

Codifying Game-Based Learning: Development and Application of LEAGUE Framework for Learning games

Rabail Tahir and Alf Inge Wang

Dept. of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway

rabail.tahir@ntnu.no

alf.inge.wang@ntnu.no

Abstract Educational games are now seen as effective learning tools. However, there is a gap in the literature regarding the core dimensions of Game-based learning (GBL) for comprehensive design, analysis, and evaluation due to inconsistent use of elements. The literature on GBL reports an extensive diversity of elements used for the design and evaluation of GBL without any categorization of micro and macro-level elements. Hardly any studies systematically decompose these aspects to derivate factors/sub-factors, obstructing identification of any clear pattern. The problem is not the scarcity of GBL research but inconsistency in terminology, scope, definition, and usage of elements leading to the absence of a holistic view of GBL for effective design and evaluation. This study bridges the gap by outlining terminology and scope with four conceptual levels and then systematically categorizing GBL elements by scope, definition, and usage. The methodology used is directed content analysis of GBL literature collected through a previous systematic literature review. Dimensionalization of GBL and further decomposition into factor/sub-factors based on theoretical constructs, has resulted in a consistent and clear pattern delineating the structure of the educational game design, analysis, and evaluation. Further codifying metrics and mapping the relationship among GBL dimensions deduce into a conceptual framework that facilitates greater insight into the process of learning with educational games, where to focus and what to evaluate. The LEAGUE framework can be applied for the design, analysis, and evaluation of educational games. The framework is put in practice by utilizing the framework components to develop three items: LEAGUE analysis instrument, LEAGUE ideation and design toolkit, and LEAGUE evaluation guide that can assist educational game designers, researchers, and evaluators. This paper exemplifies the analysis of learning games using the LEAGUE analysis instrument with one case study as an example.

Keywords: game-based learning, educational games, learning games, framework, dimensions, factors, relations, game analysis, evaluation, design

1. Introduction

Game-based learning (GBL) is an innovative educational paradigm that utilizes games as a mode for transferring learning (Tan, Ling et al., 2007). Educational games are considered to have the potential of deeply engaging learners with any topic, allowing active participation in the learning process (Wallner and Kriglstein, 2011). Although much effort is put in game-based learning studies internationally, robust and comprehensive design for effective learning games remains unclear. Game designers create exciting games but neglect the quality of teaching materials in a game. On the other hand, educators focus on educational materials but do not know about how to create exciting games (Shi, Y.-R. and Shih J.-L., 2015). Bellotti, Kapralos et al. (2013) stated that educational games, like any educational tool, must be able to show that necessary learning has occurred. Therefore, it is crucial to systematically evaluate them to affirm their impact (Marciano, Miranda et al., 2014). The diverse characteristics of learning games make their evaluation a difficult task. However, evaluation is the only means to verify that educational goals are achieved and spot any functional vulnerability (Djelil, Sanchez et al., 2014). Dondi and Moretti (2007) highlighted two critical issues related to GBL evaluation. *First*, the construction of a general framework is extremely difficult unless it is an abstract one. *Second*, the differences between the processes of analytical evaluation (looking at aspects individually based on identification of single aspects) and global or holistic evaluation (considering the characteristics of GBL product all together). There are some difficulties with both approaches. The analytical approach lacks a theoretical model essential for unifying all the different aspects of GBL evaluation and might cause unnecessary fragmentation. Whereas in holistic approach, the judgement is too comprehensive and there is a high degree of subjectivity which presents a need to use other complementary methods for validation.

Many researchers have attempted to describe what the critical elements are to create a learning game. The review of these theories draws only one conclusion: There is no consensus among researchers about the terminology and the comparable importance of GBL elements (Oprins, Visschedijk et al., 2015). Furthermore, it has been a constant challenge to understand the relationships between the different aspects embedded in GBL (Ahmad, Rahim et al. 2015). Our previous review study on GBL highlighted the following problem areas: *1) Most GBL frameworks and studies focus on exploring any single aspect of GBL, making it difficult to identify all core dimensions; 2) Use of a wide diversity of elements for designing and evaluating educational games does not allow the identification of any clear pattern; 3) Very few studies systematically decompose GBL aspects based on their theoretical construct, not allowing the hierarchical decomposition in terms of scope; and 4) The inconsistency in definition, usage, scope and terminology (e.g. dimensions, factors, etc.) of elements in GBL literature.* Therefore, to systematically analyze GBL concepts, there is a need for proper categorization of the wide variety of elements available in the literature (Petri and von Wangenheim, 2017).

This study attempts to overcome the identified problems by performing directed content analysis on the dataset of existing GBL literature collected through a systematic literature review (Tahir and Wang, 2017). The GBL elements extracted from the systematic review are hieratically decomposed (using operations such as coding, categorization, abstraction, comparison and integration) into core dimensions, factors and sub-factors based on scope, frequency of occurrence, relationship between codes, underlying meaning across codes, and mapping to existing theoretical frameworks and constructs defined by researchers in the domain of GBL. The metrics and relations between core dimensions are also detailed (using a similar process) for a complete analysis. The result is a conceptual framework named LEAGUE (**L**earning, **E**nvironment, **A**ffective-cognitive reactions, **G**ame factors, **U**sability, **U**sêr) that list the core GBL elements in a hierarchy of scope. The comprehensive hierarchal structure of the framework makes the application of the framework suitable for analysis, design, and evaluation of learning games. The framework components are used to develop three tools: an analysis instrument, an ideation and design toolkit, and an evaluation guide to assist the GBL community. However, this paper will only focus on the application of the framework for the analysis of learning games. The paper is organized as follows: Section 2 discusses related work, Section 3 describes the method, Section 4 presents the results of directed content analysis in the form of LEAGUE conceptual framework, Section 5 illustrates the application of the framework and, finally Section 6 concludes the study with discussion and future research.

2. Related work

The multidimensionality of GBL demands to consider several aspects important for its design and evaluation (Furió, D. et al., 2013). However, there is still a debate around which aspects to consider (Oprins, Visschedijk et al., 2015).

2.1 Systematic reviews and evaluation studies on educational games

Many review studies in GBL (Perttula, Kiili et al., 2017; Djelil, Sanchez et al., 2014; Petri and von Wangenheim, 2017; Tahir and Wang, 2017) have reported the use of a wide diversity of evaluation aspects for educational games. These aspects are inconsistently defined, and most studies do not systematically decompose into their constituents (Oprins, Visschedijk et al., 2015; Petri and von Wangenheim, 2016). For example, some studies consider the concept interactivity as one of the main dimensions of GBL (Annetta, 2010), while other studies use interactivity in a narrow scope as a factor to achieve a GBL dimension (Djelil, Sanchez et al., 2014). Moreover, others use it as a sub-factor of a factor interface (Omar and Jaafar, 2010). There is no clear distinction between micro- and macro-level elements.

Our insight into evaluation studies showed a similar trend where most researchers used predefined ad-hoc criteria, selecting different aspects for evaluating educational games. Moreover, existing GBL models/frameworks are less used in empirical research (Tahir and Wang, 2017). Virvou and Katsionis (2008) evaluated usability and likeability in the VR-ENGAGE game for education. Pourabdollahian, Taisch et al. (2012) employed flow dimensions for measuring learner engagement in serious games for manufacturing education. Papastergiou (2009) focused on evaluating learning effectiveness and motivation of GBL in computer science education. Giannakos (2013) and Yu, Hsiao et al. (2005) used learning performance as a measure in their evaluation studies. The aspect of usability has also been widely used for evaluating educational games, but different studies used different factors to access this aspect (Liao and Shen, 2012; Mei, Ku et al., 2015; Wallner and Kriglstein, 2011). De Lima, de Lima Salgado et al. (2015) evaluated user experience and motivation in educational games.

2.2 Game-based learning concepts in existing frameworks

Although several GBL design and evaluation models/frameworks exist, it is essential to highlight that each of these models/theories focuses on analyzing and understanding educational games using different aspects, where most researchers focus only on one or two specific aspect(s). Thus they could supplement one another, but individually these studies are relatively narrow and may account only for a portion of a complete picture of GBL design and evaluation (Fu, Su et al., 2009; Tan, Ling et al., 2007). Here are some main aspects explored in various GBL frameworks/models:

Learning: Most of the researchers in GBL mainly focus on education/learning aspects. Four-dimensional framework by (De Freitas and Oliver, 2006) focuses on learning to help tutors evaluate the potential of employing simulation/GBL in practice. Connolly, Stansfield et al. (2009) describe an evaluation framework that focuses on the pedagogical aspect, introducing attributes to measure the GBL environment with attention on the learner and learning. Another evaluation framework proposed by (Wang, Liu et al., 2015) also emphasizes learning perspective with respect to learning results, learner motivations, and learner experience.

Flow: Conversely, some researchers focus on flow and enjoyment aspects in educational games. Kiili (2005) introduced an experiential gaming model to facilitate flow experience serving as a link between game design and educational theory, but not offering a complete game design. EGameFlow proposed by (Fu, Su et al., 2009) is a scale for assessing the level of enjoyment delivered by e-learning games. Kiili, Lainema et al. (2014) presented a flow framework to analyze overall playing experience of educational games through dimensions of flow.

Game design: Serious game design assessment framework implemented by (Mitgutsch and Alvarado, 2012) structures different game design elements to analyze the formal conceptual design of serious games. It recommends on how to shape serious games assessment in terms of design. Chorianopoulos and Giannakos (2014) presented the design principles for serious games in mathematics. Shi and Shih (2015) also focused on game design aspects proposing 11 game factors for GBL design that described a thinking process to design and evaluate educational games using game elements.

Usability: Some researchers focus on usability (Mohamed, Yusoff et al., 2012; Rêgo and de Medeiros, 2015; Omar and Jaafar, 2010) and present heuristics for evaluating the usability of educational games. These researchers incorporated concepts of learning, gameplay, interface, and

enjoyment within heuristics for evaluating GBL usability. Yue and Zin (2009) proposed six usability evaluation constructs for the design of history educational games.

Pedagogy and game design: Some researchers have a combined focus on learning and game design as two critical aspects of educational game design and evaluation. Some of the frameworks include educational games design framework by Ibrahim and Jaafar (2009), a framework for the analysis and design of educational games by Alevi, Myers et al. (2010), an adaptive digital GBL framework proposed by Tan P.-H. et al. (2007), a RETAIN model presented by Zhang, Fan et al. (2010), a GBL evaluation model (GEM) by Oprins, Visschedijk et al. (2015), and a Game object model (GOM) proposed by Amory (2007). Rooney (2012) presented a framework consisting of play, fidelity, and pedagogy for serious game design.

2.3 Content analysis

Qualitative content analysis is a data reduction and sense-making effort that requires data samples to comprise of purposively selected texts that can inform research objectives and attempts to identify meanings and core consistencies (Hsieh and Shannon, 2005; Loh, Sheng et al., 2015). Three different approaches exist for content analysis: Conventional, Directed, and Summative. In the *directed approach*, analysis begins with relevant research findings/theory as guidance for the initial codes, and the goal is either to validate or conceptually extend a theory/ framework. Depending on the research question, it has two strategies to begin coding. If the aim is to identify and categorize all possible instances of any specific phenomenon, then it might be useful first to read and highlight the text representing the instances of that phenomenon and then start coding. The second strategy immediately begins coding with predetermined codes (Zhang and Wildemuth, 2005).

3. Development of framework

This study applied directed content analysis based on the work of (Hsieh and Shannon, 2005). The general objective of this study is both to validate and conceptually extend the existing research on GBL design and evaluation by analyzing, interpreting, and organizing the many aspects to fill the gap in current literature regarding inconsistency in systematic categorization and use of features for GBL design and evaluation.

The content analysis was guided by the following research questions based on problem statements identified in the introduction: *RQ1* What are the core dimensions for the design and evaluation of educational games?; *RQ2* Which factors are important for achieving each of these core dimensions?; *RQ3* What are the sub-factors for assessing these factors based on theoretical constructs (if any)?; *RQ4* What metrics can be used to quantify these factors/ sub-factors for educational game evaluation?; and *RQ5* Are the GBL dimensions interrelated?

3.1 Directed content analysis

The process of content analysis followed the steps defined by (Zhang and Wildemuth, 2005). The steps included: preparing data, defining the unit of analysis, developing coding strategy, testing coding strategy, coding all data, assessing coding consistency, drawing a conclusion from coded data, and reporting method and findings. Content analysis is applied because it plays an important role for theory development. The goal of a directed approach to content analysis was to fill the gap in literature regarding consensus and inconsistency in GBL elements and extend conceptually the existing theory. The existing frameworks reviewed were instrumental in attempting to ensure that there were no omissions and aid in the categorisation process. The directed content analysis was an iterative process involving progressing through extracted data to further analysis using the following

set of operations: coding, categorization, comparison, abstraction, integration, and iteration (Spiggle, 1994; Engl and Nacke, 2013) in such a way that preceding operations shaped the following ones. The analysis was not performed linearly but moving back and forth between stages. The complete process is presented in Figure 1.

3.1.1 Corpus for analysis

The data set for directed content analysis comprised of data extracted from 58 articles on GBL evaluation literature from our previous systematic literature review (Tahir and Wang 2017). The selected articles comprised of GBL frameworks, evaluation studies, and reviews. The corpus completely focused on GBL literature and not on the integration of gaming and learning fields to be in line with (Loh, Sheng et al., 2015). According to Loh, Sheng et al. (2015) the answer to the question “does learning plus game equals to serious games?” is 20 % Yes and 80 % No because only some measures can be commonly found in all three industries. Other than that, the measures are unlikely to transfer from one industry well into another. Therefore, the core elements that are optimal for use in design and evaluation of game-based learning must be specifically focused on GBL literature to properly assess, measure, and improve educational games. The data items extracted from selected papers include: Dimensions, factors, sub-factors, metrics, interrelated dimensions/factors/sub-factors, relation type and/or description, and definitions of dimensions/factors/sub-factors. All the information was entered into an Excel spreadsheet.

3.1.2 Defining unit of analysis

To remove the inconsistency in the terminology used in varying scope across studies, we introduced and defined four conceptual hierarchical levels concerning scope (dimensions, factors, subfactors, and metrics) for analysis of GBL components. Hierarchy is important when defining attributes for a specific application domain (Kececi and Abran 2001). The scope of terminology is defined as follows: the term “**Dimension**” refers to a broader concept but isolated within its kind and not a composition of different aspects, representing the main goals/objects of GBL. Each dimension represents one specific aspect of GBL. The term “**Factor**” refers to the elements important for achieving a specific dimension, and the term “**Sub-factor**” refers to further categorized elements that constitute that specific factor. The term “**Metrics**” is the gauge to measure a factor/sub-factor either through objective or subjective data. This can be depicted (high to low level) as: Dimension > Factors > Sub-factors > Metrics.

3.1.3 Coding strategy

We adopted the first strategy for coding (see Section 2.3) because the aim is to identify and categorize all possible instances of GBL components in the selected corpus systematically and consistently. Therefore, before starting the analysis, we read the text and extracted the text data for each of the four conceptual levels (dimensions, factors, subfactors and metrics) in the spreadsheet that appeared to represent them on first impression (as used in each study), and then started coding for each level (top-down). The definition of these concepts that appeared in text were also extracted. As we wanted to be sure to capture all possible occurrences of GBL elements therefore first extracting all the identified text without coding might be a good way to increase trustworthiness. Hence, the analysis starts with identifying core dimensions and proceeds with factors & sub-factors (with reference to the level above). The extracted data for dimensions were coded using the predetermined codes that emerged from existing GBL literature during initial review (see section 2.2). The predetermined or initial categories used for coding dimensions were learning, game design, flow, and usability (see Section 2.2). Any dimension that could not be

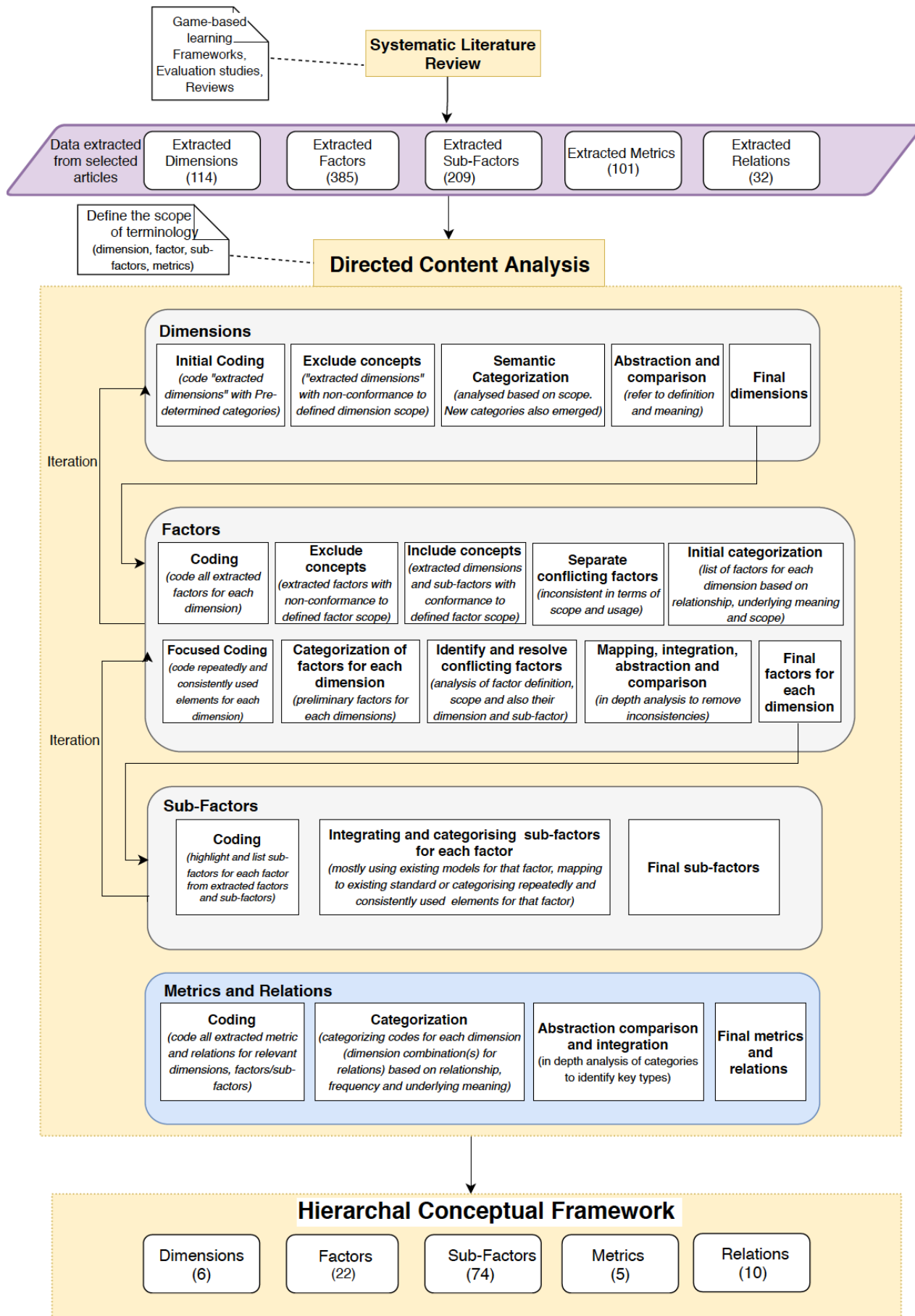


Figure 1: Directed content analysis process

categorized with the initial coding scheme was given a new code. Therefore, new categories appeared during categorization and were further analyzed, using a set of operations mentioned in Section 3.1, until the final core dimensions were attained. During analysis and categorization, comparison was performed to explore the similarities and differences across incidents. Furthermore, several concrete instances were found that shared common features and therefore abstraction was used to group the previously identified categories into more general higher order conceptual classes. Abstraction was also performed on unit of data that was recognised as an empirical indicator of more general construct. After the complete analysis six core dimensions were identified (listed in section 4). The subsequent analysis focused on analyzing the sub-categories, including factors for each dimension, sub-factors for each factor, metrics, and relations using similar operations. The analysis of sub-factors mostly resulted in the integration of constructs where possible by using existing concrete theories/models (e.g. sub-factors of flow were integrated by Csikszentmihalyi's flow model) for aiding the process and enhancing the validity of final GBL components which are theoretically grounded. For metrics and relations, the extracted data was listed for relevant dimension, factor and subfactors. They were then categorized for each dimension (combination of dimension in case of relation e.g. learning and game factors, learning and affective-cognitive reactions etc.) based on frequency, underlying meaning and relationship. The categories formed were further analyzed using operations such as comparison, abstraction and integration to identify the generic metrics types for assessing any factor/sub-factor and key relations between dimensions.

The coding was checked for consistency at each level, where both authors discussed and finalized the categories formed. The analysis and findings resulted in a conceptual framework presented in the next section.

4. The LEAGUE conceptual framework

This section presents the results of directed content analysis in the form of a hierarchical integrated conceptual framework called LEAGUE (see Figure 2).

4.1 Dimensions

Figure 3 shows the six dimensions in LEAGUE identified as key constituents of GBL design and evaluation. The dimensions are presented in the order of letters in LEAGUE and not with respect to the importance of one over the other.

4.2 Factors and sub-factors

Each dimension in the framework has a set of factors. Factors are intermediate-level concepts, and the framework entails such 22 factors (see Figure 4). Factors in the framework are further systematically categorized into *sub-factors* based on their theoretical construct, allowing a hierarchical decomposition. The sub-factors are easier to quantify and also serve to reduce the subjectivity often associated with assessing the factors. However, the choice of components for evaluation should depend on the overall evaluation objective and type of data required. Sub-factors are mostly devised by integration and mapping of conceptual elements using well-developed and widely accepted models/theories in areas where researchers had consensus in the literature. Figure 2 presents the complete hierarchy and association, including sub-factors to each factor.



Figure 2: LEAGUE hierarchal structure and components

Learning/Pedagogical	
Characteristics of an educational game that promote and facilitate learning.	(De Freitas and Oliver 2006; Yoon and Park 2013; Rêgo and de Medeiros 2015; Pappa and Pannese 2010; Omar and Jaafar 2010)
Environment	
Conditions for practical use of GBL and accessibility of educational game in terms of its environment	(Zaibon and Shiratuddin 2010; Djelil, Sanchez et al. 2014; Pappa and Pannese 2010)
Affective Cognitive Reactions (ACR)	
Set of emotions, attitudes and feelings triggered by educational game.	(Rêgo and de Medeiros 2015; Zhang, Fan et al. 2010; Rooney 2012; Djelil, Sanchez et al. 2014; Annetta 2010; Pourabdollahian, Taisch et al. 2012)
Game Factors	
Elements and features of a game environment (the game world) of a educational game that stimulate the users.	(Oprins, Visschedijk et al. 2015; Yue and Zin 2009; De Freitas and Oliver 2006; Yoon and Park 2013; Aleven, Myers et al. 2010; Zaibon and Shiratuddin 2010; Tan, Ling et al. 2007; Rooney 2012; Djelil, Sanchez et al. 2014; Liao and Shen 2012)
Usability	
The extent to which an educational game is usable by the learners to accomplish specific goals (including how to learn, understand, control, and user satisfaction).	(Yue and Zin 2009; Rêgo and de Medeiros 2015; Zaibon and Shiratuddin 2010; Djelil, Sanchez et al. 2014)
User	
Characteristics of the target users/user group of an educational game to investigate if the game matches the indented users.	(Tan, Ling et al. 2007; De Freitas and Oliver 2006; Oprins, Visschedijk et al. 2015; Tan, Ling et al. 2007; Djelil, Sanchez et al. 2014)

Figure 3: Dimensions in the LEAGUE framework

The sub-factors of learning objectives (L1, Figure 2) comes from (Aleven, V., et al., 2010). Learning task/activity (L2/3, Figure 2) is the specific task (designed in line with desired learning objectives and employed learning theory) that outlines the interaction of learners, using specific game characteristics, orientated at specific outcomes (El-Sattar and Hussein, 2016). The sub-factors of learning outcomes (L4, Figure 2) are adapted from the GEM model (Oprins, Visschedijk et al., 2015). The sub-factors of enjoyments (A1, Figure 2) are assimilated from EGameFlow (Fu, F.-L., et al., 2009), which uses flow as a structural foundation, and therefore has some common sub-factors as flow (Rêgo and de Medeiros, 2015; Tseloudi and Tsiatsos, 2015). The sub-factors of engagement (A2, Figure 2) are adopted from the framework by (Pourabdollahian, Taisch et al., 2012). The sub-factors of motivation (A3, Figure 2) are adapted from the well-established ARCS model (Su, Chen et al., 2013). The sub-factors of flow (A4, Figure 2) are adapted from the flow framework (Kiili, Lainema et al., 2014) and present the original component of flow presented by Csikszentmihalyi (Perttula, Kiili et al., 2017). The sub-factors of the interface (U1, Figure 2) are integrated by mapping the factors found in GBL literature to Nielsen's heuristics (Yue and Zin, 2009). This resulted in 9 sub-factors, one heuristic "help users recognize, diagnose, and recover from errors" could not be mapped to GBL literature. The analysis further clarified that in educational games, the focus is on error prevention and confirmation messages rather than error messages. The review of GBL literature showed the lack of psychosocial indicators used for evaluating educational games. Although the importance of psychological needs and psychosocial stages is highlighted in (Tan, Ling et al., 2007), further details are not provided.

Therefore, the psychosocial well-being indicators (Ê3/2, Figure 2) are obtained from PSWBI (Negovan, 2010). The PSWBI scale is validated with students for psychometric properties, construct validity, reliability, and internal consistency. However, its use for educational games is to be explored.

Rabail Tahir and Alf Inge Wang

Learning Objective: Goals and purpose for transfer of knowledge and skills by educational games to users.	(Law and Sun 2012; Alfadhli and Alsumait 2015; Aleven, Myers et al. 2010; Marciano, Miranda et al. 2014)
Learning Strategies: Pedagogical theories, learning models and approaches to achieve learning objectives.	(De Freitas and Oliver 2006; Alfadhli and Alsumait 2015; El-Sattar and Hussein 2016)
Learning Content: Educational material (facts, data and information) provided used for learning in educational games.	(Marciano, Miranda et al. 2014; Alfadhli and Alsumait 2015)
Learning Outcome: Desired learning output (student achievements) from educational games.	(Alfadhli and Alsumait 2015; Wang, Liu et al. 2015; Oprins, Visschedijk et al. 2015)
Technical: The accessibility and easiness of a user to enter the game world with the used technology and specified technical criteria.	(Zaibon and Shiratuddin 2010; Pappa and Pannese 2010)
Context: The particular context where GBL will take place using the educational game.	(De Freitas and Oliver 2006; Perttula, Kiili et al. 2017)
Enjoyment: The extent of playing the educational game is considered pleasurable for an individual user.	(Giannakos 2013)
Engagement: User activity absorption and interest in activity or task, and users' subjective acceptance of games reality combined with degree of focus on this realism.	(Oprins, Visschedijk et al. 2015)
Motivation: Level of user involvement in participating and using an educational game by devoting additional time and effort.	(Djelil, Sanchez et al. 2014; Oprins, Visschedijk et al. 2015)
Flow: State of absolute absorption in an activity which represents an optimal experience.	(Kiili, Lainema et al. 2014; Su, Chen et al. 2013; Fu, Su et al. 2009)
Game Definition: Game goals depending on appropriate set of game rules defining the gaming tasks in games.	(Fu, Su et al. 2009; Oprins, Visschedijk et al. 2015; Djelil, Sanchez et al. 2014)
Game Narrative: Describes what happens in the virtual world including character development and plot.	(Omar and Jaafar 2010; Shi and Shih 2015; Mitgutsch and Alvarado 2012)
Game Mechanics: Defines the functioning, operation and interaction with the game world.	(Yue and Zin 2009; Omar and Jaafar 2010; Mitgutsch and Alvarado 2012; Shi and Shih 2015)
Game Resources: All resources provided to the player (should have both utility and scarcity).	(Alfadhli and Alsumait 2015; Wallner and Kriglstein 2011; Tan, Ling et al. 2007)
Game Aesthetics: Audio-visual language selected, conceptualized and used (images, text, audio, video, animation and multimedia).	(Mitgutsch and Alvarado 2012; Omar and Jaafar 2010)
Game Play: Process by which user reaches the goal which defines the set of challenges of problems the user face to win the game.	(Omar and Jaafar 2010)
Interface: Component through which users interact with a game (elements utilized by user to interact with the educational game that influences its usability).	(Yue and Zin 2009; Omar and Jaafar 2010)
Learnability: Capability of an educational game to enable the user to easily learn to use specific functions.	(Yoon and Park 2013; Tseloudi and Tsiatsos 2015)
Satisfaction: Comfort and acceptability of an educational game to its users.	(Yoon and Park 2013; Tseloudi and Tsiatsos 2015; Ibrahim and Jaafar 2009)
Learner Profile: Attributes of a particular learner/ group.	(De Freitas and Oliver 2006; Djelil, Sanchez et al. 2014)
Cognitive Development: Games should suit learners' cognitive development level.	(Tan, Ling et al. 2007)
Psychological Needs: Games should suit users' psychological needs.	(Tan, Ling et al. 2007)

Figure 4: Factors in the LEAGUE framework

The sub-factors of technical (E1, Figure 2) include: *technology type* (technology used for GBL), *technology issues* (e.g., issue of mobility in mobile technology) and meet technical requirements (Zaibon and Shiratuddin, 2010; Pappa and Pannese, 2010). The sub-factors of context (E2, Figure 2) are adopted from the framework by (De Freitas, S. and M. Oliver, 2006). The sub-factors not directly integrated by using existing theories/models were analyzed using a set of operations mentioned in

Section 3.1 and selected based on scope, frequency of occurrence, and the relationship between codes and underlying meaning across codes.

4.3 Metrics

The metrics represent the lowest level in the hierarchy, which are used to collect evaluation data (Figure 2). The output of a metric interprets the status of sub-factor/factor: the degree to which the educational game possesses a given attribute. The choice of metrics depends on the type of data required, either subjective or objective, qualitative or quantitative. We identified 83 metrics from the corpus of analysis, which were then coded and categorized into five types. The complete exhaustive list of metrics for each factor/subfactor is not provided here. Instead, the aim is to give guidance on the key metrics types used in GBL evaluation that can be utilized and adapted for different evaluation studies depending on the evaluation goals and selected factors/sub-factors. As a result, GBL metrics are coded into five main categories presented in Figure 5. The first three metrics will result in objective data, while the last two will be useful for collecting subjective data. To illustrate, we introduce some typical examples for each category in Figure 6.

Metrics	References
Scores	(Wang, Liu et al. 2015; Petri and von Wangenheim 2017; Papastergiou 2009; Tseloudi and Tsiatsos 2015; Yue and Zin 2009; Giannakos 2013; Yu, Hsiao et al. 2005)
Time	(Connolly, Stansfield et al. 2009; Eagle 2009; Rêgo and de Medeiros 2015; Virvou and Katsionis 2008; Ronimus, Kujala et al. 2014)
Number of occurrences	(Connolly, Stansfield et al. 2009); Mohamed, Yusoff et al. 2012; Eagle 2009; Yoon and Park 2013; Rêgo and de Medeiros 2015; Marciano, Miranda et al. 2014; Virvou and Katsionis 2008; Wallner and Kriglstein 2011)
Rating	(Mohamed, Yusoff et al. 2012; Yu, Hsiao et al. 2005; Fu, Su et al. 2009; Tseloudi and Tsiatsos 2015; Kiili, Lainema et al. 2014; Shiratuddin and Zaibon 2011; Yoon and Park 2013; Zaibon and Shiratuddin 2010; Oprins, Visschedijk et al. 2015; Marciano, Miranda et al. 2014; Shiratuddin and Zaibon 2011; de Lima, de Lima Salgado et al. 2015; Mei, Ku et al. 2015; Pourabdollahian, Taisch et al. 2012; Shi and Shih 2015; Giannakos 2013; Ronimus, Kujala et al. 2014)
reviews/responses/opinions	(De Freitas and Oliver 2006; Connolly, Stansfield et al. 2009; Wang, Liu et al. 2015; Omar and Jaafar 2010)

Figure 5: Metrics in the LEAGUE framework

Metrics	Examples
Scores	Pre/post test score (Petri and von Wangenheim 2017), game performance/score, retention score (Wang, Liu et al. 2015) etc.
Time	Game session time (Rêgo and de Medeiros 2015), learning time, task completion time (Connolly, Stansfield et al. 2009), Time spent on each problem (Eagle 2009) etc.
Number of occurrences	No of usability problems, no of navigation problem (Virvou and Katsionis 2008), no of failures (Yoon and Park 2013), no of levels played (Connolly, Stansfield et al. 2009), no of overall attempts (Wallner and Kriglstein 2011), no of distraction occurrences (Virvou and Katsionis 2008) etc.
Rating	3,4,5,6,7 point Likert scale ((Kiili, Lainema et al. 2014; Shiratuddin and Zaibon 2011; Fu, Su et al. 2009; Zaibon and Shiratuddin 2010; Yu, Hsiao et al. 2005), severity ratings (Mohamed, Yusoff et al. 2012), User task ranking (Connolly, Stansfield et al. 2009) etc.
Reviews/responses /opinions	Expert review (Connolly, Stansfield et al. 2009), self-assessment (Bellotti, Kapralos et al. 2013), Mood and attitude statements (Connolly, Stansfield et al. 2009), game review based on selected (sub)factors (De Freitas and Oliver 2006), comments on positive/negative game aspects (Wang, Liu et al. 2015) etc.

Figure 6: Examples of the metrics

4.4 Relations: High-level abstraction of game-based learning

The highest abstraction of the framework is displayed in Figure 8. We identified ten key relations from directed content analysis presented in Figure 7 (see Figure 8 for the direction of relations).

The structure of GBL is depicted by the hierarchal layout presented in Figure 2. The high-level abstraction of the LEAGUE framework describes the internal operation of GBL and classifies the six dimensions into generic and domain-specific. Learning, Game Factors, and Affective-Cognitive Reactions are the core domain-specific dimensions that represent the GBL phenomenon and process. Environment, Usability, and Users are the generic dimensions that influence the core dimensions and are essential for any software application to be effective for its users. An educational game is a game for education purposes that imparts learning by involving learners in the learning process. Game Factors generate Affective-Cognitive Reactions that absorb users in playing the game and positively influence the Learning. The main trick for an effective GBL approach is to keep generic dimensions in line while tweaking the Learning and Game Factors dimensions to integrate, create a balance and work in accordance with each other for enhancing the Affective Cognitive Reactions in order to meet the purpose of the educational game. We have introduced a term T-relation (see Figure 8) for the association between Learning, Game Factors, and Affective-Cognitive Reactions as the core process of GBL, where the integration of game and learning enhance affective reactions (Kiili, K., 2005). The generic dimensions not only influence the GBL phenomenon (domain-specific dimensions) but are also linked with each other. Usability should address the intended users and also cater to the technical and context related specifications of the environment. The technical specification and context (environment) should also map to the learner profile and capabilities (user). The overall process of GBL is a complex phenomenon and requires a multidisciplinary approach.

Dimensions	Relation	References
Learning & Game Factors	Integration/ Balance	(De Freitas and Oliver 2006; Aleven, Myers et al. 2010; Zhang, Fan et al. 2010; Wallner and Kriglstein 2011; El-Sattar and Hussein 2016; Pappa and Pannese 2010)
ACR & Learning	Positive effect	(Kiili, Lainema et al. 2014; Thomas, Schott et al. 2004; Oprins, Visschedijk et al. 2015); Perttula, Kiili et al. 2017; Bellotti, Kapralos et al. 2013; Yu, Fu et al. 2009; Tseloudi and Tsiatsos 2015; Giannakos 2013)
Game Factors & ACR	Generate	(Rooney 2012; Abdul Jabbar and Felicia 2015; Ronimus, Kujala et al. 2014)
(Integration of Game Factors and Learning) & ACR	Enhance/facilitate /increase	(Kiili 2005; Zhang, Fan et al. 2010; Rooney 2012)
Usability & (Learning, Game factors, ACR)	Influence	(Gibson and Bell 2013; Shi and Shih 2015)
User & (Learning, Game factors, ACR)	Influence	(De Freitas and Oliver 2006; Abdul Jabbar and Felicia 2015; Mei, Ku et al. 2015; Yu, Hsiao et al. 2005; Mitgutsch and Alvarado 2012; Shi and Shih 2015)
Environment & (Learning, Game factors, ACR)	Influence	(De Freitas and Oliver 2006; Oprins, Visschedijk et al. 2015; Dondi and Moretti 2007; Virvou and Katsionis 2008)
Usability & Environment	Address/Cater	(Virvou and Katsionis 2008)
Usability & User	Address/ Cater	(Virvou and Katsionis 2008; Mei, Ku et al. 2015)
User & Environment	Map	(De Freitas and Oliver 2006)

Figure 7: Relations in LEAGUE

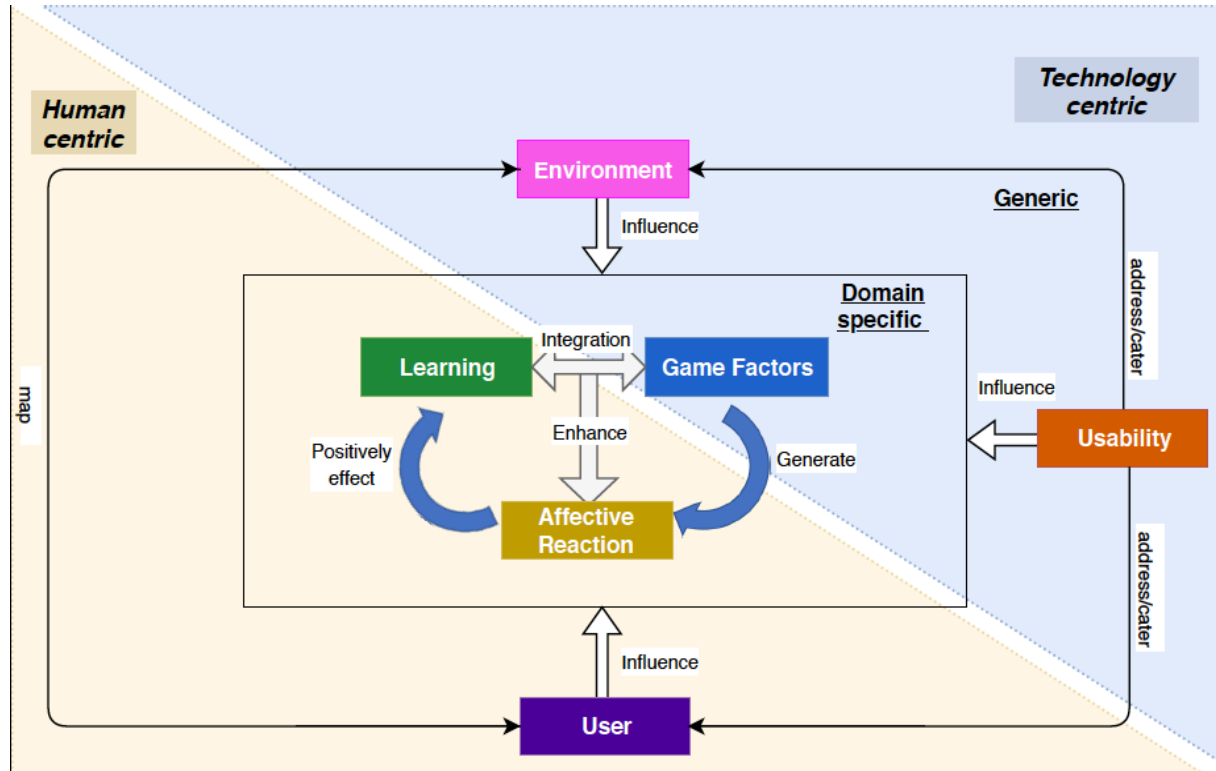


Figure 8: High-level abstraction of LEAGUE

There is another viewpoint to the LEAGUE framework, which divides it into two views: technology-centric, and human-centric. The technology-centric view includes three dimensions (Game Factors, Usability, and Environment) related to technological aspects of game-based learning and the human-centric view also includes three dimensions (Learning, Affective-Cognitive Reaction, and User) related to human aspects (cognitive, behavioral, identity) of game-based learning. The idea here is to model the technology-centric dimensions in such a way that they facilitate human-centric dimensions.

5. Application of the LEAGUE framework

The framework organizes the GBL dimensions regarded in the literature as significant in producing an effective learning game. For each dimension, a set of configurable factors, sub-factors, and the basic vocabulary is provided to facilitate the application and use of framework components in multiple ways.

The framework can be applied in the process of design, analysis, and evaluation of game-based learning. The dimensions in GBL (depending on the evaluation objective) might be considered in isolation (picking and selecting components) for a specific evaluation study. However, GBL dimensions are linked to each other in terms of cause and effect and can be viewed as a collective whole to understand the process and help in design and analysis. The framework is put in practice by utilizing the framework components to develop three items: LEAGUE analysis instrument, LEAGUE ideation and design toolkit, and LEAGUE evaluation guide for analysis, design and evaluation of learning games respectively. Using each of the three items is divided into a set of stages or activities to scaffold the process. This paper will primarily focus on using the framework for the analysis of learning games.

5.1 Analysis of learning games using LEAGUE analysis instrument

This section presents how the framework is applied for the analysis of learning games using the LEAGUE analysis instrument. The instrument is aligned with the framework in terms of the elements

required to specify an educational game. An example study is also presented which demonstrated its use. It can be useful to understand the potential of educational games in a specified environment by knowing the strengths and areas in which it could improve. The learning game is analyzed based on whether it contains the elements deemed important for game-based learning by splitting the subject matter into its basic and fundamental components.

The analysis instrument consists of three parts: a primary analysis form, a secondary form to reinforce in-depth analysis, and a reflection form. The primary analysis form (see Figure 9) breaks the complex GBL paradigm into smaller parts based on framework components to gain a better understanding of how learning takes place in the learning game. Each question in the form regarding all six dimensions is answered for the game being analyzed. The factors laid out in the primary form may help analyze the learning game in terms of individual elements essential for an effective game-based learning approach and highlight any areas that are weak or neglected in the game. At the end of the form there are two rows to list the overall strength and weakness of the game with respect to each dimension by critically analyzing the given answers.

Learning	Environment	Affective-Cognitive Reactions	Game Factors	Usability	UsÊr
L1. What are the learning objectives of the game?	E1. What technical aspects are required for the game to work and best support learning?	A1. How can the game provide enjoyment to the users?	G1. What are the game objectives to integrate learning?	U1. How interface of the game is made easy to use for target users?	Ê1. What are the attributes of the target users of the game?
L2. Which learning strategies are being used to enable learning through the game?	E2. What is the context for playing the game for learning?	A2. How can the game engage the users?	G2. What narrative is used to make game compelling and integrate learning?	U2. How does the game provide easy learnability to its target users?	Ê2. Which cognitive needs (of target users) are considered in the game?
L3. What learning content is being used in the game for target users?		A3. How can the game motivate the users?	G3. What mechanics are used to make game compelling and support learning?	U3. How does the game provide satisfaction to its target users?	Ê3. Which Psychological needs (of target users) are considered in the game?
L4. What learning outcome(s) can be acquired from the game?		A4. How can the game generate flow?	G4. What resources are provided to the users to function effectively that also support learning?		
			G5. What aesthetics are used to make game compelling for target users?		
			G6. What game play is used to make game compelling for target users and support learning?		
Strength of the analyzed game					
Weakness of the analyzed game					

Figure 9: Primary analysis form

Learning	Environment	Active-Cognitive Reactions	Game Factors	Usability	Usêr
<p>(Prior knowledge): What knowledge or skills do players need to have before starting the game?</p> <p>(Learning and retention): What knowledge or skills game is reasonably expected to teach players?</p> <p>(Potential transfer): What knowledge and skills might game teach players that go beyond what they actually encountered in the game?</p> <p>Learning style: Which learning styles need to be incorporated in the game?</p> <p>Learning theory: Which pedagogic models/learning theory is most effective for attaining learning objectives?</p> <p>Learning task/activity: What learning activity/task learners perform in the game to achieve learning objectives?</p> <p>Learning material: What learning material need to be incorporated in the game?</p> <p>Instructional support: How much instructional support is provided to support learning objectives?</p> <p>Difficulty levels: How many levels of difficulty are provided to players to attain learning objectives?</p> <p>Knowledge/skill/attitude: What knowledge, skills or attitudes are acquired by the learners by playing this game?</p> <p>Competencies: Does learner gain any competencies by playing this game?</p> <p>Performance: Does playing this game enhance performance of learners?</p>	<p>Technology Type: Which type of technology is used in the game to best support learning? (E.g. mobile, VR, AR, Kinect, etc.)</p> <p>Technology related issues: Are there any issues related to the used technology that need special attention or will affect learning?</p> <p>Technical requirement: Does the game has any specific technical requirements that must be meet? (E.g. functional req., Portability & conformance to standards, Structure & Organization, Technical quality, Information produced etc.)</p> <p>Place: Does the game need to be played on some specific place and how does game facilitate that? E.g. home, school, combination of several.</p> <p>Setting: Does the game require any specific settings to be played and does that affect learning? (E.g. access to equipment, technical support, additional resources etc.)</p>	<p>Which elements are used in the game to induce enjoyment?</p> <ul style="list-style-type: none"> •Immersion •Social interaction •Challenges •Goal clarity •Feedback •Concentration •Control •Knowledge improvement <p>Which elements are used in the game to induce engagement?</p> <ul style="list-style-type: none"> •Immersion •Challenges •Control •Purpose •Interest <p>Which elements are used in the game to induce motivation?</p> <ul style="list-style-type: none"> •Attention •Relevance •Confidence •Learner satisfaction <p>Which elements are used in the game to generate flow?</p> <ul style="list-style-type: none"> •Challenges •Clear clarity •Feedback •Playability •Control •Rewarding experience •Concentration •Loss of self-consciousness •Time distortion 	<p>Game Goals: What are the goals (that match with the learning objectives) of the game that players need to achieve?</p> <p>Game rules: What are the appropriate set of rules/constraints the players must follow to reach the goals of the game?</p> <p>Game tasks: What are the tasks (defined by the set of rules and goals) players has to complete in the game?</p> <p>Player characters: Does game use characters to tell narrative?</p> <p>Storyline: Does game need a storyline to effectively impart learning?</p> <p>Fantasy: Does game use fantasy to involve players?</p> <p>Game Interactions: How do players interact with the game?</p> <p>Game Controls: How do players control the game?</p> <p>Game tutorial or help: Does game provide tutorial or help to players if they need support?</p> <p>Rewards or sources: Does game provide rewards or other suitable resources to the players (in accordance with game objectives) to function in the game world?</p> <p>Game customizability: Does game provide any customizability options important for the players?</p> <p>Multimedia elements: What multimedia elements are used to support learning and involve players? E.g. text, audio, images, animations, video, interactive content etc.</p> <p>Game visualization: What level of visualization is used to support learning and involve players? E.g. 2D, 3D etc.</p> <p>Challenges: Does the game provide suitable challenges to target players?</p> <p>Strategy: Does game require players to make a strategy to achieve the game goals?</p> <p>Pace & adequate levels: Does game provide different levels with appropriate difficulty and pace?</p> <p>Game Feedback: Does game provide appropriate feedback for in-game activity?</p>	<p>What elements are used to make the interface easy to use?</p> <ul style="list-style-type: none"> •Provide Interface Feedback •Use of metaphors and objects •Provide user control and settings •Provide consistency •Provide error presentation •Provide easy navigation •Provide adaptability and accessibility •Provide attractive screen design •Provide Help and support <p>U2. How does this game provide easy learnability to its target players?</p> <p>U3. How does this game provide satisfaction to its target players?</p>	<p>Bio-demographics: What are the bio-demographics of the players that will mostly play this game? (E.g. age, gender, ethnicity, economic background etc. that should be kept in mind.)</p> <p>Experience: What level of experience they have with technology and learning domain?</p> <p>Personality: What are the learning styles and preferences of target players?</p> <p>Cognitive Stage: What is the cognitive stage of players of this game?</p> <p>Cognitive Load: How does this game support the level of cognitive load suitable for them?</p> <p>Psychosocial Stage: What is the psychosocial stage of players of this game?</p> <p>Psychosocial Well-being: How does this game improve their psychosocial Well-being?</p>

Figure 10: Secondary in-depth analysis form

Table 1: Reflection form

No.	Reflection questions
R1.	Are game factors (game objectives, narrative, mechanics etc.) and learning factors (learning objectives, strategy, content etc.) well integrated into the game?
R2.	Are selected game factors (narrative, mechanics, play etc.) effective for generating affective-cognitive reactions (engagement, enjoyment etc.) in target users of this game?
R3.	Does the usability of this game cater to the needs of the target users?
R4.	Does the usability (interface, learnability etc.) of this game cater to any specific needs of the environment (technology, context etc.) in which it will be played?
R5.	Are specific technical and context requirements (if any) for playing this game easy to manage by the target users?
R6.	Are learning factors (objective, content, strategy etc.) appropriate for the target user?
R7.	Are game elements (mechanics, narrative, play etc.) used appropriately for the target user?
Revisions/improvements needed in the analyzed game:	

The secondary form (see Figure 10) supports an in-depth analysis of each element of primary form by splitting it further into simpler questions to deeply analyze its constituents and facilitate the thinking process to construct the answer for each question in the primary form. The secondary form can be used on its own or can also be used only as a guide to give concrete answers in primary form. Building on this analysis, the relationship between individual components will be considered using the reflection part. The reflection questions (see Table 1) determine the effectiveness of the learning game as a collective whole by analyzing the relations between individual components. This is giving the opportunity to reflect on the design choices made in the game and if they are effective or need improvement. For example, if the game is appropriate for target users in terms of usability, learning content, strategy or game elements used; if the game can be used to support different context and informal or formal learning considerations; and if the game has a balance and harmony in learning and game factors used etc. Similarly, for reflection form each of the seven reflection questions are answered for the game being analyzed and revisions or improvements needed in the game are listed in the last row based on the reflections made. Therefore, the steps of the framework application for analyzing a learning game using LEAGUE analysis instrument consists of the following steps in the specified order: 1) description of the different aspects of the game by answering the questions in the primary form (Figure 9), using secondary form (Figure 10) for in-depth analysis and detailed description (if needed), 2) identification of strengths and weaknesses of the game by analyzing the answers, 3) reflection on the relationships between core elements of the game (using Table 1) and 4) critically analyzing the reflections made to highlight the refinements or improvements needed in the game to make it effective.

The instrument can be used to support an analysis process undertaken by any stakeholder (designers/developers, researchers, and intermediates like teachers/parents etc.) to ensure that they take into account the key issues and essential factors associated with game-based learning to support practice. The analysis of learning games using the three forms provides the analytical as well as holistic picture of whether the learning game is an effective GBL approach and not just the tool itself. This analysis will include all aspects worth considering, from the right content and strategy to appropriate game elements and software for target users in order to apply the game within the specified context. To illustrate the application of the framework for the analysis of learning games, the instrument is used to analyze an empathy game as an example.

5.1.1 Case: Empathy game

The illustrated game is an online empathy game for primary school children (8-14 years old), which can be played in school or at home with friends. The game is about making stories for different

characters using personality traits to develop a strategy for achieving a goal. One of the players plays the role of the selected character, and the other two players help the character to complete the tasks assigned according to personality traits and goals, which creates the challenge factor. The game uses this activity to teach empathy and enable players to develop the ability to understand and share the feelings of another. The game is suitable for interactive and collaborative learning.

Completing the analysis instrument highlights the particular challenges, strengths, and weaknesses in terms of essential GBL components needed to embed desired learning through the game into effective practice. The beta version of this game was analyzed using the LEAGUE analysis instrument, and the results are presented in Figure 11 for the primary form and Figure 12 for the reflection part. The secondary form was used as a guide to thinking more in-depth and give a concrete answer for questions in the primary form. The last part of the primary form (the strength and weakness of the analyzed game) and the reflection form (the revisions/improvements needed in the analyzed game) are not shown in the figures but instead described below.

Strength of the analyzed game: In this particular game, a notable strength is a collaborative learning approach that might be used to support the cyclical transition from storytelling towards developing empathy through discussion and reflection on actions. However, learning is not tied to any curriculum content, developing a strategy to achieve a goal using personality traits, and the outcome facilitates reflection and additionally may support creative thinking and imagination. This approach has the potential to additionally teach cooperation, listening, and improve storytelling after time as it uses social interaction between players, as well as this help to engage learners who prefer self-directed and interactive ways of learning. Another strength is the use of characters to tell the story. The game does not require significant technical support and can be easily played with a device with an internet connection. The game also uses bright colors and simple and consistent screens with straightforward controls and navigation.

Weakness of the analyzed game: On the other hand, game aesthetics lack the effective use of multimedia. The game does not provide any help or support for playing the game first time and also lacks the use of rewards or other resources to generate additional purpose to engage in storytelling or facilitate reflection through the use of props. One of the least successful aspects is that the game does not provide any instructional support to facilitate children to feel emotions or differentiate between them to generate empathy unless they self-reflect on their story. The players are not given the control to move back in the game even if they accidentally press a button or miss a task, which is a significant drawback along with no feedback is given for in-game activity. Game is mainly text-based, not much visualization in the game. The tasks are not very clear and challenging to understand at least the first time (reading through text) and could be supported with multimedia usage, e.g., audio, animation, or videos. Also, the game does not provide tasks with an increasing difficulty level.

Reflections made: The reflection tool supports a deeper reflection on the interrelation between different essential elements to apprehend whether the learning game is effective for the purpose. In this case, depending on device availability, learners may want to play the game in school as well as home context supporting formal as well as informal learning processes, reinforcing the learning outcomes. However, the game requires three players to start the game but does not support connecting with players online, so this might be a problem in a context where more children are not present. The player should know how to read as the game is mostly text-based, and also have the vocabulary to create a story as storytelling in the main activity in the game, which makes it more

suitable for talkative children who can put their thoughts into words. The challenge for helping a character to achieve the goal give players confidence to verbalize their thoughts into a story without being self-consciousness as they are playing for another character and not themselves. The game uses characters and traits that resemble different types of children in the target age group, and thus provides a safe space for children to talk about situations that may be hard to discuss outside of a game environment. The complete analysis of the reflection part is presented in Figure 12.

Revisions/improvements needed in the analyzed game: From the analysis, the game designer can identify the need for a better fit between the use of game elements to facilitate the learning outcomes and to emphasis on creating greater challenges for the children and more importantly support increased reflection upon empathy learning through instructional support and feedback and thereby offer improved opportunities to work in a team.

6. Discussion and conclusion

This study expands on GBL design and evaluation literature to overcome the shortcomings in current research (problem areas highlighted in the introduction) by conducting a directed content analysis. The results of the analysis are translated into a conceptual hierarchal framework LEAGUE, which shows that the multidimensionality of GBL requires evaluation of several aspects referred to as core dimensions (RQ1), including Learning, Game Factors, Affective-Cognitive Reactions, Usability, User and Environment. Each dimension focuses on certain factors and sub-factors that constitute that aspect, and metrics are required to assess them. The framework presents 22 factors (addressing RQ2), 74 sub-factors (RQ3), and five metrics categories (RQ4). The dimensions of GBL are related to each other, and it is essential to assess the relations presented as a high abstraction of LEAGUE for more significant insights into educational games (RQ5). The framework provides a detailed picture of GBL that will guide not only researchers and evaluators but also designers and developers of educational games. The proposed framework is built on components grounded in theory. Each component has a strong basis for formation that is supported by theoretical constructs in GBL literature and not merely based on suspicion.

Most of the existing GBL frameworks focus on some specific elements, which make them difficult to use in practice when the target genre differs from default game genres used in research (Shi and Shih 2015) or when the objective is design and analysis of complete GBL experience and not just focusing on few individual aspects of it. Thus, the existing models and frameworks could supplement one another, but individually these are relatively narrow and focus on a portion of the complete picture of GBL design and evaluation (such as Learning Mechanics-Game Mechanics (LM-GM) model can help identify the learning and game mechanics to draw the LM-GM map for a game but neglect the other elements that equally account of an effective learning game). The specificities of LEAGUE in relation to other frameworks can be highlighted by the comprehensiveness of the framework (detailing the individual parts in order to allow analysis in terms of presence/absence as well as overall picture and interconnection between the core dimensions) that can be equally useful for analytical and holistic evaluation providing a theoretical model essential for unifying all the different aspects of GBL and thereby solving the two critical issues related to GBL design and evaluation highlighted by Dondi and Moretti (2007) (see introduction). Furthermore, dimensions presented in LEAGUE are higher-level concepts and not restricted by the game genre.

The framework is put into practice by developing three tools to support GBL practitioners and researchers. The framework can be applied for design, analysis, and evaluation of educational games using the LEAGUE analysis instrument, ideation and design toolkit, and evaluation guide. This

Learning	Environment	Active-Cognitive Reactions	Game Factors	Usability	User
<p>L1. Players by putting thoughts into words can better develop skills on empathy, be able to reflect on everyday situations of self and others, and might also learn cooperation, listening skills, imagination and storytelling by playing game.</p> <p>L2. The players learn by telling stories with group discussion and reflection. The game employs behaviorism and social constructivism based on collaborative learning approach. Players develop a story based on personality traits for achieving a life goal by putting oneself in the shoes of the character.</p> <p>L3. Prompts (goals) guiding on what to talk about and with what foundation (traits). The game doesn't provide any instructional support and difficulty levels.</p> <p>L4. Change in attitude towards others in understanding their situations and not judging based on their few traits. Might support to develop competences in cooperation, listening and improve storytelling after time.</p>	<p>E1. A tablet or PC with web browser and internet connection is required.</p> <p>E2. The game can be played anywhere (aim is to use in schools) but only collaboratively as it requires three players at a time and the game doesn't facilitate connecting with other players online. One device is required to play the game with three people.</p>	<p>A1. Game uses social interaction between players, challenge for helping character to achieve the goal and possibility of immersion in developing a story of the character. The game should also provide feedback and control over task which is major drawback. The sub-goals are not very clear and difficult to understand at least first time.</p> <p>A2. Game induces interest through character, provide control over building story and choice of goal.</p> <p>A3. Game involve attention by making players listen to each other to develop their part of story, provide relevance of characters to target players' life; gives confidence to verbalize their thoughts into a story.</p> <p>A4. Game provide control over the story progression, concentration on listing to peer's part of story, and loss of self-consciousness as by playing for another character.</p>	<p>G1. Complete a story by developing a strategy to achieve a goal. The players need to use the traits given for that character for achieving the goal and build on the story of the other players. The players choose a character and goal to work on developing a story both individually and in groups as specified for the task. However, the players always fail to achieve the goal.</p> <p>G2. The game uses characters for story telling that is supposed to generate empathy. The game does not use fantasy elements, but players can create a fantasy world around character using their imagination.</p> <p>G3. Use simple clicking for selecting a character, goal and traits. Main game mechanic is social interaction that uses high level of interactivity by playing in team. Game is controlled with a continue button, but no back button is provided to go back and change strategy in story.</p> <p>G4. There is no game tutorial or help to understand gameplay. No resources (other than the given traits) or rewards are provided. The players can customize their character only by choosing a name.</p> <p>G5. Game is largely text-based, cartoonish images for characters and colorful background. Not much visualization in the game largely roleplay driven rather than immersive 2D/3D interface.</p> <p>G6. The only challenge is to build a story around the given traits and goals of selected character. Players need to make a strategy for achieving goal and change it when they fail. Game does provide tasks with increasing difficulty level, it may be easy or hard depending on the matching of the traits and goal. No feedback is given for in-game activity.</p>	<p>U1. Provide consistency in UI, easy navigation, adaptability from tablet to computer screen, attractive and playful color scheme and image design, give player the control on when to move forward but they cannot go back if needed. No feedback is provided for interface action accept disabling button. A lot of text on screen. No help/support.</p> <p>U2. All screens and activity are similar every time except the content, after first play they are likely to feel comfortable in the next round.</p> <p>U3. Use of bright colors and cartoonish characters. Includes variety of characters and traits that resemble different types of children.</p>	<p>E1. Children aged 8-14, all genders, ethnicities and backgrounds. No experience required as interactions are simple. Player should know to read and have vocabulary to create a story as game is largely text-based and storytelling. Potentially would support all types of personalities but is more fit for talkative children.</p> <p>E2. The game is made to be rather short, so it wouldn't get boring.</p> <p>E3. Empathy development is important for target users. Game provides a safe space for children to share stories and interact with each other by having some shared tasks. Gives confidence to talk about situations that may be hard to discuss outside of a game environment.</p>

Figure 11: Analysis instrument for empathy game: primary form

No.	Reflection questions
R1	A good integration is the use of personality traits as a game rule to make a strategy. The connection between the game task and reflection upon learning process was the least successful aspect. No use of instructional support for game task and no use of game elements to facilitate reflection, feel empathy for the character, justify their actions or visualize different personality traits. The game does not provide any feedback to facilitate empathy or trigger feelings, it only depends on the storyteller. Induce the fantasy world created by children's' story by making characters react to story and using them in story progression as animation to help children trigger emotions and feel empathy. The players always fail to achieve the goal as a game objective, which initially might help to reflect on traits and feel emotion as sadness for characters' failure but later will confuse them to differentiate between using different traits as they always fail and may end up generating negative emotions only.
R2	Use of characters and social interaction is a major hit otherwise the game is a bit weak in terms of game factors to generate affective reactions. Should give players more control to go back if they misunderstood the task or want to elaborate on their story. Game is monotonous in terms of difficulty level and same result every time. The only challenge is the combination of traits and goal to tell the story that might become boring after players have tried a couple of combinations. No customizability options or use of rewards/other resources to generate interest. No in-game feedback. Multimedia usage is limited, using immersive animation, videos or 3D visualization can increase enjoyment and engagement.
R3	The game is designed to be rather short, so it wouldn't get boring as intension span for children is short. The use of cartoonish characters is in accordance with age group. However, too much text on screen makes screen design less attractive for children. It is difficult to read all text and understand the task and role of each player at least first time. Considering age group, the interface should provide the control to move back in game, if children skip one task because they misunderstood and move forward they are likely to miss the whole game as tasks are interlinked.
R4	The interface of the game is consistent and provides adaptability from tablet to PC. The game needs to be played with three users and therefore provides easy navigation but there should be some feedback and metaphors in screen design (in addition to text) to guide which player is leading.
R5	Game does not require any significant technical support and is therefore appropriate for target age group. The classroom context is more appropriate to play this game as there are other students to play with. In other context the only issue is to have other children to play with as game requires three players. The game should therefore facilitate connecting with friends online who wants to play as the game cannot be played alone.
R6	Learning objective that is empathy development is important for target children. Use of collaborative learning approach provides a safe space for children to share stories and interact with each other by having some shared tasks. It is also appropriate as some children are not talkative and cannot verbalize their thoughts, so teamwork will help and give them confidence to talk about situations that may be hard to discuss outside of a game environment. Overall simple content is used in the game, but some traits are difficult to understand for children age 8-14. Instructional support for generating empathy is missing as children in this age group might need a feedback or some guidance to develop empathy based on their story or to reflect on it but the game does not facilitate children in feeling the emotions or prompts to trigger emotions.
R7	Children can relate to different characters and traits used in the game as they target this age group. Using role playing is appropriate as even with very basic or no interactive image children can pretend to be someone/something else very easily. The use of multimedia is very less for children and no help or support is provided to understand task. Text should be supported with use of animations in the form of tutorial or interactive content. Individual game tasks might be difficult for some children who find it difficult to verbalize their thoughts or are less imaginative and may require assistance.

Figure 12: Analysis instrument for empathy game: reflection form

paper described how the framework could be applied for the analysis of learning games, exemplified with a case study of an empathy game. The LEAGUE analysis instrument can be used by the stakeholders of educational games, including designers, researchers, and intermediates like teachers/parents, to ascertain its effectiveness. Educational game designers/developers can use it to analyze the educational game (both early stage game prototypes or later alpha/beta versions of the game) to identify the loop-holes and make improvements in the design. It is often recommended to carry out analysis early because it is easy to make changes and improvements at an early stage of development as they get more expensive later, and it is useful to carry out analysis before actual evaluation. Educational game researchers are interested in gaining insights regarding the game's effectiveness and suitability in different domains with respect to its designated purpose and application context. Therefore, they can use this instrument to learn more about the different elements used in the games and the relationships between them and gain experience from both successful and failed game concepts in order to improve in designing effective educational games and critically question the effects and consequences games may have on target users, especially in the case of vulnerable groups. Teachers/parents need to be convinced of the positive effect of game-based learning because otherwise, they will choose not to use them. Therefore, they can use the analysis instrument for analyzing the educational game to assess the potential and develop trust and conviction for justification to use the game as an efficient tool or not.

The proposed framework is also employed in workshop sessions for ideation and design of learning games for various learning domains using the LEAGUE toolkit, and to conduct evaluation studies. The complete process of using the framework for design and evaluation of learning games along with the results will be presented in another paper, and framework components will be further validated and developed. Future research will focus on automating or partially automating GBL evaluation using the proposed framework and game data logs. The future work will also focus on developing a web-based ideation and evaluation tool using the proposed framework components that will facilitate the educational game ideation and evaluation process during different phases of the development lifecycle and help professional and game companies working with game-based learning.

References

- Ahmad, M., et al. (2015). Towards an Effective Modelling and Development of Educational Games with Subject-Matter: A Multi-Domain Framework. IT Convergence and Security (ICITCS), 2015 5th International Conference on, IEEE.
- Aleven, V., et al. (2010). Toward a framework for the analysis and design of educational games. Digital Game and Intelligent Toy Enhanced Learning (DIGTEL), 2010 Third IEEE International Conference on, IEEE.
- Amory, A. (2007). "Game object model version II: a theoretical framework for educational game development." Educational Technology Research and Development 55(1): 51-77.
- Annetta, L. A. (2010). "The 'I's' have it: A framework for serious educational game design." Review of General Psychology 14(2): 105.
- Bellotti, F., et al. (2013). "Assessment in and of serious games: an overview." Advances in Human-Computer Interaction 2013: 1.
- Chorianopoulos, K. and M. N. Giannakos (2014). "Design principles for serious video games in mathematics education: from theory to practice." International Journal of Serious Game 1(3): 51-59.
- Connolly, T., et al. (2009). "Towards the development of a games-based learning evaluation framework." Games-based learning advancements for multisensory human computer interfaces: Techniques and effective practices. Hershey PA: IGI Global.
- De Freitas, S. and M. Oliver (2006). "How can exploratory learning with games and simulations within the curriculum be most effectively evaluated?" Computers & Education 46(3): 249-264.
- De Lima, L. G. R., et al. (2015). Evaluation of the user experience and intrinsic motivation with educational and mainstream digital games. Proceedings of the Latin American Conference on Human Computer Interaction, ACM.
- Djelil, F., et al. (2014). Towards a learning game evaluation methodology in a training context: A literature review. European Conference on Games Based Learning, Academic Conferences International Limited.
- Dondi, C. and M. Moretti (2007). "A methodological proposal for learning games selection and quality assessment." British Journal of Educational Technology 38(3): 502-512.

Rabail Tahir and Alf Inge Wang

- El-Sattar, A. and H. K. Hussein (2016). Learning Islamic Principles with Serious Games. Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology, ACM.
- Engl, S. and L. E. Nacke (2013). "Contextual influences on mobile player experience—A game user experience model." *Entertainment Computing* 4(1): 83-91.
- Fu, F.-L., et al. (2009). "EGameFlow: A scale to measure learners' enjoyment of e-learning games." *Computers & Education* 52(1): 101-11.
- Giannakos, M. N. (2013). "Enjoy and learn with educational games: Examining factors affecting learning performance." *Computers & Education* 68: 429-439.
- Hsieh, H.-F. and S. E. Shannon (2005). "Three approaches to qualitative content analysis." *Qualitative health research* 15(9): 1277-1288.
- Ibrahim, R. and A. Jaafar (2009). Educational games (EG) design framework: combination of game design, pedagogy and content modeling. Electrical Engineering and Informatics, 2009. ICEEI'09. International Conference on, IEEE.
- Kececi, N. and A. Abran (2001). "ANALYZING, MEASURING & ASSESSING SOFTWARE QUALITY WITHIN A LOGIC-BASED GRAPHICAL FRAMEWORK." Université du Québec a Montreal.
- Kiili, K. (2005). "Digital game-based learning: Towards an experiential gaming model." *The Internet and higher education* 8(1): 13-24.
- Kiili, K., et al. (2014). "Flow framework for analyzing the quality of educational games." *Entertainment Computing* 5(4): 367-377.
- Liao, Y. H. and C.-Y. Shen (2012). Heuristic evaluation of digital game based learning: a case study. Digital Game and Intelligent Toy Enhanced Learning (DIGITEL), 2012 IEEE Fourth International Conference on, IEEE.
- Loh, C. S., et al. (2015). Serious games analytics: Theoretical framework. Serious games analytics, Springer: 3-29.
- Marciano, J. N., et al. (2014). "Evaluating multiple aspects of educational computer games: literature review and case study." *International Journal of Computer Games Technology* 2014: 14.
- Mei, S.-Y., et al. (2015). Interface Evaluation of a Game-Based Learning System: Experts vs. Novices. Advanced Applied Informatics (IIAI-AAI), 2015 IIAI 4th International Congress on, IEEE.
- Mitgutsch, K. and N. Alvarado (2012). Purposeful by design?: a serious game design assessment framework. Proceedings of the International Conference on the foundations of digital games, ACM.
- Mohamed, H., et al. (2012). Quantitative analysis in a heuristic evaluation for usability of educational computer game (UsaECG). Information Retrieval & Knowledge Management (CAMP), 2012 International Conference on, IEEE.
- Negovan, V. (2010). "Dimensions of students' psychosocial well-being and their measurement: Validation of a students' Psychosocial Well Being Inventory." *Europe's Journal of Psychology* 6(2): 85-104.
- Omar, H. M. and A. Jaafar (2010). "Conceptual framework for a heuristics based methodology for interface evaluation of educational games." *Computer and Information Science* 3(2): 211.
- Oprins, E., et al. (2015). "The game-based learning evaluation model (GEM): measuring the effectiveness of serious games using a standardised method." *International Journal of Technology Enhanced Learning* 7(4): 326-345.
- 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.
- Pappa, D. and L. Pannese (2010). "Effective design and evaluation of serious games: The case of the e-VITA project." *Knowledge Management, Information Systems, E-Learning, and Sustainability Research*: 225-237.
- Perttula, A., et al. (2017). "Flow experience in game based learning—a systematic literature review." *International Journal of Serious Games* 4(1).
- Petri, G. and C. G. von Wangenheim (2016). "How to Evaluate Educational Games: a Systematic." *Journal of Universal Computer Science* 22(7): 992-1021.
- Petri, G. and C. G. von Wangenheim (2017). "How games for computing education are evaluated? A systematic literature review." *Computers & Education* 107: 68-90.
- Pourabdollahian, B., et al. (2012). "Serious games in manufacturing education: Evaluation of learners' engagement." *Procedia Computer Science* 15: 256-265.
- Rêgo, M. B. B. A. B. and I. de Medeiros (2015). "HEEG: Heuristic Evaluation for Educational Games."
- Rooney, P. (2012). "A theoretical framework for serious game design: exploring pedagogy, play and fidelity and their implications for the design process."
- Shi, Y.-R. and J.-L. Shih (2015). "Game factors and game-based learning design model." *International Journal of Computer Games Technology* 2015: 11.
- Spiggle, S. (1994). "Analysis and interpretation of qualitative data in consumer research." *Journal of consumer research* 21(3): 491-503.
- Su, C.-H., et al. (2013). "Rough set theory based fuzzy TOPSIS on serious game design evaluation framework." *Mathematical Problems in Engineering* 2013.
- Tahir, R. and A. I. Wang (2017). State of the art in Game Based Learning: Dimensions for Evaluating Educational Games. European Conference on Games Based Learning, Academic Conferences International Limited.
- Tan, P.-H., et al. (2007). Adaptive digital game-based learning framework. Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts, ACM.
- Tseloudi, C. and T. Tsiatsos (2015). Panic in the gallery: An online educational game for art history: Design and evaluation of a matching game. Information, Intelligence, Systems and Applications (IISA), 2015 6th International Conference on, IEEE.
- Virvou, M. and G. Katsionis (2008). "On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE." *Computers & Education* 50(1): 154-178.
- Wallner, G. and S. Kriglstein (2011). Design and evaluation of the educational game DO Geometry: a case study. Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, ACM.

Rabail Tahir and Alf Inge Wang

- Wang, Y., et al. (2015). "An Evaluation Framework for Game-Based Learning."
- Yu, K.-C., et al. (2005). The implementation and evaluation of educational online gaming system. Information Technology: Research and Education, 2005. ITRE 2005. 3rd International Conference on, IEEE.
- Yue, W. S. and N. A. M. Zin (2009). Usability evaluation for history educational games. Proceedings of the 2nd international Conference on interaction Sciences: information Technology, Culture and Human, ACM.
- Zaibon, S. B. and N. Shiratuddin (2010). Heuristics evaluation strategy for mobile game-based learning. Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE), 2010 6th IEEE International Conference on, IEEE.
- Zhang, H.-f., et al. (2010). Research on the design and evaluation of educational games based on the RETAIN model. Knowledge Acquisition and Modeling (KAM), 2010 3rd International Symposium on, IEEE.
- Zhang, Y. and B. M. Wildemuth (2005). "Qualitative Analysis of Content by." Analysis 1(2): 1-12.