**.282****	**.414****	**1**				
Support4	8	0.096	.162*	.176**	0.054	**.148***	**.408****	**.313****	1			
Afford1	9	.132*	.131*	0.069	0.070	.171**	.232**	.281**	**.261****	**1**		
Afford2	10	.132*	0.059	0.118	0.108	.188**	.345**	.281**	**.294****	**.467****	**1**	
Afford3	11	0.038	0.053	0.028	0.099	.142*	.253**	.183**	**.251****	**.466****	**.490****	**1**

8.3. Discussion

Items measuring the compatibility of gamified learning tools were removed, including the compatibility to class size, teaching schedule, type of knowledge, and teachers' knowledge. The items were not captured as ones measuring one factor, namely, there are various determinants contributing to the compatibility of gamified learning tools. The survey items were also criticised for being not easy enough to understand. For example, the "type of knowledge" was categorised based on Bloom's revised taxonomy, which includes remembering, understating, application, analysis, evaluation, and creation (Anderson & Bloom, 2001); however, though Bloom's revised taxonomy was a familiar theory for experts, it was a complicated one that needs further explanations for the participants. In the expert interview, this issue has been discussed but remained unsolved, which is possible to bring negative influence to the development of the scale. Participants' understanding of the keyword "teacher knowledge" may also vary: some may interpret it as the knowledge of the complex process of gamification implementation that involves tool selection and content adjustment, while some others may interpret it as the awareness of a potential gamified FLL tool. Different interpretations lead to varied responses, which results in data with high variance and low reliability. The results indicated that the compatibility of gamified learning tools contains various sub-factors, which need further investigation and discussion.

All the survey items measuring curriculum relatedness were included, including whether the gamified FLL tool fit the curriculum, whether it is clear how the gamified learning tools can be used to fit the curriculum, whether the provided content is relevant to the curriculum, and whether the provided content is relevant with the assessment.

The perceived usefulness of the outcomes (PU-Product) highlighted gamified learning tools' functions in triggering cognitive engagement, such as fostering knowledge and enhancing understanding; however, the item indicates "using the tool helps get better academic achievement" was removed. Data of the current study indicated that cognitive engagement and academic achievement in the current context might not indicate the same thing. *Cognitive engagement* is defined as a psychological state in which students are willing to put time and efforts to "truly understand a topic" and in which students "persist studying over a long period of time" (Rotgans & Schmidt, 2011, p. 465), which includes "less observable" and "more internal" indicators, such as self-regulation, the value of learning, personal goals and learning autonomy (Appleton et al., 2006, p. 429). Improved cognitive engagement does not equivalent to improved academic achievements. Previous studies presented the tendency of assessing the effectiveness of gamified learning tools by measuring students' academic achievement, which can be replaced by the assessment of students' cognitive engagement level.

The fourth item measuring PU-Process was removed, which is to investigate whether "teachers enjoy using gamified learning tools in schools". It is interesting to see that the participants gave the lowest score on this item (Mean = 5.01, while that for other items measuring PU-Process range from 5.40 to 5.58), indicating that they do not believe the gamification process is as enjoyable for teachers as it is for students. Expert 2 commented that the gamification implementation in China is still at the initial stage, which involves practical difficulties, so teachers, including those who are highly interested, are possible to be frustrated in using gamified learning tools. Since the applied data collection method was convenience sampling, the current study involves teachers who are not that interested in gamification, so the results tend to be even more pessimistic. Correspondently, the perceived benefits of using gamification are important for teachers, which should be addressed in improving teachers' acceptance intention of this innovative technology.

Items measuring social influences were included. Teachers believed that using gamified learning tools in class is a symbol of being creative and "in the current", which is positive for their teaching career. What's more, the schools using gamification were also believed to be more capable or international.

Effort expectancy is the main determinant of a perceived easy-to-use gamified FLL tool. Teachers expect a mature tool that can be used in teaching without complicated design or modifications; they also expected a systematic gamified learning curriculum for FLL which enables them to "just follow it". Teachers did not want troubles in classroom discipline management, nor potential negotiation issues such as those connected with infrastructures.

Support from technicians was nominated as the factor influencing the PEoU of gamified learning tools, though it was excluded from the current scale (item Support_3). Interview responses showed that gamification implementation in China now is the trail initiated by individual teachers rather than a top-down revolution supported by schools or the government, so technical support is absent. What's more, the use of gamified learning tools for FLL is not only a technical issue: it highlights educational designs that combine gamification features with pedagogical needs; therefore, what teachers need are not technicians but experienced teachers who know how to achieve meaningful gamification for enhanced learning engagement or outcomes. Support from parents was tested statistically irrelevant with other survey items, showing that though teachers and the school might appreciate the use of gamification in FLL classrooms, parents might not. The phenomenon can be caused by the negative stereotype of gamification or the lack of school-parent communications.

The affordability of gamified learning tools is reported as an issue. Though there are increasingly more free tools for gamified language learning, such as Duolingo, the ones customised for teacher-leading and classroom use are scarce. There is a correspondent dilemma: the number of schools that would like to purchase the gamified learning tools is limited, so the tool development cost cannot be lowered; at the same time, the financial cost on gamified learning tools remains high, so schools tend to be reluctant in purchasing them.

The EFA analysis results removed two constructs for PEoU-GLT: teacher control and tool availability. Teachers' perceived lack of control can be addressed by providing teacher education on gamification knowledge, especially practical implementations. Though there are lots of options in the market, it was still reported that there is a lack of *suitable* gamified learning tools to support formal education in secondary school classrooms, which can be addressed by paying attention to the needs of secondary school teachers in designing the gamified learning tools.

CHAPTER 9

Conclusion

09

Conclusion

This chapter is the conclusion of the current thesis, which consists of four main sections: the section briefing the research findings, the section introducing implications, limitations, and discussions on future studies.

CHAPTER NINE: CONCLUSION

9.1. Summary

The findings of this thesis consist of four main parts, which include a pilot study, an interview of teachers' perceptions of gamified learning tools in secondary schools, a scale development aimed to measure teachers' technology acceptance intention to gamified learning tools, and two scales that measure teachers' perceived usefulness and perceived ease of use of gamified learning tools in secondary schools (PU-GLT and PEoU-GLT).

This is a cross-sectional study that employed the exploratory sequential mixed method, which combines the collection and analysis of both qualitative and quantitative data in a sequence of phases (Creswell & Clark, 2017). The qualitative data were collected via interviews, which contributed to the summary of teachers' perceptions of gamified learning tools and the construction of three scales (the scale measuring teachers' technology acceptance intention to gamification, the PU-GLT scale, and the PEoU-GLT scale); the quantitative data were collected via surveys, which helped evaluate the designed three scales.

The interviews were conducted to collect teachers' opinions and comments regarding to the benefits, concerns, and barriers in implementing gamified learning tools in secondary schools. The interview responses were analysed with the use of grounded theory techniques, which contributed to the construction of scales in the following two sections.

The interview responses revealed four factors contributing to teachers' acceptance intention of gamified learning tools, which contain 16 themes. The four factors are teacher factors, tool factors from the educational perspective, tool factors from the IT perspective, and external supports. The 16 themes are teacher control, teacher skill, teacher attitude, risk avoidance, resource availability, compatibility, curriculum relatedness, perceived usefulness on process (PU-Process), perceived usefulness on academic outcomes (PU-Product), cost-effectiveness, perceived ease of use, cost, negative effects, gamification design, interface aesthetics, and students', parental and administrative support.

This author also summarised a framework describing factors contributing to teachers' acceptance intention of gamified learning tools. The framework shows that four factors contribute to teachers' overall attitude, which leads to behavioural intention and the actual acceptance. The four factors are perceived enjoyment, perceived usefulness (PU), perceived ease of use (PEoU), and perceived risks. There are sub-factors that furtherly explain PU and PEoU: PU contains relatedness, perceived usefulness on process, perceived usefulness on academic outcomes, and social image; PEoU contains accessibility, compatibility, complexity, and user control.

Based on the interview responses and the summarised framework, this author constructed and validated a two-dimensional scale measuring teachers' perceived usefulness (PU) and perceived ease of use (PEoU) of gamified learning tools (GLT) in secondary schools. The two scales were named PU-GLT and PEoU-GLT respectively (see Appendix B and C). Data analysis showed that there are four determinants of PU-GLT (curriculum relatedness, PU-Product, PU-Process, and social influence) and three determinants of PEoU-GLT (effort expectancy, support, and affordability). Determinants studies provide valuable information regarding the understanding and improvement of user technology acceptance intention.

9.2. Limitations

One main limitation of this study is that the interview study involved the convenience sampling approach, which may be biased, but this could be minimised by selecting participants with related gamification experiences. Moreover, though the scale version one has been confirmed, the number of items measuring the sub-constructs is limited: according to MacCallum et al. (1999) and Raubenheimer (2004), there should be at least three items for each construct for further studies. This is also why in shortening the scale some assuming-important items were excluded from the statistical analysis, such as the items measuring the relatedness of gamified learning tools.

Unavoidably, as almost all the studies, the current one does not "exhaustively enumerate all potential variables" of technology acceptance intention (Jiang, Chen, & Lai, 2010, p. 95), even though the decision process can be shaped by numerous factors (Quazi & Talukder, 2011). More research investigating gamification acceptance intention from different perspective are welcomed.

What's more, this study did not conduct model-fit tests based on the validated scales, because according to DeVellis (2003), model-fit tests are not essential for scale development studies. However, it is still a limitation, which can be addressed in future studies.

9.3. Implications

This thesis generated five main outputs:

1) A redefinition of gamified learning and gamified learning tools.
2) A summary of factors influencing whether secondary teachers accept or resist using gamified learning tools.
3) A theoretical framework illustrating factors that influence teachers' acceptance intention to gamified learning tools.
4) A validated scale that measures teachers' acceptance intention to gamified learning tools (the TAI-GLT scale).
5) Two validated scales that measure teachers perceived usefulness (PU) and perceived ease of use (PEoU) of gamified learning tools in English-as-a-second-language learning contexts (the PU-GLT scale and PEoU-GLT scale).

The summarised theoretical framework and the correspondent interview response summary contribute to the description of factors influencing whether secondary teachers accept or resist using gamified learning tools. As suggested by Sánchez-Mena and Martí-Parreño (2017), managers or principals of all levels of educational institutions should pay attention to the hindering factors in teachers' acceptance intention and provide the corresponding supports if they are interested in implementing gamified learning. To be specific, the managers or principals could select the gamified learning tools that enable teachers to take control of the learning process, provide teacher training on gamification knowledge, help teachers select gamified learning tools that are compatible with the current teaching system, choose gamified learning tools that are related to curriculum and perceived effective, as well as providing financial supports. What's more, the managers or principals should pay special attention to the avoidance of negative effects brought by digital devices.

Findings of the interviews can also serve as detailed explanations of theoretical frameworks for technology acceptance, such as the Technology Acceptance Model (TAM) that contains perceived usefulness and perceived ease of use. Even though a large number of studies have been conducted based on the TAM model, the connotation of perceived usefulness and perceived ease of use could be significantly different in gamification contexts. Future researchers can also conduct quantitative analyses based on findings generated in the current study. One main limitation of this exploratory study is it involved the convenience sampling approach, which may be biased, but this could be minimised by selecting participants with related gamification experiences.

The three validated scales can be directly used for data collection. They can also be modified to fit varied contexts.

9.4. Future studies

Future studies are suggested to investigate the determinants of core constructs contributing to users' acceptance intention of gamified learning tools, as well as the correspondent survey items. As suggested by Baydas and Cicek (2019), the developed scales could be used to test different gamified learning tools for different courses and education levels to provide a wider range of results. General scales measuring the determinants of gamified learning tools in contexts other than language learning are also needed. Future studies can also aim to reduce the survey items or access a larger sample size. Compared with short-term data collection in labs, long-term data collection in real-world teaching practices is also suggested (Nacke & Deterding, 2017).

New surveys can be established based on other technology-acceptance-related frameworks or theories, such as the Diffusion of Innovation Theory (DIT) (Rogers, 2003). The DIT theory is a popular and mature adoption model in the last three decades (Dearing, 2009; Sahin, 2006), which describes how innovations would diffuse within the social system through communicative processes and how individuals eventually accept an innovation (Celik, Sahin, & Aydin, 2014). Or, to make it simple, the DIT theory "seeks to explain how an innovation is adopted in a population" (Shaban & Egbert, 2018, p. 235). The DIT theory has been nominated as a promising one in helping "gain insight into factors that influence the adoption" of educational innovations (Ball, Ogletree, Asunda, Miller, & Jurkowski, 2014, p. 240), which can be valued in future gamification-related scale-development studies.

What's more, align with what this author attempted to do, researchers proposed that the contributing factors to users' technology acceptance intention can be classified into different groups, such as tool factors, individual factors, organisation factors, social factors, and context-related factors (Luo et al., 2021; Park et al., 2012; H.-K. Wu et al., 2008). Future studies can dig into each type of factors for better understanding of user technology acceptance intention. For example, Rogers (2003) specified five attributes of innovations that influence users' technology acceptance intention, which include relative advantage, compatibility, complexity, trialability, and observability. These five attributes describe the nature of the technology (rather than the perception of individuals or the influence of contexts), so addressing these attributes can directly influence the final acceptance intention or acceptance behaviour. New scales can be developed based on the five attributes of innovations theory.

In completing the current thesis, this author also recognised several more themes for future studies:

Discuss more on the new definition of game elements and gamification mechanisms. In the current study, this author redefined game elements as "the obvious game-like elements that are frequently used in typical digital or board games"

and gamification mechanisms as "the underlying guidelines that make an activity engaging, which originate from human needs and desires". More studies are needed to discuss the rationale of the redefinition.

Focus on educational content rather than game elements. Research findings revealed that a portion of previously involved gamified learning tools were poor in integrating gamification features with pedagogical content. Rego (2015) criticised that in some trials, learners must already have some knowledge to participate in the activity, which makes the process nothing educational. Similarly, in some cases, the researchers are providing "chocolate-covered broccoli" by simply adding game elements (e.g. colourful graphics and animation) to "dull and repetitive tasks" (Purgina et al., 2020, p. 128), which makes gamification essentially "not very different from those of the traditional educational settings consisting of blackboards and textbooks" (Sendurur et al., 2017, p. 119).

One solution is to focus on educational content rather than game elements. It was suggested that gamification should be a "holistic, creative and structured *process*" (Baldauf et al., 2017), or a "good and careful design on *learning materials*" (Kurniawan et al., 2019). It was not specified how to gamify the learning process or learning materials, which should be addressed in future studies. Another solution is to understand how game elements should be "selected, deployed, implemented, and integrated" to provide "gameful experience" in pedagogical practices (Werbach, 2014). However, it is still not clear how to push the frontier further to achieve meaningful gamification.

Investigate the impact of an individual game element. Many studies treat gamification as a uniform concept, while in practice, the specific designs and realizations of gamification environments can be quite diverse. Since gamification can take many forms and can combine game design elements in many different ways, it is inappropriate to study the motivational effects of gamification as a generic construct (Sailer et al., 2017; Seaborn & Fels, 2015). Therefore, researchers should experimentally investigate the effectiveness of individual game elements on learning motivation and performance (Deterding, Sicart, et al., 2011; Linehan, Kirman, Lawson, & Chan, 2011; Mekler, Brühlmann, Tuch, & Opwis, 2017). Researchers could achieve this goal by removing certain game elements from a gamified learning tool and compare the results of controlled and experimental groups (Seaborn & Fels, 2015).

Previous studies presented incomparability issues since they involved varied game elements in different contexts for distinguished purposes (Dindar et al., 2020). Accordingly, future studies can explore the impact of an individual game element, as well as filtering the unsuitable game elements for educational purposes. Researchers also concluded that gamification was considered positive "when specific requirements are met" (Baldauf et al., 2017), so more efforts are expected to explore what specific requirements are needed for a successful gamification implementation.

Establish automatic assessment systems in e-learning platforms. As concluded by DomíNguez et al. (2013), gamification has some limitations when students cannot get immediate feedback, therefore, the existence of automatic evaluation or unsupervised scoring system is important in e-learning platforms. The automatic assessment system could be advanced into one providing adaptive tasks which are tailored to students' ability (Lee & Hammer, 2011).

Investigate how gamification impacts different types of students. Though some researchers are positive about the effectiveness of gamification in education, they still call for cautions in implementing this approach as not everyone benefits from it (Buckley & Doyle, 2017; Hwang, Sung, Hung, Huang, & Tsai, 2012). There is a need to study learner differences and how it influences the effectiveness of gamification (Straub, 2009).

Investigate the methods promoting cognitive engagement in gamified education. The results show that previous studies have laid too much attention on measuring students' behavioural and affective engagement with cognitive engagement understudied (e.g., focused attention, self-reflection, analysis, evaluation, etc.) (Hew et al., 2016). Future

studies could lay more emphasis on the exploration of four issues, including “is there any method to assess the cognitive engagement in gamified learning”, “whether the involvement of behavioural and affective engagement connected with positive learning outcomes or learning experiences”, “is there any method to assess the cognitive engagement in gamified learning” and “how reliable and valid are the assessments of cognitive engagement in gamified education”.

Explore the techniques of gamifying class without digital games. As highlighted by Hanus and Fox (2015), one important conceptual implication of gamification is that ‘game’ is not necessary for gamified learning. Rather than using digital games or other forms of games, teachers could gamify the learning process and therefore “make the class itself a game” (Hanus & Fox, 2015, p. 152). In this way, gamification could break the constraints of game features such as high cost, heavy reliance on technology, time-wasting and poor integration with the curriculum (Cain & Piascik, 2015; Fernandes et al., 2012; Lee & Hammer, 2011; B. Morris et al., 2013).

The establishment of authentic learning environments in gamification studies. Gafni et al. (2017) and Purgina et al. (2020) mentioned the voluntariness issue during the experiment progress. Even though the participation was fully voluntary, the participants might be reluctant in completing the tasks because they are not comfortable or familiar with the presented materials (Gafni et al., 2017). Therefore, there is a need to establish authentic learning environments in experimental settings. It applies to teachers, too, as they need the opportunities to see how gamification functions in authentic learning contexts as well as how to apply the gamification knowledge in real-world classroom teaching (Shaban & Egbert, 2018).

More aspects have been highlighted in previous studies, including studying how gamification strategies improve language learning (Rego, 2015), how individual game element functions in the learning process (B. J. Morris et al., 2019), how gamification promotes behaviour change in the learning process (Rego, 2015), what negative impacts of gamification on students (Lu, 2008), how to gamify the learning process rather than simply adding game elements (Werbach, 2014), how users perceive the use of gamified learning tools (Honório et al., 2018), when and how to use gamification for the optimal effectiveness (Welbers et al., 2019), etc.

ABBREVIATION

&

GLOSSARY

ABBREVIATION

CFA: Confirmatory factor analysis

EFA: Exploratory factor analysis

FLL: foreign language learning

EFL: English-as-a-foreign-language

Gamified FLL tools: Gamified learning tools for foreign language learning

GBL: Fame-based learning

GLT: Gamified learning tool

PEoU: Perceived ease of use

PEoU-GLT: The perceived ease of use of gamified learning tools

PU: Perceived usefulness

PU-GLT: The perceived usefulness of gamified learning tools

PU-Process: The perceived usefulness of a tool about its contribution to the learning process

PU-Product: The perceived usefulness of a tool about its contribution to learning products

RQ: Research question

TAI-GLT: The scale measuring teachers' acceptance intention (TAI) of gamified learning tools (GLT)

TAM: Technology-acceptance model

GLOSSARY

- **Educational game.** Educational game refers the game software that is either developed explicitly as being educational or the game software that can be used for educational purposes (Berg Marklund, 2015).
- **Edutainment.** Edutainment is the combination of educational software "lightly sprinkled with game-like interfaces and cute dialog" (Zyda, 2005, p. 29).
- **Explicit gamification.** Gamification involving strategies that utilise applications that are obviously game-like (Chou, 2015).
- **Game design element.** "An essential or characteristic part of making sense to an activity where one or several players agree to voluntary participate in playing according to rules, in order to achieve a defined goal and respond to a represented feedback system. More clearly speaking, game design elements are the individual characteristics used to create a meaningful game" (Ranz, 2015, p. 10). The basic building blocks of gamification applications, which are largely equivalent to game design patterns (Sailer et al., 2017, p. 372).
- **Game mechanics.** Game mechanics means the interaction with games and the gamer focusing on internal management of interactions.
- **Game-based learning (GBL).** Game-based learning is to describe the act of "conducting pedagogical work with a game as the core educational tool" (Berg Marklund, 2015, p. 16).
- **Gamification.** Gamification is use game mechanisms or game elements in game contexts to non-game contexts (Deterding, Sicart, et al., 2011). It is defined as "the use of design (rather than game-based technology or other game-related practices) elements (rather than fully developed games) characteristic for games (rather than play or playfulness) in non-game contexts (regardless of specific usage intentions, contexts, or implementation media)" (Deterding, Sicart, et al., 2011, p. 5).
- **Implicit gamification.** Human-focused design that utilises game elements in non-game contexts (Chou, 2015).
- **Item**. "An item is a single question or statement in a questionnaire and is usually the smallest building block of a questionnaire, module, scale or test" (Robertson, 2017, p. 7). .
- **Lite games.** Lite games defined by Shang et al. (2005) refer to educational websites or software that employed game features.
- **Mainstream games.** Widely known as computer games, video games, or mainstream commercial games. Mainstream games are developed "solely for fun for the user and to maximize profit for the publisher" (Kirriemuir & McFarlane, 2004, p. 20).
- **Scale.** "A scale is a composite measure of several items (questions/statements) that have a logical or empirical structure among them. This allows us to measure the direction and intensity of a construct" (Robertson, 2017, p. 6).
- **Serious game.** A serious game is a video game in which "the primary goal is to educate, rather than just entertain players" (Fitz-Walter, 2015, p. 26).
- **Simulation.** Also named simulation game. A simulation game is a game, which "has elements like score, performance rating, conflict, and payoff, and simulates a real world situation for decision-making or alternative evaluation" (Deshpande & Huang, 2011, p. 400).

REFERENCES

REFERENCES

Abdulghani, H. M., Al-Drees, A. A., Khalil, M. S., Ahmad, F., Ponnamperuma, G. G., & Amin, Z. (2014). What factors determine academic achievement in high achieving undergraduate medical students? A qualitative study. *Medical teacher*, *36*(sup1), S43-S48.

Abrams, S. S., & Walsh, S. (2014). Gamified vocabulary. *Journal of Adolescent & Adult Literacy*, *58*(1), 49-58.

Adukaite, A., van Zyl, I., Er, Ş., & Cantoni, L. (2017). Teacher perceptions on the use of digital gamified learning in tourism education: The case of South African secondary schools. *Computers & Education*, *111*, 172-190.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Al Kurdi, B., Alshurideh, M., Salloum, S., Obeidat, Z., & Al-dweeri, R. (2020). An empirical investigation into examination of factors influencing university students' behavior towards elearning acceptance using SEM approach.

Albirini, A. (2006). Teachers' attitudes toward information and communication technologies: The case of Syrian EFL teachers. *Computers & Education*, *47*(4), 373-398.

Ale, K., & Chib, A. (2011). Community factors in technology adoption in primary education: Perspectives from rural India. *Information Technologies & International Development*, *7*(4), pp. 53-68.

Aljukhadar, M., & Senecal, S. (2015). Determinants of an organization's website ease of use: The moderating role of product tangibility. *Journal of Organizational Computing and Electronic Commerce*, *25*(4), 337-359.

Alsabawy, A. Y., Cater-Steel, A., & Soar, J. (2016). Determinants of perceived usefulness of e-learning systems. *Computers in human behavior*, *64*, 843-858.

Anderson, L. W., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives* (Abridg ed.). New York: Longman.

Angelia, F. (2019). *Improving English Learning through Game Using 6–11 MDA Framework.* Paper presented at the 2019 12th International Conference on Information & Communication Technology and System (ICTS).

Anthoine, E., Moret, L., Regnault, A., Sébille, V., & Hardouin, J.-B. (2014). Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. *Health and quality of life outcomes*, *12*(1), 1-10.

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*(5), 369-386.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of school psychology*, *44*(5), 427-445.

Arakawa, Y., & Matsuda, Y. (2016). Gamification mechanism for enhancing a participatory urban sensing: Survey and practical results. *Journal of Information Processing*, *24*(1), 31-38.

Aziz, M., & Macredie, R. D. (2005). *Proposing a perceived ease of use factors taxonomy for information system use.* Paper presented at the Proceedings. IEEE SoutheastCon, 2005.

Baaren, E., Van De Wijngaert, L., & Huizer, E. (2008). *'I want my HDTV'? Underlying factors of perceived usefulness for high definition television.* Paper presented at the European Conference on Interactive Television.

Baek, Y. K. (2008). What hinders teachers in using computer and video games in the classroom? Exploring factors inhibiting the uptake of computer and video games. *CyberPsychology & Behavior*, *11*(6), 665-671.

Baggaley, A. R. (1983). Deciding on the ratio of number of subjects to number of variables in factor analysis. *Multivariate Experimental Clinical Research*.

Baldauf, M., Brandner, A., & Wimmer, C. (2017). *Mobile and gamified blended learning for language teaching: studying requirements and acceptance by students, parents and teachers in the wild.* Paper presented at the Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia.

Ball, J., Ogletree, R., Asunda, P., Miller, K., & Jurkowski, E. (2014). Diffusion of innovation elements that influence the adoption and diffusion of distance education in health. *American journal of health studies*, *29*(3), 240-246. Retrieved from https://go.exlibris.link/TcT314cF

Bandura, A. (1997). The anatomy of stages of change. *American journal of health promotion: AJHP*, *12*(1), 8.

Baptista, G., & Oliveira, T. (2019). Gamification and serious games: A literature meta-analysis and integrative model. *Computers in human behavior*, *92*, 306-315.

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. *Computers in human behavior*, *71*, 550-585.

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child development*, 887-907.

Baydas, O., & Cicek, M. (2019). The examination of the gamification process in undergraduate education: A scale development study. *Technology, Pedagogy and Education*, *28*(3), 269-285.

Becker, K. (2007). Digital game ‐ based learning once removed: Teaching teachers. *British journal of educational technology*, *38*(3), 478-488.

Beggs, T. A. (2000). *Influences and Barriers to the Adoption of Instructional Technology.* Paper presented at the Mid-South Instructional Technology Conference 2000.

Berg Marklund, B. (2015). *Unpacking Digital Game-Based Learning: The complexities of developing and using educational games.* University of Skövde,

Birks, M., & Mills, J. (2015). *Grounded theory: A practical guide*: Sage.

Blohm, I., & Leimeister, J. (2013). Gamification: Design of IT-based enhancing services for motivational support and behavioral change. Business and Information Systems Engineering, 5 (4), 275–278. In.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher*, *13*(6), 4-16.

Bourgonjon, J., De Grove, F., De Smet, C., Van Looy, J., Soetaert, R., & Valcke, M. (2013). Acceptance of game-based learning by secondary school teachers. *Computers & Education*, *67*, 21-35.

Bourgonjon, J., Valcke, M., Soetaert, R., & Schellens, T. (2009). *Exploring the acceptance of video games in the classroom by secondary school students.* Paper presented at the 17th International Conference on Computers in Education.

Bourgonjon, J., Valcke, M., Soetaert, R., & Schellens, T. (2010). Students' perceptions about the use of video games in the classroom. *Computers & Education*, *54*(4), 1145-1156.

Boykin, A. W., & Noguera, P. (2011). *Creating the opportunity to learn: Moving from research to practice to close the achievement gap*: Ascd.

Brandyberry, A. A., Li, X., & Lin, L. (2010, August 12-15, 2010). *Determinants of Perceived Usefulness and Perceived Ease of Use in Individual Adoption of Social Network Sites.* Paper presented at the Americas Conference on Information System (AMCIS), Lima, Peru.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, *45*(4), 624-634.

Brown, I. T. (2002). Individual and technological factors affecting perceived ease of use of web ‐ based learning technologies in a developing country. *The Electronic Journal of Information Systems in Developing Countries*, *9*(1), 1-15.

Buckley, P., & Doyle, E. (2017). Individualising gamification: An investigation of the impact of learning styles and personality traits on the efficacy of gamification using a prediction market. *Computers & Education*, *106*, 43-55.

Bunchball. (2010). *Gamification 101: An introduction to the use of game dynamics to influence behavior*. Retrieved from https://australiandirectmarketingassociation.files.wordpress.com/2011/10/gamification101.pdf

Cain, J., & Piascik, P. (2015). Are serious games a good strategy for pharmacy education? *American journal of pharmaceutical education*, *79*(4), 47.

Carlson, K. D., & Herdman, A. O. (2012). Understanding the Impact of Convergent Validity on Research Results. *Organizational Research Methods*, *15*(1), 17-32.

Carr, J. (2011). *Educational games to engage the un-engagable.* (PhD thesis). University of Nottingham,

Cavanagh, K. V. (2019). *Gamification as a Vehicle to Encourage Behavior Change.* (PhD Thesis). Case Western Reserve University,

Celik, I., Sahin, I., & Aydin, M. (2014). Reliability and Validity Study of the Mobile Learning Adoption Scale Developed Based on the Diffusion of Innovations Theory. *International Journal of Education in Mathematics Science and Technology (IJEMST)*, *2*(4), 300-316.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of research in personality*, *37*(4), 319-338.

Chapman, J. R., & Rich, P. J. (2018). Does educational gamification improve students' motivation? If so, which game elements work best? *Journal of Education for Business*, *93*(7), 315-322.

Chen, C.-M., Li, M.-C., & Chen, T.-C. (2018). *A collaborative reading annotation system with gamification mechanisms to improve reading performance.* Paper presented at the 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI).

Chen, C.-M., Li, M.-C., & Chen, T.-C. (2020). A web-based collaborative reading annotation system with gamification mechanisms to improve reading performance. *Computers & Education*, *144*, 103697.

Chen, Y.-L., & Hsu, C.-C. (2020). Self-regulated mobile game-based English learning in a virtual reality environment. *Computers & Education*, *154*, 103910.

Chen, Y., Burton, T., Mihaela, V., & Whittinghill, D. (2015). Cogent: a case study of meaningful gamification in education with virtual currency. *iJET–International Journal of Emerging Technologies in Learning*, *10*, 133-147.

Cheng, Y.-M., Lou, S.-J., Kuo, S.-H., & Shih, R.-C. (2013). Investigating elementary school students' technology acceptance by applying digital game-based learning to environmental education. *Australasian Journal of Educational Technology*, *29*(1), 96-110.

Chou, Y.-k. (2015). *Actionable gamification: Beyond points, badges, and leaderboards*: Octalysis Group.

Chou, Y.-k. (2019). Actionable gamification: Beyond points, badges, and leaderboards. In: Packt Publishing Ltd.

Churchill Jr, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of marketing research, 16*(1), 64-73.

Cigdem, H., & Ozturk, M. (2016). Factors affecting students' behavioral intention to use LMS at a Turkish post-secondary vocational school. *International Review of Research in Open and Distributed Learning, 17*(3), 276-295.

Clements, A. J., Ahmed, S., & Henderson, B. (2017). *Student Experience of Gamified Learning: A Qualitative Approach.* Paper presented at the European Conference on Games Based Learning.

Colman, A. M., Norris, C. E., & Preston, C. C. (1997). Comparing rating scales of different lengths: Equivalence of scores from 5-point and 7-point scales. *Psychological Reports, 80*(2), 355-362.

Comrey, A. L. (1973). *A first course in factor analysis*: New York: Academic Press.

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of consulting and clinical psychology, 56*(5), 754.

Conklin, J. (2005). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives complete edition. In: Sage Publications, Ltd., Phi Delta Kappa International.

Considine, G., & Zappalà, G. (2002). The influence of social and economic disadvantage in the academic performance of school students in Australia. *Journal of sociology, 38*(2), 129-148.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation, 10*(1), 1-9.

Cox III, E. P. (1980). The optimal number of response alternatives for a scale: A review. *Journal of marketing research, 17*(4), 407-422.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*: Sage publications.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into practice, 39*(3), 124-130.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*: Sage publications.

Crow, T., & Parsons, D. (2015). *A mobile game world for Māori language learning.* Paper presented at the International Conference on Mobile and Contextual Learning.

Cruaud, C. (2018). The playful frame: gamification in a French-as-a-foreign-language class. *Innovation in Language Learning and Teaching, 12*(4), 330-343.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience* (Vol. 1990): Harper & Row New York.

Csikzentimihalyi, M. (1975). *Beyond boredom and anxiety: Experiencing flow in work and play*: Jossey-Bass.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*: Teachers College Press.

da Silva Brito, R., Contreras Pinochet, L. H., Luiz Lopes, E., & de Oliveira, M. A. (2018). Development of a gamification characteristics measurement scale for mobile application users. *Internext: Revista Electrônica de Negócios Internacionais da ESPM, 13*(1), 1-16.

Dalal, D. K., & Carter, N. T. (2014). Negatively worded items negatively impact survey research. In *More Statistical and Methodological Myths and Urban Legends: Doctrine, Verity and Fable in Organizational and Social Sciences* (pp. 112-132): New York, NY: Routledge.

Davidson, S. (2018). *A Multi-dimensional model of enjoyment: Development and validation of an enjoyment scale (ENJOY).* (Doctor of Philosophy). Embry-Riddle Aeronautical University,

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management science, 35*(8), 982-1003.

Dawes, J. (2002). Five point vs. eleven point scales: Does it make a difference to data characteristics. *Australasian Journal of Market Research, 10*(1), 39-47.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International journal of market research, 50*(1), 61-104.

De-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education, 75*, 82-91.

de Freitas, S., & Liarokapis, F. (2011). Serious games: A new paradigm for education? In M. Ma, Oikonomou, A. and Jain, L.C. (Ed.), *Serious Games and Edutainment Applications* (pp. 9-23): Springer-Verlag, London, UK.

De Grove, F., Bourgonjon, J., & Van Looy, J. (2012). Digital games in the classroom? A contextual approach to teachers' adoption intention of digital games in formal education. *Computers in human behavior, 28*(6), 2023-2033.

de los Arcos, B., Faems, B., Comas-Quinn, A., & Pulker, H. (2017). Teachers' use and acceptance of gamification and social networking features of an open repository. *European Journal of Open, Distance and E-learning, 20*(1), 127-138.

Dearing, J. W. (2009). Applying Diffusion of Innovation Theory to Intervention Development. *Research on social work practice, 19*(5), 503-518. doi:10.1177/1049731509335569

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry, 11*(4), 227-268.

Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaee, E., & Noroozi, O. (2019). Using gamification to support learning English as a second language: a systematic review. *Computer Assisted Language Learning*, 1-24.

Demirbilek, M., & Tamer, S. L. (2010). Math teachers' perspectives on using educational computer games in math education. *Procedia-Social and Behavioral Sciences, 9*, 709-716.

Denden, M., Tlili, A., Essalmi, F., & Jemni, M. (2017). *An investigation of the factors affecting the perception of gamification and game elements.* Paper presented at the 2017 6th International Conference on Information and Communication Technology and Accessibility (ICTA).

Deng, Z., Hong, Z., Ren, C., Zhang, W., & Xiang, F. (2018). What predicts patients' adoption intention toward mHealth services in China: empirical study. *JMIR mHealth and uHealth, 6*(8), e172.

Denny, P. (2013). *The effect of virtual achievements on student engagement.* Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems.

Deshpande, A. A., & Huang, S. H. (2011). Simulation games in engineering education: A state-of-the-art review. *Computer applications in engineering education, 19*(3), 399-410. doi:10.1002/cae.20323

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). *From game design elements to gamefulness: defining gamification.* Paper presented at the Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). *Gamification. using game-design elements in non-gaming contexts.* Paper presented at the CHI'11 extended abstracts on human factors in computing systems.

DeVellis, R. (2003). *Scale development: Theory and applications*: Thousand Oaks, CA: Sage Publication.

Dick, W., & Hagerty, N. (1971). Topics in Measurement: Reliability and Validity.

Dindar, M., Ren, L., & Järvenoja, H. (2020). An experimental study on the effects of gamified cooperation and competition on English vocabulary learning. *British journal of educational technology, 52*(1), 142-159.

DomíNguez, A., Saenz-De-Navarrete, J., De-Marcos, L., FernáNdez-Sanz, L., PagéS, C., & MartíNez-HerráIz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*, 380-392.

Emel'yanenko, V. D., Vetoshko, A. N., Malinnikov, S. G., Malashenko, I. V., & Vetoshko, L. I. (2016). Man's Values and Ideologies as a Basis of Gamification. *International Journal of Environmental and Science Education, 11*(18), 12576-12592.

Eppmann, R., Bekk, M., & Klein, K. (2018). Gameful Experience in Gamification: construction and validation of a Gameful Experience Scale [GAMEX]. *Journal of Interactive Marketing, 43*, 98-115.

Everitt, B. S. (1975). Multivariate analysis: The need for data, and other problems. *The British Journal of Psychiatry, 126*(3), 237-240.

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*: Oxford University Press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods, 4*(3), 272-299.

Fenollar, P., Román, S., & Cuestas, P. J. (2007). University students' academic performance: An integrative conceptual framework and empirical analysis. *British Journal of Educational Psychology, 77*(4), 873-891.

Fernandes, J., Duarte, D., Ribeiro, C., Farinha, C., Pereira, J. M., & da Silva, M. M. (2012). iThink: A game-based approach towards improving collaboration and participation in requirement elicitation. *Procedia Computer Science, 15*, 66-77.

Fink, A., & Litwin, M. S. (1995). *How to measure survey reliability and validity* (Vol. 7): Sage.

Finstad, K. (2010). Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of Usability Studies, 5*(3), 104-110.

Fisbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory and research. *Massachusetts, Addison-Wiley Publishing Company*.

Fitz-Walter, Z. J. (2015). *Achievement unlocked: Investigating the design of effective gamification experiences for mobile applications and devices.* (Doctoral thesis). Queensland University of Technology,

Fleming, N. D. (2001). *Teaching and learning styles: VARK strategies*: IGI global.

Flick, U. (2018). *An introduction to qualitative research*: sage.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research, 18*(1), 39-50.

Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education, 52*(1), 101-112.

Gafni, R., Achituv, D. B., & Rahmani, G. (2017). Learning foreign languages using mobile applications. *Journal of Information Technology Education: Research, 16*, 301-317.

Ge, X., & Ifenthaler, D. (2018). Designing engaging educational games and assessing engagement in game-based learning. In *Gamification in Education: Breakthroughs in Research and Practice* (pp. 1-19): IGI Global.

González, C. S., Gómez, N., Navarro, V., Cairós, M., Quirce, C., Toledo, P., & Marrero-Gordillo, N. (2016). Learning healthy lifestyles through active videogames, motor games and the gamification of educational activities. *Computers in human behavior, 55*, 529-551.

Gorsuch, R. L. (1983). *Factor Analysis* (2 ed.): Hillsdale, NJ: Erlbaum.

Graafland, M., Schraagen, J. M., & Schijven, M. P. (2012). Systematic review of serious games for medical education and surgical skills training. *British journal of surgery, 99*(10), 1322-1330.

Groh, F. (2012). Gamification: State of the art definition and utilization. *RTMI*, 39-46.

Haghighat, R. (2007). The development of the brief social desirability scale (BSDS). *Europe's Journal of Psychology, 3*(4). doi:https://doi.org/10.5964/ejop.v3i4.417

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis* (Vol. 5): Prentice hall Upper Saddle River, NJ.

Hakulinen, L., Auvinen, T., & Korhonen, A. (2013). *Empirical study on the effect of achievement badges in TRAKLA2 online learning environment.* Paper presented at the 2013 Learning and teaching in computing and engineering.

Ham, M., Jeger, M., & Frajman Ivković, A. (2015). The role of subjective norms in forming the intention to purchase green food. *Economic research-Ekonomska istraživanja, 28*(1), 738-748.

Hamari, J. (2013). Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic commerce research and applications, 12*(4), 236-245.

Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in human behavior, 71*, 469-478.

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional flow scale-2. *Computers in human behavior, 40*, 133-143.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education, 80*, 152-161.

Hashim, K. F., Tan, F. B., & Rashid, A. (2015). Adult learners' intention to adopt mobile learning: A motivational perspective. *British journal of educational technology, 46*(2), 381-390.

Hattie, J., & Yates, G. C. (2013). *Visible learning and the science of how we learn*: Routledge.

He, J., Van de Vijver, F. J., Dominguez Espinosa, A., Abubakar, A., Dimitrova, R., Adams, B. G., . . . Bobowik, M. (2015). Socially desirable responding: Enhancement and denial in 20 countries. *Cross-Cultural Research, 49*(3), 227-249.

Heryadi, Y., & Muliamin, K. (2016). *Gamification of M-learning Mandarin as second language.* Paper presented at the 2016 1st International Conference on Game, Game Art, and Gamification (ICGGAG).

Herzig, P., Strahringer, S., & Ameling, M. (2012). *Gamification of ERP systems-Exploring gamification effects on user acceptance constructs.* Paper presented at the Multikonferenz Wirtschaftsinformatik.

Hew, K. F., Huang, B., Chu, K. W. S., & Chiu, D. K. (2016). Engaging Asian students through game mechanics: Findings from two experiment studies. *Computers & Education, 92*, 221-236.

Higgins, E. T. (2006). Value from hedonic experience and engagement. *Psychological review, 113*(3), 439-460.

Hinkin, T. R., Tracey, J. B., & Enz, C. A. (1997). Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality & Tourism Research, 21*(1), 100-120.

Holsapple, C. W., & Wu, J. (2007). User acceptance of virtual worlds: The hedonic framework. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 38*(4), 86-89.

Holton, J. A., & Walsh, I. (2016). *Classic grounded theory: Applications with qualitative and quantitative data*: Sage Publications.

Holton, J. A., & Walsh, I. (2017). Foundational pillars of classic grounded theory. *Classic Grounded Theory: Applications with qualitative and quantitative data. Thousand Oaks, CA: Sage Publications*.

Homer, R., Hew, K. F., & Tan, C. Y. (2018). Comparing digital badges-and-points with classroom token systems: Effects on elementary school ESL students' classroom behavior and English learning. *Journal of Educational Technology & Society, 21*(1), 137-151.

Hong, J.-C., Hwang, M.-Y., Liu, Y.-H., & Tai, K.-H. (2020). Effects of gamifying questions on English grammar learning mediated by epistemic curiosity and language anxiety. *Computer Assisted Language Learning*, 1-25.

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63-84.

Honório, J., Moura, J. A. B., Brito, P., Menezes, T., & de Barros, M. A. (2018). *Gamification in Location-based M-Learning: Students' Perceptions of Game Elements.* Paper presented at the 10th International Conference on Computer Supported Education.

Hu, J. H., Clark, T., & Ma, W. W. (2003). Examining technology acceptance by school teachers: a longitudinal study. *Information & Management, 41*(2), 227-241. doi:https://doi.org/10.1016/S0378-7206(03)00050-8

Huang, H.-M., & Liaw, S.-S. (2018). An analysis of learners' intentions toward virtual reality learning based on constructivist and technology acceptance approaches. *International Review of Research in Open and Distributed Learning, 19*(1), 1-25.

Huizenga, J., Ten Dam, G., Voogt, J., & Admiraal, W. (2017). Teacher perceptions of the value of game-based learning in secondary education. *Computers & Education, 110*, 105-115.

Hunicke, R., LeBlanc, M., & Zubek, R. (2004). *MDA: A formal approach to game design and game research.* Paper presented at the Proceedings of the AAAI Workshop on Challenges in Game AI.

Hwang, G.-J., Sung, H.-Y., Hung, C.-M., Huang, I., & Tsai, C.-C. (2012). Development of a personalized educational computer game based on students' learning styles. *Educational Technology Research and Development, 60*(4), 623-638.

Ibanez, M.-B., Di-Serio, A., & Delgado-Kloos, C. (2014). Gamification for engaging computer science students in learning activities: A case study. *IEEE Transactions on learning technologies, 7*(3), 291-301.

Ibrahim, R., & Jaafar, A. (2011). User acceptance of educational games: A revised unified theory of acceptance and use of technology (UTAUT). *World Academy of Science, Engineering and Technology, 77*, 551-557.

IJsselsteijn, Kort, d., & Poels. (n.d.). The game experience questionnaire. Retrieved from https://research.tue.nl/en/publications/the-game-experience-questionnaire

Ismail, R., & Ibrahim, R. (2018, 2018). *Teachers' Perception on Digital Game: A Preliminary Investigation towards Educational Game Application for Islamic Religious Primary Schools*.

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of sport and exercise psychology, 18*(1), 17-35.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies, 66*(9), 641-661.

Jiang, Y., Chen, D., & Lai, F. (2010). Technological-Personal-Environmental (TPE) Framework: A Conceptual Model for Technology Acceptance at the Individual Level. *Journal of International Technology and Information Management, 19*(3), 5. Retrieved from https://scholarworks.lib.csusb.edu/jitim/vol19/iss3/5

Johnson, R. B. (1997). Examining the validity structure of qualitative research. *Education, 118*(2), 282.

Kapp, K. M. (2012). *The gamification of learning and instruction*: Wiley San Francisco.

Ketelhut, D. J., & Schifter, C. C. (2011). Teachers and game-based learning: Improving understanding of how to increase efficacy of adoption. *Computers & Education, 56*(2), 539-546.

Khlaif, Z. (2018). Teachers' perceptions of factors affecting their adoption and acceptance of mobile technology in K-12 settings. *Computers in the Schools, 35*(1), 49-67.

Kim, J. T., & Lee, W.-H. (2015). Dynamical model for gamification of learning (DMGL). *Multimedia Tools and Applications, 74*(19), 8483-8493.

Kirkley, S. E., & Kirkley, J. R. (2005). Creating next generation blended learning environments using mixed reality, video games and simulations. *TechTrends, 49*(3), 42-53.

Kirriemuir, J., & McFarlane, A. (2004). Literature review in games and learning.

Koh, E., Kin, Y. G., Wadhwa, B., & Lim, J. (2012). Teacher perceptions of games in Singapore schools. *Simulation & Gaming, 43*(1), 51-66.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in human behavior, 35*, 179-188.

Kurniawan, H. C., Sitohang, B., & Rukmono, S. A. (2019). *Gamification of Mobile-based Japanese Language Shadowing.* Paper presented at the 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT).

Kurt, S. (2012). How do teachers prioritize the adoption of technology in the classroom? *Teachers and Teaching, 18*(2), 217-231.

Lam, Y. W., Hew, K., & Chiu, K. F. (2018). Improving argumentative writing: Effects of a blended learning approach and gamification. *Language learning & technology, 22*(1), 97-118.

Lam, Y. W., Hew, K. F., & Chiu, K. F. (2018). Improving argumentative writing: Effects of a blended learning approach and gamification. *Language learning & technology, 22*(1), 97-118.

Lameras, P., Arnab, S., Dunwell, I., Stewart, C., Clarke, S., & Petridis, P. (2017). Essential features of serious games design in higher education: Linking learning attributes to game mechanics. *British journal of educational technology, 48*(4), 972-994. doi:10.1111/bjet.12467

Landers, R. N., & Armstrong, M. B. (2017). Enhancing instructional outcomes with gamification: An empirical test of the Technology-Enhanced Training Effectiveness Model. *Computers in human behavior, 71*, 499-507.

Lee, & Hammer. (2011). Gamification in Education: What, How, Why Bother. *Academic Exchange Quarterly, 15*(2), 1-5.

Lee, M.-C. (2010). Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation–confirmation model. *Computers & Education, 54*(2), 506-516.

Lee, Y.-H., Hsieh, Y.-C., & Hsu, C.-N. (2011). Adding Innovation Diffusion Theory to the Technology Acceptance Model: Supporting Employees' Intentions to use E-Learning Systems. *Educational technology & society, 14*(4), 124-137. Retrieved from https://go.exlibris.link/Q5y62R9K

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the" overjustification" hypothesis. *Journal of Personality and social Psychology, 28*(1), 129.

Lewis, J. R. (1993). Multipoint scales: Mean and median differences and observed significance levels. *International Journal of Human‐Computer Interaction, 5*(4), 383-392.

Licorish, S. A., Owen, H. E., Daniel, B., & George, J. L. (2018). Students' perception of Kahoot!'s influence on teaching and learning. *Research and Practice in Technology Enhanced Learning, 13*(1), 1-23.

Lieberoth, A. (2015). Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture, 10*(3), 229-248.
Lijuan Wang, C., & Ha, A. (2009). Pre-service teachers' perception of Teaching Games for Understanding: A Hong Kong perspective. *European Physical Education Review, 15*(3), 407-429.
Lim, C. P. (2008). Global citizenship education, school curriculum and games: Learning Mathematics, English and Science as a global citizen. *Computers & Education, 51*(3), 1073-1093.
Lin, M.-H., & Chen, S.-H. (2013). *A case study of senior user acceptance for Orange Technology game.* Paper presented at the 2013 1st International Conference on Orange Technologies (ICOT).
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*: Newbury Park, CA: Sage.
Linehan, C., Kirman, B., Lawson, S., & Chan, G. (2011). *Practical, appropriate, empirically-validated guidelines for designing educational games.* Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems.
Lister, M. (2015). Gamification: The effect on student motivation and performance at the post-secondary level. *Issues and Trends in Educational Technology, 3*(2).
Liu, C.-R., Wang, Y.-C., Huang, W.-S., & Tang, W.-C. (2019). Festival gamification: Conceptualization and scale development. *Tourism Management, 74*, 370-381.
Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL: the Journal of EUROCALL, 31*(3), 293-311.
Lopez, C. E., & Tucker, C. S. (2019). The effects of player type on performance: A gamification case study. *Computers in human behavior, 91*, 333-345.
Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of computer assisted learning, 24*(6), 515-525.
Luo, Z. (2021). Gamification for educational purposes: What are the factors contributing to varied effectiveness? *Education and Information Technologies*, 1-25. doi:https://doi.org/10.1007/s10639-021-10642-9
Luo, Z., Brown, C., & O'Steen, B. (2021). Factors contributing to teachers' acceptance intention of gamified learning tools in secondary schools: An exploratory study. *Education and Information Technologies, 26*, 6337-6363. doi:https://doi.org/10.1007/s10639-021-10622-z
Lyons, E. E., & Coyle, A. E. (2007). *Analysing qualitative data in psychology*: Sage Publications Ltd.
MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods, 4*(1), 84-99.
Malaquias, R. F., Malaquias, F. F., & Hwang, Y. (2018). Understanding technology acceptance features in learning through a serious game. *Computers in human behavior, 87*, 395-402.
Marczewski, A. (2015). Game-based solution design: Differences between gamification, simulations, serious games and games. Retrieved from https://www.gamified.uk/gamification-framework/differences-between-gamification-and-games/
Marino, M. T., Israel, M., Beecher, C. C., & Basham, J. D. (2013). Students' and teachers' perceptions of using video games to enhance science instruction. *Journal of Science Education and Technology, 22*(5), 667-680.
Martí-Parreño, J., Méndez-Ibáñez, E., & Alonso-Arroyo, A. (2016). The use of gamification in education: a bibliometric and text mining analysis. *Journal of computer assisted learning, 32*(6), 663-676.
Martí-Parreño, J., Seguí-Mas, D., & Seguí-Mas, E. (2016). Teachers' attitude towards and actual use of gamification. *Procedia-Social and Behavioral Sciences, 228*, 682-688.
Maydeu-Olivares, A., Fairchild, A. J., & Hall, A. G. (2017). Goodness of Fit in Item Factor Analysis: Effect of the Number of Response Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(4), 495-505. doi:10.1080/10705511.2017.1289816
McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games*: TEEM (Teachers evaluating educational multimedia), Cambridge.
McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*: Penguin.
McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica, 22*(3), 276-282.
McKenzie, K., & Schweitzer, R. (2001). Who succeeds at university? Factors predicting academic performance in first year Australian university students. *Higher education research & development, 20*(1), 21-33.
Mei, Y. C., & Aun, N. B. (2019). Factors Influencing Consumers' Perceived Usefulness of M-Wallet in Klang Valley, Malaysia. *Review of Integrative Business and Economics Research, 8*, 1-23.
Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in human behavior, 71*, 525-534.
Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science, 6*(1), 42.
Mihas, P. (2019). *Learn to Use an Exploratory Sequential Mixed Method Design for Instrument Development*: SAGE Publications, Limited.
Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information systems research, 2*(3), 192-222.

Moral-Bofill, L., Lópezdelallave, A., Pérez-Llantada, M. C., & Holgado-Tello, F. P. (2020). Adaptation to Spanish and psychometric study of the Flow State Scale-2 in the field of musical performers. *PloS one, 15*(4), e0231054.

Morris, B., Croker, S., Zimmerman, C., Gill, D., & Romig, C. (2013). Gaming science: the "Gamification" of scientific thinking. *Frontiers in psychology, 4*, 607.

Morris, B. J., Dragovich, C., Todaro, R., Balci, S., & Dalton, E. (2019). Comparing badges and learning goals in low-and high-stakes learning contexts. *Journal of Computing in Higher Education, 31*(3), 573-603.

Murad, D., Wang, R., Turnbull, D., & Wang, Y. (2018). *SLIONS: A karaoke application to enhance foreign language learning.* Paper presented at the Proceedings of the 26th ACM international conference on Multimedia.

Nacke, L. E., & Deterding, S. (2017). The maturing of gamification research. *Computers in human behavior, 71*, 450-454. doi:10.1016/j.chb.2016.11.062

NFER. (2009). NFER Teacher Voice Omnibus February 2009 Survey- Using computer games in the classroom Futurelab. Retrieved from https://www.nfer.ac.uk/media/2752/futl27teachervoicereport.pdf

Nicolle, P. S., & Lou, Y. (2008). Technology adoption into teaching and learning by mainstream university faculty: A mixed methodology study revealing the "how, when, why, and why not". *Journal of Educational Computing Research, 39*(3), 235-265.

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence-based nursing, 18*(2), 34-35.

Noraddin, E., & Kian, N. T. (2015). Three learning potentials in digital games: perception of Malaysian university teachers. *Journal of E-learning and Knowledge Society, 11*(2).

Nunnally, J. C. (1978). Psychometric Theory 2nd ed. In: Mcgraw hill book company.

Okioga, C. K. (2013). The impact of students' socio-economic background on academic performance in Universities, a case of students in Kisii University College. *American International Journal of Social Science, 2*(2), 38-46.

Osipov, I. V., Volinsky, A. A., Nikulchev, E., & Prasikova, A. Y. (2016). Communication and gamification in the web-based foreign language educational system: Web-based foreign language educational system. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 11*(4), 22-34.

Pallant, J. (2016). *SPSS survival manual: a step by step guide to data analysis using IBM SPSS* (6th ed.). Maidenhead, Berkshire, England: McGraw-Hill Education.

Palomo-Duarte, M., Berns, A., Dodero, J. M., & Cejas, A. (2014). *Foreign language learning using a gamificated APP to support peer-assessment.* Paper presented at the Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality.

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*(1), 1-12.

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Journal of Educational Technology & Society, 12*(3), 150-162.

Park, S. Y., Nam, M. W., & Cha, S. B. (2012). University students' behavioral intention to use mobile learning: Evaluating the technology acceptance model. *British journal of educational technology, 43*(4), 592-605.

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health services research, 34*(5 Pt 2), 1189-1208.

Pedersen, M. K., Rasmussen, N. R., Sherson, J. F., & Basaiawmoit, R. V. (2017). Leaderboard effects on player performance in a citizen science game. *arXiv preprint arXiv:1707.03704*.

Peng, W., Song, H., Kim, J., & Day, T. (2016). The influence of task demand and social categorization diversity on performance and enjoyment in a language learning game. *Computers & Education, 95*, 285-295.

Phan, M. H., Keebler, J. R., & Chaparro, B. S. (2016). The development and validation of the game user experience satisfaction scale (GUESS). *Human factors, 58*(8), 1217-1247.

Platonova, V., & Bērziša, S. (2017). Gamification in Software Development Projects. *Information Technology and Management Science, 20*(1), 58-63.

Prensky, M. (2001). Digital game-based learning, McGraw-Hill & Paragon House, New York.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psychologica, 104*(1), 1-15.

Proctor, M. D., & Marks, Y. (2013). A survey of exemplar teachers' perceptions, use, and access of computer-based games and technology for classroom instruction. *Computers & Education, 62*, 171-180.

Purgina, M., Mozgovoy, M., & Blake, J. (2020). WordBricks: Mobile technology and visual grammar formalism for gamification of natural language grammar acquisition. *Journal of Educational Computing Research, 58*(1), 126-159.

Quazi, A., & Talukder, M. (2011). Demographic Determinants of Adoption of Technological Innovation. *The Journal of computer information systems, 52*(1), 34-42. doi:10.1080/08874417.2011.11645520

Rachels, J. R., & Rockinson-Szapkiw, A. J. (2018). The effects of a mobile gamification app on elementary students' Spanish achievement and self-efficacy. *Computer Assisted Language Learning, 31*(1-2), 72-89.

Rahmi, B., Birgoren, B., & Aktepe, A. (2018). A meta analysis of factors affecting perceived usefulness and perceived ease of use in the adoption of e-learning systems. *Turkish Online Journal of Distance Education, 19*(4), 4-42.

Rai, V., & Beck, A. L. (2017). Play and learn: serious games in breaking informational barriers in residential solar energy adoption in the United States. *Energy research & social science*, *27*, 70-77.

Ranellucci, J., Rosenberg, J. M., & Poitras, E. G. (2020). Exploring pre‐service teachers' use of technology: The technology acceptance model and expectancy–value theory. *Journal of computer assisted learning*, *36*(6), 810-824.

Ranz, M. (2015). The impact of gamification on intrinsic motivation: An experimental study of administrative tasks. Retrieved from https://www.mysciencework.com/publication/show/impact-gamification-intrinsic-motivation-experimental-study-administrative-tasks-a8e6728c

Rapp, A. (2015). A qualitative investigation of gamification: Motivational factors in online gamified services and applications. In *Gamification: Concepts, Methodologies, Tools, and Applications* (pp. 32-48): IGI Global.

Raubenheimer, J. (2004). An item selection procedure to maximize scale reliability and validity. *SA Journal of Industrial Psychology*, *30*(4), 59-64.

Raza, S. A., Umer, A., & Shah, N. (2017). New determinants of ease of use and perceived usefulness for mobile banking adoption. *International Journal of Electronic Customer Relationship Management*, *11*(1), 44-65.

Reeves, B., & Read, J. L. (2009). *Total engagement: How games and virtual worlds are changing the way people work and businesses compete*: Harvard Business Press.

Rego, I. (2015). Mobile Language Learning: How Gamification Improves the Experience. In Y. A. Zhang (Ed.), *Handbook of mobile teaching and learning* (pp. 705-720).

Reinicke, B. A., & Marakas, G. M. (2005). *Exploring the psychological determinants of perceived ease of use and usefulness.* Paper presented at the Proceedings of the 38th Annual Hawaii International Conference on System Sciences.

Revythi, A., & Tselios, N. (2019). Extension of technology acceptance model by using system usability scale to assess behavioral intention to use e-learning. *Education and Information Technologies*, *24*(4), 2341-2355.

Reynolds, E. D., & Taylor, B. (2020). Kahoot!: EFL instructors' implementation experiences and impacts on students' vocabulary knowledge. In: CALL-EJ.

Riva, S., Camerini, A.-L., Allam, A., & Schulz, P. J. (2014). Interactive sections of an Internet-based intervention increase empowerment of chronic back pain patients: randomized controlled trial. *Journal of medical Internet research*, *16*(8), e180.

Robertson, G. (2017). *Developing valid and reliable survey scales*. Retrieved from https://i2ifacility.org/system/documents/files/000/000/010/original/Developing_valid_and_reliable_survey_scales_i2i_October_2017.pdf?1507625556

Rodgers, K. B., & Rose, H. A. (2001). Personal, family, and school factors related to adolescent academic performance: A comparison by family structure. *Marriage & family review*, *33*(4), 47-61.

Rogers, E. M. (2003). *Diffusion of Innovations* (5 ed.): New York, NY: Free Press, Simon & Schuster.

Rogers, E. M. (2010). *Diffusion of innovations*: Simon and Schuster.

Rotgans, J. I., & Schmidt, H. G. (2011). Cognitive engagement in the problem-based learning classroom. *Advances in health sciences education*, *16*(4), 465-479.

Roy, R. V., & Zaman, B. (2017). *Why Gamification Fails in Education and How to Make It Successful: Introducing Nine Gamification Heuristics Based on Self-Determination Theory*: Serious Games and Edutainment Applications.

Rueckert, D., Pico, K., Kim, D., & Calero Sánchez, X. (2020). Gamifying the foreign language classroom for brain‐friendly learning. *Foreign Language Annals*.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

Sahin, I. (2006). Detailed Review of Rogers' Diffusion of Innovations Theory and Educational Technology-Related Studies Based on Rogers' Theory. *TOJET the Turkish online journal of educational technology*, *5*(2). Retrieved from https://go.exlibris.link/PGB3hgTX

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in human behavior*, *69*, 371-380.

Sailer, M., & Homner, L. (2020). The Gamification of Learning: a Meta-analysis. *Educational psychology review*, *32*(1), 77-112. doi:10.1007/s10648-019-09498-w

Salazar, M. S. (2015). The dilemma of combining positive and negative items in scales. *Psicothema*, *27*(2), 192-199.

Salkind, N. J. (2010). Trangulation. In *Encyclopedia of Research Design*. Thousand Oaks, California.

Sánchez-Mena, A., & Martí-Parreño, J. (2017). Drivers and Barriers to Adopting Gamification: Teachers' Perspectives. *Electronic Journal of e-Learning*, *15*(5), 434-443.

Sanchez, D. R., & Langer, M. (2020). Video game pursuit (VGPu) scale development: Designing and validating a scale with implications for game-based learning and assessment. *Simulation & Gaming*, *51*(1), 55-86.

Sbaraini, A., Carter, S. M., Evans, R. W., & Blinkhorn, A. (2011). How to do a grounded theory study: a worked example of a study of dental practices. *BMC medical research methodology*, *11*(1), 1-10.

Schacht, M., & Schacht, S. (2012). Start the game: Increasing user experience of enterprise systems following a gamification mechanism. In *Software for People* (pp. 181-199): Springer.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of human-computer studies, 74*, 14-31.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*: Teachers college press.

Selim, H. M. (2007). Critical success factors for e-learning acceptance: Confirmatory factor models. *Computers & Education, 49*(2), 396-413.

Sendurur, E., Efendioglu, E., Çaliskan, N. Y., Boldbaatar, N., Kandin, E., & Namazli, S. (2017). The M-Learning Experience of Language Learners in Informal Settings. *International Association for Development of the Information Society*.

Shaban, A., & Egbert, J. (2018). Diffusing education technology: A model for language teacher professional development in CALL. *System (Linköping), 78*, 234-244. doi:10.1016/j.system.2018.09.002

Shang, J., Li, F., & Li, H. (2005). "Qing Youxi": Jiaoyu Youxi de Xiwang he Weilai ["Lite" games: The future of educational games]. *e-Education Research, 1*, 24-26.

Silpasuwanchai, C., Ma, X., Shigemasu, H., & Ren, X. (2016). *Developing a comprehensive engagement framework of gamification for reflective learning.* Paper presented at the Proceedings of the 2016 ACM Conference on Designing Interactive Systems.

Smith, H. E., Stair, K. S., Blackburn, J. J., & Easley, M. (2018). Is There an App for That?: Describing Smartphone Availability and Educational Technology Adoption Level of Louisiana School-Based Agricultural Educators. *Journal of Agricultural Education, 59*(1), 238-254.

Stanculescu, L. C., Bozzon, A., Sips, R.-J., & Houben, G.-J. (2016). *Work and play: An experiment in enterprise gamification.* Paper presented at the Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing.

Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement: The influence of school-and individual-level factors on academic achievement. *Education and urban society, 40*(2), 179-204.

Straub, E. T. (2009). Understanding technology adoption: Theory and future directions for informal learning. *Review of educational research, 79*(2), 625-649.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*: Sage publications.

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: a practical guide to their development and use*: Oxford University Press, USA.

Su, C.-H., & Cheng, C.-H. (2013). A mobile game-based insect learning system for improving the learning achievements. *Procedia-Social and Behavioral Sciences, 103*, 42-50.

Sugar, W., Crawley, F., & Fine, B. (2004). Examining teachers' decisions to adopt new technology. *Educational technology & society, 7*(4), 201-213.

Sun, J. C.-Y., & Hsieh, P.-H. (2018). Application of a gamified interactive response system to enhance the intrinsic and extrinsic motivation, student engagement, and attention of English learners. *Journal of Educational Technology & Society, 21*(3), 104-116.

Sun, P. P., & Mei, B. (2020). Modeling preservice Chinese-as-a-second/foreign-language teachers' adoption of educational technology: a technology acceptance perspective. *Computer Assisted Language Learning*, 1-24.

Sun, Y., & Jeyaraj, A. (2013). Information technology adoption and continuance: A longitudinal study of individuals' behavioral intentions. *Information & Management, 50*(7), 457-465.

Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics: International edition. *Pearson2012*.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5): Pearson Boston, MA.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273-1296.

Tamtama, G. I. W., Suryanto, P., & Suyoto. (2020). Design of English Vocabulary Mobile Apps Using Gamification: An Indonesian Case Study for Kindergarten. *iJEP, 10*(1), 150-162.

Tang, S. H., & Hall, V. C. (1995). The overjustification effect: A meta‐analysis. *Applied Cognitive Psychology, 9*(5), 365-404.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53-55.

Teo, T. (2011a). Factors influencing teachers' intention to use technology: Model development and test. *Computers & Education, 57*(4), 2432-2440.

Teo, T. (2011b). Modeling the determinants of pre‐service teachers' perceived usefulness of e‐learning. *Campus-Wide Information Systems, 28*(2), 124-140.

Teo, T., Lee, C. B., & Chai, C. S. (2008). Understanding pre‐service teachers' computer attitudes: applying and extending the technology acceptance model. *Journal of computer assisted learning, 24*(2), 128-143.

Thom, J., Millen, D., & DiMicco, J. (2012). *Removing gamification from an enterprise SNS.* Paper presented at the Proceedings of the acm 2012 conference on computer supported cooperative work.

Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of counseling psychology, 34*(4), 414-424.

Toda, A. M., Oliveira, W., Klock, A. C., Palomino, P. T., Pimenta, M., Gasparini, I., . . . Cristea, A. I. (2019). *A taxonomy of game elements for gamification in educational contexts: Proposal and evaluation.* Paper presented at the 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT).

Tondello, G. F., Wehbe, R. R., Diamond, L., Busch, M., Marczewski, A., & Nacke, L. E. (2016). *The gamification user types hexad scale.* Paper presented at the Proceedings of the 2016 annual symposium on computer-human interaction in play.

Tuparova, D., Tuparov, G., Veleva, V., & Nikolova, E. (2018). *Educational computer games and gamification in informatics and information technology education—Teachers' points of view.* Paper presented at the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO).

Urbina, S. (2014). *Essentials of psychological testing*: John Wiley & Sons.

van Braak, J., & Tearle, P. (2007). The computer attributes for learning scale (CALS) among university students: Scale development and relationship with actual computer use for learning. *Computers in human behavior, 23*(6), 2966-2982. doi:10.1016/j.chb.2006.08.014

Van Rosmalen, P., Wilson, A., & Hummel, H. G. (2015). Games for and by teachers and learners. In *Gamification: Concepts, Methodologies, Tools, and Applications* (pp. 1706-1733): IGI Global.

Vandercruysse, S., Vandewaetere, M., Cornillie, F., & Clarebout, G. (2013). Competition and students' perceptions in a game-based language learning environment. *Educational Technology Research and Development, 61*(6), 927-950.

Vanduhe, V. Z., Nat, M., & Hasan, H. F. (2020). Continuance intentions to use gamification for training in higher education: Integrating the technology acceptance model (TAM), social motivation, and task technology Fit (TTF). *IEEE Access, 8*, 21473-21484.

Varannai, I., Sasvári, P. L., & Urbanovics, A. (2017). The use of gamification in higher education: an empirical study. *International Journal of Advanced Computer Science and Applications, 8*(10), 1-6.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences, 39*(2), 273-315.

Venkatesh, V., & Davis, F. D. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision sciences, 27*(3), 451-481.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science, 46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Vollstedt, M., & Rezat, S. (2019). An introduction to grounded theory with a special focus on axial coding and the coding paradigm. *Compendium for early career researchers in mathematics education, 13*, 81-100.

Wang, L. C., & Ha, A. (2009). Pre-service teachers' perception of Teaching Games for Understanding: A Hong Kong perspective. *European Physical Education Review, 15*(3), 407-429.

Weiser, P., Bucher, D., Cellina, F., & Luca, V. D. (2015). *A Taxonomy of Motivational Affordances for Meaningful Gamified and Persuasive Technologies.* Paper presented at the ICT for Sustainability (ICT4S).

Welbers, K., Konijn, E. A., Burgers, C., de Vaate, A. B., Eden, A., & Brugman, B. C. (2019). Gamification as a tool for engaging student learning: A field experiment with a gamified app. *E-Learning and Digital Media, 16*(2), 92-109.

Werbach, K. (2014). *(Re) defining gamification: A process approach.* Paper presented at the International conference on persuasive technology.

Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative health research, 11*(4), 522-537.

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in human behavior, 32*, 123-132.

Wingo, N. P., Roche, C. C., Baker, N., Dunn, D., Jennings, M., Pair, L., . . . Willig, J. H. (2019). "Playing for Bragging Rights": A Qualitative Study of Students' Perceptions of Gamification. *Journal of Nursing Education, 58*(2), 79-85.

Wong, G. K. (2016). The behavioral intentions of Hong Kong primary teachers in adopting educational technology. *Educational Technology Research and Development, 64*(2), 313-338.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The counseling psychologist, 34*(6), 806-838.

Wu, H.-K., Hsu, Y.-S., & Hwang, F.-K. (2008). Factors affecting teachers' adoption of technology in classrooms: Does school size matter? *International Journal of Science and Mathematics Education, 6*(1), 63-85.

Wu, Y.-L., & Wang, W.-T. (2014). Understanding the effects of mobile gamification on learning performance.

Yee, N. (2006). Motivations for play in online games. *CyberPsychology & behavior, 9*(6), 772-775.

Yıldırım, İ., & Demir, S. (2016). Students' views about gamification based curriculum for the lesson of "Teaching Principles and Methods". *Uluslararası Eğitim Programları ve Öğretim Çalışmaları Dergisi, 6*(11), 85-101.

You, C. J., & Dörnyei, Z. (2016). Language learning motivation in China: Results of a large-scale stratified survey. *Applied Linguistics*, *37*(4), 495-519.

Younas, A., & Porr, C. (2018). A step-by-step approach to developing scales for survey research. *Nurse Researcher*, *26*(3), 14-19.

Yusoff, A., Crowder, R., & Gilbert, L. (2010). *Validation of serious games attributes using the technology acceptance model.* Paper presented at the 2010 Second International Conference on Games and Virtual Worlds for Serious Applications.

Zeng, D., & Zheng, H. (2017). *The Educational Value of the Second Class in Colleges.* Paper presented at the 2017 7th International Conference on Education, Management, Computer and Society (EMCS 2017).

Zhang, M., Yin, S., Luo, M., & Yan, W. (2017). Learner control, user characteristics, platform difference, and their role in adoption intention for MOOC learning in China. *Australasian Journal of Educational Technology, 33*(1), 114-133.

Zhao, Y., & Cziko, G. A. (2001). Teacher adoption of technology: A perceptual control theory perspective. *Journal of technology and teacher education*, *9*(1), 5-30.

Ziesemer, A., Müller, L., & Silveira, M. (2013). *Gamification aware: users perception about game elements on non-game context.* Paper presented at the Proceedings of the 12th Brazilian Symposium on Human Factors in Computing Systems.

Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, *38*(9), 25-32.

APPENDICES

APPENDICES

Appendix A: The original TAI-GLT scale

Table 46. Scale items and the reference (TAI-GLT)

Construct	Code	#	Sub-construct	Item	Source
Perceived enjoyment (ENJ)	ENJ1	1	Flow	Using the tool is pleasurable to students.	(Davidson, 2018)
	ENJ2	2	Flow	Students can forget about time passing while learning with the gamified learning tools.	(Fu et al., 2009)
	ENJ3	3	Flow	Students can become involved in learning with the gamified learning tools.	(Fu et al., 2009)
Perceived usefulness (PU)	PU1	4	Overall PU	My job would be difficult to perform without the tool.	(Davis, 1989)
	PU2	5	Overall PU	Using the tool improves my performance.	(Davis, 1989)
	PU3	6	Overall PU	Overall, I find the gamified learning tool useful in my teaching.	(Bourgonjon et al., 2013)
	PU4	7	Relatedness	In my teaching, usage of the tool is relevant.	(Venkatesh & Davis, 2000)
	PU5	8	Relatedness	The tool would fit the current curriculum.	(Adukaite et al., 2017)
	PU6	9	PU-Product	Gamified learning tools offer opportunities to experiment with knowledge.	(Bourgonjon et al., 2013)
	PU7	10	PU-Product	The quality of the output the students get from the gamified learning tool is high.	(Venkatesh & Davis, 2000)
	PU8	11	PU-Process	Students are happy when using gamified learning tools in schools.	From interview responses
	PU9	12	PU-Process	Teachers enjoys using gamified learning tools in schools.	From interview responses
	PU10	13	Social influence	People who influence my behaviour think that I should use gamified learning tools.	(Venkatesh & Davis, 2000)
	PU11	14	Social influence	People who are important to me think that I should use gamified learning tools.	(Venkatesh & Davis, 2000)
Perceived ease of use (PEoU)	PEoU1	15	Overall PEoU	The gamified learning tool is rigid and inflexible to interact with.	(Davis, 1989)
	PEoU2	16	Overall PEoU	The gamified learning tool often behaviours in unexpected ways.	(Davis, 1989)
	PEoU3	17	Overall PEoU	I find it cumbersome to use the gamified learning tool.	(Davis, 1989)
	PEoU4	18	Accessibility	It is difficult to get access to a suitable gamified learning tool.	(Park et al., 2012)
	PEoU5	19	Compatibility	The gamified learning tools have bad compatibility with the current devices.	(Park et al., 2012)
	PEoU6	20	Complexity	The gamified learning tools are complex to use.	From interview responses
	PEoU7	21	User control	I feel lost control when I'm using the gamified learning tools in teaching.	(Zhang et al., 2017)
Perceived risks (RISK)	RISK1	22	(A lack of) support	The principals, parents or students are possible to refuse using the gamified learning tool.	From literature and interview responses

	RISK2	23	Classroom control	Using the gamified learning tool in school is easy to lose control.	From interview responses
	RISK3	24	Side effects	There are side effects in using the gamified learning tool (e.g. hurt eyesight, distract learners, etc.).	From interview responses
Facilitating conditions (FC)	FC1	25	Necessary resource	I have the resources necessary to use gamified learning tools.	(Wong, 2016)
	FC2	26	Necessary knowledge	I have the knowledge necessary to use gamified learning tools.	(Wong, 2016)
	FC3	27	Workflow	Gamified learning tools fit well into my workflow	(Wong, 2016)
	FC4	28	Technical support	A specific person or group (e.g. technical support team) is available for assistance with difficulties using gamified learning tools.	(Wong, 2016)
Control variable (CV)	CV1	29	Experience	I have used gamified learning tools before.	(Bourgonjon et al., 2013; De Grove et al., 2012)
	CV2	30	Voluntariness	My use of the gamified learning tool is voluntary.	(Venkatesh & Davis, 2000)
	CV3	31	Personal innovativeness	Among my peers, I am usually the first to try out new information technologies.	(Zhang et al., 2017)

Note: The items in italics are those deleted in scale evaluation; all items were evaluated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Appendix B: The validated TAI-GLT scale

Table 47. The TAI-GLT scale (PU section)

13-item PU-GLT scale			5-item PU-GLT scale		
1	PU 12 (ENJ1)	Using the tool is pleasurable to students.	1	PU 14 (ENJ3)	Students can become involved in learning with the gamified learning tools.
2	PU 13 (ENJ2)	Students can forget about time passing while learning with the gamified learning tools.	2	PU3	Overall, I find the gamified learning tool useful in my teaching.
3	PU 14 (ENJ3)	Students can become involved in learning with the gamified learning tools.	3	PU6	Gamified learning tools offer opportunities to experiment with knowledge.
4	PU2	Using the tool improves my performance.	4	PU7	The quality of the output the students get from the gamified learning tool is high.
5	PU3	Overall, I find the gamified learning tool useful in my teaching.	5	PU8	Students are happy when using gamified learning tools in schools.
6	PU4	In my teaching, usage of the tool is relevant.			
7	PU5	The tool would fit the current curriculum.			
8	PU6	Gamified learning tools offer opportunities to experiment with knowledge.			
9	PU7	The quality of the output the students get from the gamified learning tool is high.			
10	PU8	Students are happy when using gamified learning tools in schools.			
11	PU9	Teachers enjoy using gamified learning tools in schools.			
12	PU10	People who influence my behaviour think that I should use gamified learning tools.			
13	PU11	People who are important to me think that I should use gamified learning tools.			

Note: All items were evaluated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Table 48. The TAI-GLT scale (PEoU section)

10-item PEoU-GLT scale			6-item PEoU-GLT scale		
1	PEoU1	The gamified learning tool is rigid and inflexible to interact with.	1	PEoU1	The gamified learning tool is rigid and inflexible to interact with.
2	PEoU2	The gamified learning tool often behaviours in unexpected ways.	2	PEoU2	The gamified learning tool often behaviours in unexpected ways.
3	PEoU3	I find it cumbersome to use the gamified learning tool.	3	PEoU4	It is difficult to get access to a suitable gamified learning tool.
4	PEoU4	It is difficult to get access to a suitable gamified learning tool.	4	PEoU5	The gamified learning tools have bad compatibility with the current devices.
5	PEoU5	The gamified learning tools have bad compatibility with the current devices.	5	PEoU6	The gamified learning tools are complex to use.
6	PEoU6	The gamified learning tools are complex to use.	6	PEoU7	I feel lost control when I'm using the gamified learning tools in teaching.
7	PEoU7	I feel lost control when I'm using the gamified learning tools in teaching.			
8	PEoU8 (RISK1)	The principals, parents or students are possible to refuse using the gamified learning tool.			
9	PEoU9 (RISK2)	Using the gamified learning tool in school is easy to lose control.			
10	PEoU10 (RISK3)	There are side effects in using the gamified learning tool (e.g. hurt eyesight, distract learners, etc.).			

Note: All items were evaluated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Appendix C: The PU-GLT scale (original and validated)

Table 49. The original PU-GLT scale for foreign language learning (FLL)

Construct	#	Code	Survey item	Origin
Compatibility	1	Compa_1	It is easy to find devices to use gamified learning tools in the class.	Interview
	2	Compa_2	Gamification fits well for the size of my class.	
	3	Compa_3	It is easy to find time to use the gamified learning tool in my teaching schedules.	
	4	Compa_4	It is easy to find a satisfactory gamified learning tool for my subject.	
	5	Compa_5	The tool is suitable for the type of knowledge I teach (e.g., memorisation, application, analysis, etc.).	
	6	Compa_6	I have the knowledge to implement gamification in my class.	
	7	Compa_7	If I would like to implement gamification in my school, I can do it without obvious difficulties.	
Curriculum relatedness	8	CR_1	The tools fit the curriculum.	Curriculum relatedness (Bourgonjon et al., 2013) + the last five items
	9	CR_2	It is clear how digital games can be used to fit the curriculum.	
	10	CR_3	The content provided by the tool is relevant to the curriculum.	
	11	CR_4	The content provided by the tool is relevant to the assessment.	
	12	CR_5	It fosters students' awareness about the learning objectives.	
	13	CR_6	It enhances students' knowledge about the learning objectives.	
	14	CR_7	It develops a better understanding of the learning objectives.	
PU-Product	15	PU-Product_1	It better visualises the knowledge points.	Learning Opportunities (Adukaite et al., 2017) revised + the last two items
	16	PU-Product_2	Using the tool helps get higher scores in the assessment.	
	17	PU-Product_3	Using the tool helps get better academic achievement.	
PU-Process	18	PU-Process_1	Gamified learning tools are fun.	Interview responses
	19	PU-Process_2	Gamified learning tools are engaging.	
	20	PU-Process_3	It increases students' interests.	
	21	PU-Process_4	Students are happy when using gamified learning tools in schools.	
	22	PU-Process_5	Teachers enjoy using gamified learning tools in schools.	
Social influence	23	Social_1	People who influence my behaviour think that I should use the tool.	TAM2 Measurement Scales-subjective norm and social image (Venkatesh & Davis, 2000, p. 201)
	24	Social_2	People who are important to me think that I should use the tool.	
	25	Social_3	People in my organization who use the tool have more prestige than those who do not.	
	26	Social_4	People in my organization who use the tool have a high profile.	
	27	Social_5	Having the tool is a status symbol in my organization.	

Note: All items were evaluated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Table 50. The validated PU-GLT scale for foreign language learning (FLL)

Determinant	#	Survey item
Curriculum relatedness	1	The gamified FLL tool fits the curriculum.
	2	It is clear how the gamified FLL tool can be used to fit the curriculum.
	3	The content provided by the gamified FLL tool is relevant with the curriculum.
	4	The content provided by the gamified FLL tool is relevant to the assessment.
PU-Product	5	Using the gamified FLL tool helps get better academic achievement.
	6	The gamified FLL tool enhances students' knowledge about the learning objectives.
	7	The gamified FLL tool develops a better understanding of the learning objectives.
PU-Process	8	The gamified FLL tool is fun.
	9	The gamified FLL tool is engaging.
	10	The gamified FLL tool increases students' interests.
Social influence	11	People who influence my behaviour think that I should use gamified learning tools.
	12	People who are important to me think that I should use gamified learning tools.
	13	People in my organization who use gamified learning tools have more prestige than those who do not.
	14	People in my organization who use gamified learning tools have a high profile.
	15	Using gamified learning tools is a status symbol in my organization.

Note: All items were evaluated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Appendix D: The PEoU-GLT scale (original and validated)

Table 51. The original PEoU-GLT scale for foreign language learning (FLL)

Construct	#	Code	Survey item
Effort expectancy	1	Effort_1	There are good gamified learning tools for FLL. I don't need to design or modify too much.
	2	Effort_2	There is a systematic gamified learning curriculum for FLL. I can just follow it.
	3	Effort_3	I am in control of classroom discipline when using gamified learning tools for FLL.
	4	Effort_4	We have the infrastructures to support gamified FLL (e.g., Wi-Fi and tablets)
	5	Effort_5	Students have digital devices and able to use them in school.
Support	6	Support_1	It is easy to find time to use the gamified learning tool for FLL.
	7	Support_2	If I would like to use gamified learning tools in school, I would not come across difficulties.
	8	Support_3	We have technicians to help solve problems in using gamified learning tools.
	9	Support_4	My colleagues appreciate my attempts in using gamified learning tools.
	10	Support_5	Parents are supportive in using gamified learning tools in school.
	11	Support_6	School policies are supportive in using gamified learning tools in school.
Affordability	12	Affordability_1	Cost is not an obstacle for students or parents.
	13	Affordability_2	Cost is not an obstacle for the school.
	14	Affordability_3	The gamified learning tools are affordable.
Teacher control	15	Control_1	I have the knowledge of where and how to find suitable gamified learning tools.
	16	Control_2	I have the knowledge of how to use gamified learning tools in school.
	17	Control_3	I am in master of the learning content in using gamified learning tools.
	18	Control_4	I am in master of the learning pace in using gamified learning tools.
Availability	19	Availability_1	If I want to use gamified learning tools, I can easily find some.
	20	Availability_2	There are mature, useable and effective gamified learning tools.
	21	Availability_3	The gamified learning tool can cover multiple aspects of learning.
	22	Availability_4	The tool is suitable for the type of knowledge I teach.
	23	Availability_5	Generally, I feel in control in using the gamified learning tools.

Note: All items were evaluated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Table 52. The validated PEoU-GLT scale for foreign language learning (FLL)

Construct	#	Survey item
Effort expectancy	1	There are good gamified learning tools for FLL. I don't need to design or modify too much.
	2	There is a systematic gamified learning curriculum for FLL. I can just follow it.
	3	I am in control of classroom discipline when using gamified learning tools for FLL.
	4	We have the infrastructures to support gamified FLL classes (e.g., Wi-Fi and tablets)
	5	Students have digital devices and able to use them in school.
Support	6	It is easy to find time to use the gamified FLL tool for FLL.
	7	If I would like to use gamified learning tools in school, I would not come across difficulties.
	8	My colleagues appreciate my attempts in using gamified learning tools.
Affordability	9	Cost is not an obstacle for students or parents in using gamified learning tools.
	10	Cost is not an obstacle for the school in using gamified learning tools.
	11	The gamified learning tools are affordable.

Note: All items were evaluated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Doctoral study is a long journey.

Congratulate myself for completing this doctoral thesis happily!

I knew I will make it.

Yes I did.

LOL.

I enjoyed the happiness of creating new knowledge;

I enjoyed the warm smiles;

I enjoyed being with people at UC;

I enjoyed the peaceful life in Christchurch;

I enjoyed the stunning scenery in New Zealand.

I enjoyed everything happened here.

How wonderful the life was in the last three years!

I thank people who provided support;

I thank people who accompanied me;

I thank people who cheered me up;

I thank all the people in my life.

You all are amazing.

Wish everybody be happy!

Zhanni Luo

THANKS FOR READING

Contributing factors of teachers' acceptance intention to gamified learning tools in secondary schools

A scale development study

Zhanni Luo


UC
UNIVERSITY OF
CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

DOCTORAL THESIS

Contributing factors of teachers' acceptance intention to gamified learning tools in secondary schools

A scale development study

Zhanni Luo

Supervised by Billy O'Steen and Cheryl Brown

Zhanni Luo