IMPACT OF DIGITAL GAME-BASED FORMATIVE ASSESSMENT ON LEARNING QUALITY AND EFFICIENCY: USING GAME-BASED APPLICATION TO ASSESS ENGLISH VOCABULARY IN PRIMARY LEVEL STUDENTS

By

Ramsha Asrar

NUST201464140MSEECS61414F

Supervisor

Ms. Farzana Ahmad

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Innovative Technologies in Education (MS ITE)

In

School of Electrical Engineering and Computer Science,

National University of Sciences and Technology (NUST),

Islamabad, Pakistan.

(August, 2018)

APPROVAL

It is certified that the contents and form of thesis entitled **Impact of Digital Game-Based Formative Assessment on Learning Quality and Efficiency: Using Game-based application to assess English Vocabulary in Primary Level Students** submitted by **Ramsha Asrar** (NUST201464140MSEECS61414F) have been found satisfactory for the requirement of the degree.

Supervisor: _Miss Farzana Ahmad

Signature: _____

Date: _____

Committee Member 1: Dr. Muhammad Mudassir Malik

Signature_____

Date: _____

Committee Member 2: Mr. Jaudat Mamoon

Signature _____

Date: _____

Committee Member 3: Miss Erum Afzal

Signature _____

Date: _____

DEDICATION

To Allah the Almighty

&

To my Parents and Faculty

CERTIFICATE OF ORIGINALITY

I hereby declare that the thesis titled "Impact of Digital Game-Based Formative Assessment on Learning Quality and Efficiency: Using Game-based application to assess English Vocabulary in Primary Level Students" is my own work and to the best of my knowledge it contains no materials previously published or written by any other person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at SEECS, NUST or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at SEECS, NUST or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Author Name: Ramsha Asrar

Signature: _____

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Ms. Farzana Ahmad, for supervising my work with patience and dedication. Her motivation encouraged me to finish the work. I am also thankful to my committee members, Sir Jaudat Mamoon, Dr. Mudassir Malik and Ms. Erum Afzal for their valuable suggestions.

Special shout-out to Faculty and students at Qazi Public School, Lahore for accommodating my research and being ever so welcome and enthusiastic and helping me in deploying this assessment in their school.

My family for their unconditional love, advice and support through all the tough times I faced during these four years and always and helping me become who I am today.

My friends and class-fellows for their help wherever needed as well as our various moments of laughter and adventures together. I cannot put to words what you all mean to me. Wish you all best of everything, please don't change, you precious souls.

Thanks to everyone who made my journey of this thesis exceptionally wonderful.

Abstract

In this study, we investigate the outcomes of incorporating a digital game-based learning formative assessment into regular educational curriculum and then evaluate its effects on students' learning quality and efficiency. The purpose of this study is to explore if there is a remarkable impact on students' learning quality and efficiency when digital game-based learning formative assessment were used in their classroom.

In Pakistan, summative assessment is practiced commonly which assess the learning of the student after a lesson or course ends. Therefore, traditional and conventional assessment methods fail to cater to the needs of students, as there is not enough feedback or proper guidance available to them. Teachers also find it hard to address to learning needs of each and every student. Teachers have the load of class administrative issues, which makes it extremely hard for the teacher to attend to every student in the class and make sure that every student has understood the concept being taught.

This research was carried out using digital game-based learning formative assessment, developed while keeping in mind the pedagogical aspects, principals of UDL and interactive elements. The application was developed in accordance with the National curriculum of Pakistan and covered 1st term Basics of Grammar. The app consisted of a quiz about Parts of speech, Vocabulary and Basic Composition. The research was carried out on students of grade 3 in a private school for lower middle class, located in Lahore. The Data samples were collected from two branches of same school, all located in same vicinity, one for girls other for boys.

To gather quantitative data through the tools of pre-test and post-

test, the study used an experimental research design. Their relevant corresponding teachers validated these tests. The total Number of Participants was 200, which were divided in two groups: control group and experiment group, each with a student set of 100 respectively. The participants in experiment group used a digital game-based learning formative assessment for English for a week. The scores obtained through pre-test and post-test were weighed up to ascertain if the indulgence of formative assessment had a statistically noteworthy influence on students' learning efficiency of both or any group. Data collected was coded into the SPSS software and statistical analyses were conducted.

In order to test the significant difference between the pre-test and post-test scores of both the groups, different tests were utilized. As the result of those analysis significant difference was found in the learning efficiency of the students. In the view of literature review and the statistical analysis it was found that lectures supported by digital game-based learning formative assessment not only increase student's understanding of those particular topics but make learning fun and enjoyable.

Table of Contents

Contents

List of Abbreviations

DGBL: Digital Game-Based Learning
HCI: Human Computer Interaction
UDL: Universal Design for Learning
MOOC: Massive Open Online Courses
MMPOG: Massive Multi-Player Online Games
CALL: Computer Assisted Language Learning
TELL: Technology Enhanced Language Learning

List of Figures

List of Tables

Chapter 1

Introduction

Background

Basic education is a fundamental human right in accordance with Article 26 of the Universal Declaration of Human Rights (1948), same is envisaged in constitution of Pakistan Article 25-A. This famous article speaks about the education of all individuals without applying any limitation of specific age group. All individuals including children, youth, and adults of any sex are entitled to this fundamental right. Despite all the laws there are approximately 57 Million illiterates in this country (Associated Press of Pakistan Corporation, 2017). This begs the question that what steps may be taken to attract more people to take admissions and to retain those who dropout. In this regard, studies and researches in education are being conducted every day to contribute towards the betterment and improvement of education scene in Pakistan.

One of the imperative principles in curriculum design is the need to achieve positive balance between intended learning outcomes, learning activities and assessment (Biggs, 1999). Assessment play a major role in any education setup as they are the true measure learning and understanding. Assessments using empirical data collection are conducted to understand student learning, teaching methodology and curriculum programs with an aim to augment them further (S. A. , 2004). Assessments conducted at state level range from kindergarten to universities are summative in nature to

audit the success of institutions and initiate reforms (Ruland, 2011). Formative assessment should be used during instruction to help students learn material initially and throughout the learning process. Summative assessments can be used at the end of a unit, chapter, quarter, or semester to assess and evaluate how much learning students have gained and retained. Formative assessments boost up student learning and understanding and are pivotal for future progress. Assessment serves a formative function when evidence about student achievement is elicited, interpreted and used by teachers to make decisions that result in better instructional set (Wiliam & Black, 2010).

With the growing day-to-day advancements in technology, it is evitable for every individual to have a good know how of using these gadgets, that too for our advantage. In this regard, introduction of basic digital tools in classrooms is eminent. Video games are now the fifth major pillar of literacy (Squire, 2008). Since then, use of digital forms in learning has become a norm. They are deployed extensively in the field of education, with an existing body of work examining the relation between games and education (Yang, 2010), (Chiang, 2011). With the growing expansion of technology, instructors and those who create educational policy are interested in introducing innovative technological tools, such as video games, virtual worlds, and Massive Multi-Player Online Games (MMPOGs) (Buckless, 2014), (Gómez, 2014). Especially for the teachings of language learning, digital game-based learning (DGBL), computer assisted language learning (CALL) and technology enhanced language learning (TELL) are some of the most commonly used practices.

Since English is considered a second language in Pakistan, therefore it becomes even difficult to retain student engagement and presence in the class. Foreign language educationists proclaim that students bring with them to the classroom a complicated set of "attitudes, experiences, expectations, beliefs, and learning strategies" (Benson, 2001). These "attitudes toward learning, and the perceptions and beliefs that determine them", might have a significant impact on students' learning performance (Bandura A. &., 1981) (Como, 1986) (Cotterall, 1995) (McCombs, 1984) and "learning outcomes" (Martin & Ramsden, 1987) (Van Rossum & Schenk, 1984) (Weinert & Kluwe, 1987). Thus, a language learner's belief might be contemplated as varying from person to person, a persuasion about language learning and how it might work. Creation of learner beliefs depends on achievement and anticipation (Akhtar, 2011).

Problem Statement

In Pakistan two main sectors for providing education are public sector, which is cheaper, but the quality of education is not up to par, and private sector, that is not affordable by average working class and is considered to provide a good quality standard of education. Consequently, the number of students enrolled in public sector institutions is far above that of private sector institutes.

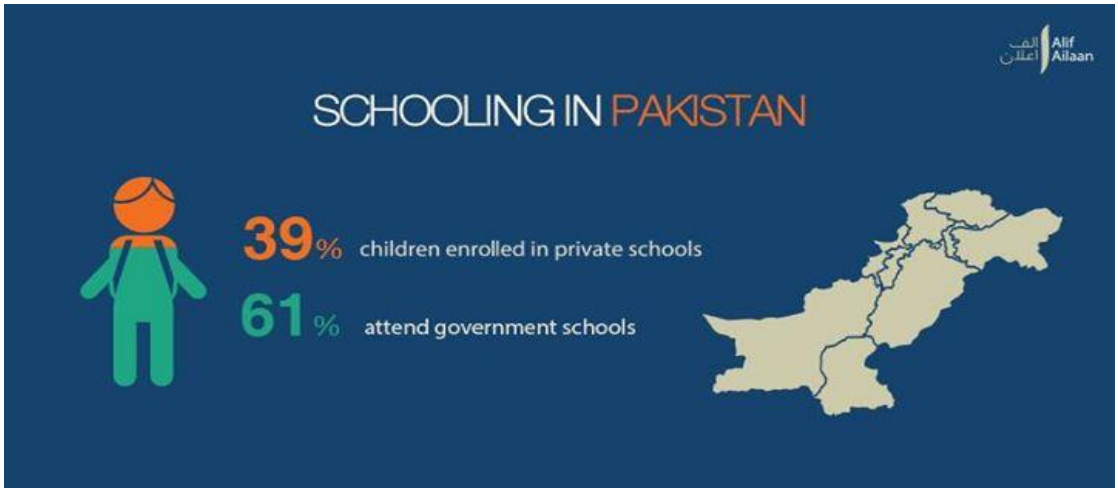


Figure 1 Percentage of enrolled students by school type

In Pakistan Summative assessment for primary education was done by National Education Assessment System (NEAS) which indicates students below mean for English reading (56%) and English writing (87%). While the overall English learning levels for public sector schools is illustrated by Alif Ailaan's report below:

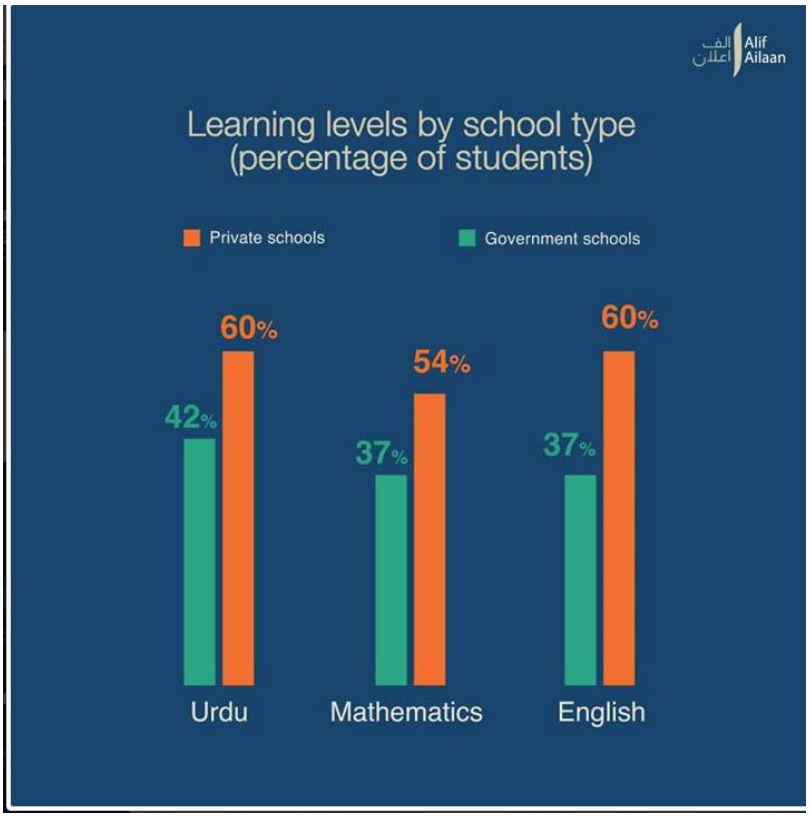


Figure 2 Learning level by school type

Our conventional learning environment does not let every child blossom as it restricts their learning styles, their curiosity and their enthusiasm. They are rarely allowed to ask for repetition of the content as the teacher must cover the syllabus within deadline. Similarly, our traditional learning classroom does not have a place for slow learner. This only promotes rote learning.

The only way to measure learning is through the summative assessments, which mentioned earlier do not provide adequate feedback neither do they measure if the student can apply the learning in a real-world scenario. Exam questions and format is repeated, rote learning is widely practiced, and it is difficult for the students to do basic comprehension and develop deep understanding (Rehmani, Impact of Public Examination System on Teaching and Learning in Pakistan, 2003). Formative assessments are the real deal

breaker to this problem. Formative assessments are defined as, activities undertaken by teachers—and by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities" (Wiliam & Black, 2010).

Government of Pakistan has announced its smart school e-learning system initiative in which tablets will be handed out to students, so they may have ease of access to all the notes, e-books as well as assessments. In order to do so they not only have to digitalize the present curriculum, but they must avail this opportunity to make digital game-based learning applications that cater the needs of the students and makes learning fun and enjoyable for them. An application that follows UDL principle, provide multiple means of engagement, representation and expression, addresses the needs of different learning styles of learners (CAST, 2014), is interactive, caters EFL issues and is formative in nature. Therefore, it is high time to address the issues mentioned above and conduct a study to measure the factors that may increase student retention, motivation and learning at the same time.

Study Overview

The purpose of this study was to measure the effects of digital game-based learning formative assessment on learning quality and efficiency. It aimed to encourage learners to acquire sense of ownership, engagement and self-assessment that motivates them whilst assessing them at the same time. For that purpose, an application was designed that will cover the contents of grammar of the curriculum of grade 3 and assess their knowledge and understanding through a formative interactive assessment.

The combination of right pedagogy and immediate effective

feedback formed the basis of this application. Students were able to view response to each wrong answer so that they can correct their decision learn from it and go towards the right answer.

The outcomes of this study will help to make recommendations to policy makers for including interactive learning applications as part of the curriculum for all public/government schools, to increase student retention rate and decrease the dropout rate in schools, to inspire other schools to adapt this methodology of teaching by showing students' academic performance and improved grades, to propose an effective pedagogy for making interactive learning applications for all subjects and grades, to propose other researchers to extend this study for longer period of time and then observe its effects on students learning.

Chapter 2

Literature Review

In Pakistan the major educational service provider is the public sector structured education system of twelve academic years. There are five different levels of education system in Pakistan. First level is the primary level that consists of grade one to grade five. Second level is the middle one that consists of grade 6 to grade 8, then comes secondary level that consists of grade nine and ten. After secondary level, intermediate level appears that includes grade eleven and twelve. After completion of intermediate level, university level appears with advanced courses.

The education system is supervised by the Ministry of Education of Pakistan Government. The federal government is involved in the development of the curriculum, authorization and funding of research projects, while the state government takes responsibility for academic foundations.

According to all reports from the year 2015, it has been analyzed that there are around 146,185 primary, 42,147 middle level and 29,874 secondary schools in Pakistan (Malik, Amin, & Ahmad, 2015).

Education for all (EFA) programs encompasses a wide range of educational dimensions focusing on the early attention of children in primary and secondary universal education. In Pakistan, primary enrollment rates are 93%, but dramatically drop to 37% when students reach secondary school. In urban areas, 37% of children and in rural areas 36% of kids are not prepared to go to school. Likewise, 11% of girls in rural areas and 10% in urban areas are not agreeable to go to school (Pakistan, 2013-2016).

Other causes of failure are low educational accomplishment of students in school and unappealing teaching approaches (Anwar, Tahir, & Batool, 2012).

A study done in Sindh, Pakistan in the year 2012 underlined a number of noteworthy reasons for students to leave school. The study that is in the teacher's perspective, the student drop-out rate is growing due to the fact that they are merely not concerned in gaining an education. Involvement gaps are found in different countries and cultures.

A lower socio-economic group and students from minority groups are leaving, but Williams noted that family history has partly influenced the education of students and that school rules and policies play an important role in the education of students (Willms, 2003).

Secondary education is an indispensable sub-sector of the whole education system. It serves not only as a moderate power provider for the economy, but also as a feeder for higher education levels (Anwar, Tahir, & Batool, 2012). When a student enters teenage

years, it is the most significant phase of life. Basic forms of perception and behavior toward life originate to take shape. Method and measure of individual performance is determined as a distinct person and as a civilian (Review, 2006). These students will participate in the national labor force and play a role in the national development of their homeland.

Those who do retain their education have to go through the process of assessments and exams in order to stay in school. Trainees and researchers examine the assessment methods used in the academic world for the qualification and quantitative assessment of learning, understanding and the effects of students' educational development. The research areas covered by students and trainers include student assessment, the educational environment, the role of ICT in education, the privatization of education, critical thinking, educational interventions, standardized tests, professional development of teachers and class interaction that continues to work towards discovery. Below, we offered a complete review of the work in the field of evaluation in general, and in particular on formative evaluation.

Black and William aim to respond to the articles related to the criteria, if the formative evaluation leads to an increase of the educational standards and the evolutionary evaluation itself can be improved (Wiliam & Black, 2010). More than 160 magazines and books have been read in the works and readers have prepared a long review.

They reported the magnitude of the effect as a measure of the average improvement in the scores of students in innovation as a

measure of the average scores of a typical group of students in the same test and showed that the size of the effect of the effectiveness test was understood between 0.4 and 0.4. and 0.7. Studies have emphasized that formative evaluation translates into equality in learning outcomes, particularly with help to students with low performance and learning difficulties. Because of continued evaluation, students who refused to take the school seriously established a better relationship. The development of feedback from students and teachers using the results of the evaluation of the active involvement of students in the process of learning, learning and the results of teaching, are ways in which evaluation can influence motivation and self-esteem of the students. Research studies suggest that the methods of class evaluation that exist in many educational settings are quantitative and lack guidance on healing, encouraging extraordinary learning from effective learning, over-emphasized signs, irresistible inspiration from collaborations and neglect of learning functions.

They argue that in their work, self-confidence and their own talents provide a clear focus on specific learning needs that can be improved by instinctively supporting self-esteem and successful cultures, particularly in classroom evaluation practices for results. downs. It is understood that there is no competition and comparison between the students. Self-assessment and peer review suggest a complementary formative assessment and are inevitable as successful innovation in the learning process. Self-assessment expertise aid pupils grow a greater understanding of learning objectives and at what point they are on the learning scale. It becomes more participative and effective learner by discussing student assessments with teachers and other students, starting a

thought process that will help them build concepts and solve inequalities. Peer review provides students with relevant feedback in their mother tongue and allows them to work in a group that is good in terms of collaboration and group discussion, giving them the opportunity to communicate their understanding of evolution to students. To influence teaching, classes and assignments should be given to students as a way to encourage feedback and encourage response and give them a coherent dialogue to guide their thoughts and beliefs.

Black and William claim that when entering such dialogue, teachers should offer enough time for students to ask, think and respond, provide suggestions to guide them, initiate discussions in small groups, and help them understand by example when answers are wrong. In this way, the dialogue begins with a formatting process that evokes a reflective image that all students should encourage participation. Feedback should be made about the strengths and weaknesses and opportunities should be provided to work on the deficiencies.

Jacoby, Heugh, Bax and Branford-White (Jacoby, 2014) examined the use of formative assessment in student participation by the use of the virtual learning environment developed using ICT. The study developed formative assessments with learning modules, each with 10 multiple-choice questions and 60 questions in total. Formal assessments were planned for an unlimited period of time and for an unlimited number of initiatives so that students had sufficient time to finish the drills.

To test the study, the student tried to examine the basic text and the material related to the exams, and the goal was to develop the students' learning skills by obtaining and researching the answers of the questions. In addition to this evaluation, he included the opportunity to motivate the answers by taking comments and thoughts on their answers. The comments were classified as positive, negative and neutral with respect to the formative evaluation. After each evaluation test, students were given immediate descriptive feedback including appropriate notes and remedies. The data showed an increase in the formative evaluation and the highest grades with consequent participation of the students in the learning process.

This may be related to the fact that students have enough time to investigate and answer questions, perform more than one evaluation test, and not only with notes, but also with more guidance and instant feedback to improve learning. In their study, they discovered that virtual learning processes designed using the formative assessment framework improved the learning experience of students by improving their comprehension skills and incorporating them into the learning process, causing in improved scores of pupils.

Rehmani focused on research on student learning assessment to improve their learning through feedback (Rehmani, Changing assessment practices in Pakistani schools: A case of AKU-EB middle school assessment framework, 2012). He explained how the change in assessment at secondary school level could be started and what the implementation challenges were. The study provided the use of project work portfolio evaluation as an alternative to traditional evaluation based on brands and notes in Pakistan.

Portfolio assessment is expected to self-evaluate, stimulate students to absorb and develop collaborative learning. Teachers can support them understand how to evaluate and provide feedback and response to increase student learning.

Activities designed and developed for the institution's board of directors, for students with a central question in terms of learning objectives and content-based goals. Separate and group activities are planned to aggravate serious thinking, problem solving skills, administrative skills and principled mindfulness. The teachers in charge provided the professional development necessary to carry out the tasks and the evaluation. Students were asked to do self-assessment and peer review using the worksheet provided by the board of directors. Each student presented a note of reflection on the work of the project. Descriptive notes and explanations and descriptive feedback are used as a support assessment essence to address learning difficulties and improve student learning. The Board anticipates that such an attractive learning process will increase students' confidence in self-sufficiency, independence and success in life.

The examination and evaluation system dominated by Pakistan's primary education system was studied by Shamaas (Khattak, Assessment in Schools in Pakistan, 2012). The research indicates that internal examinations and evaluations of the educational institution do not positively influence student learning due to teachers who have not received the necessary training for assessment. Furthermore, evaluations are summative and students are held every twelve months to promote the next class or keep them updated.

These ratings are subjective and do not provide sufficient feedback to students for development. Self-assessment and peer review are not performed and, therefore, the student is not aware of the learning objectives and the place where he is facing the learning curve. The design of the evaluations does not take into account the students and the learning outcomes, but instead encourages the learning of the memorization and memorization of the manuscripts. In this context, the author suggests the use of multiple assessment tools to balance the formative assessment approach and the summative approach. This will ensure that the cognitive domain is included in the learning process and will improve the student's performance and understanding.

Nobert, Caroline, Yishay and Harvey describe the role of technology in learning applications that use formative evaluation (Norbert Pachler C. D., 2010). The study presents five case studies of a group of students using projectors, audio files, mobile phones, web tools and a series of comparators to help present and organize students' thoughts. Students / teachers encourage individual reflection, provide teachers with graphical feedback on students' response frequency and give systematic feedback on writing on multiple lines, respectively. Without this work, they constitute an important part of technology in formative assessment when the student opens the thought window for teachers. Students have the capability to characterize their thoughts and plans in various ways and to be more inventive and advanced.

The technological intervention stimulated the mental processes and allowed them to create otherwise unbearable things. Teachers have

also been able to provide better and more instant feedback that is more effective, empathetic and personalized than daily written feedback practices. Consistent with Black and William, they thought that pedagogical change was a fundamental goal of formative evaluation. The unexpected double situation cooperates with students who are encouraged to enter the learning process and teachers have responded to their learning using pedagogical knowledge. The authors concluded that training e-assessments are a series of processes that involve technological and social resources to engage students and teachers in order to improve learning and learning. Learning not to use technology alone, but how students and teachers use it to model formative assessment makes it a central method that needs to be enhanced as part of education advancement.

Normally, an experimental design is applied so that DGBL learning outcomes' impact be evaluated by comparing and / or interfering with a game-based approach to another type of education. Game-based approach may be compared to different interventions, indicating that the results will depend on the resultant differentiation (Bleumer, et al., 2012). According to Campbell, it is a preliminary post-test approach that has both control and experimental groups, to determine if learning has taken place (Campbell, 1963) .

The questionnaires are used to measure the motivations of the participants and their commitment to participate in the intervention in order to evaluate the motivational aspects of the DGBL (Hainey, 2010). Surveys are used to evaluate other effective results as well such as attitudes. Furthermore, researches have defined the evaluation of the game - secret evaluation - a technique aimed at

measuring the player's progress in a precise and dynamic way (Shute V. , 2011).

Depending on the learning objectives rooted in digital games DGBL efficacy studies can be famously categorized in three types (Bleumer, et al., 2012). In particular, digital-games often focus at the transmission of information (cognitive learning outcomes), the accession of skills (learning outcomes based on skill) or the attitude and change of behavior (results of emotional learning). Games for information transfer are generally applied in education. For instance, some research has found a significant effect of using digital games for mathematical teaching (Bai & et al., 2012) and language learning (Yip & Kwan, 2006). Digital games intended to acquiring skills are usually applied in an educational and institutional context. Many studies have demonstrated the effectiveness of game skills in the managerial skills implementation (Corsi & et al., 2006). Similarly Behavioral changes games are generally applied in the medical scenario. For instance health games that affect children's nutrition and physical activity (Baranowski & et al., 2008). To raise awareness on a particular topic, such as poverty, games for attitudinal or behavioral changes are applied (Neys & et al., 2012). (Kraiger, Ford, & Salas E, 1993) require different learning outcomes, different types of assessment. The integration of studies for the different learning outcomes will produce more heterogeneous results depending on the outcome of the evaluated result. For this reason, we will focus on this type of learning outcomes, ie the results of cognitive learning.

Thiru, Weena and Allan assessed the use of self-assessment exams presented as formative evaluations of student performance

(Aravinthan & L., 2011). During the course, a total of 14 studies were designed using the qualifications of the formative evaluation and repeatedly published in the students. Unlimited studies were carried out for the students and the scores were not counted for the final grade. The exam assignments were designed to include basic concepts and were both qualitative and quantitative, providing immediate opinion to the learners.

In the case of a false answer, students were given precise feedback and delivered detailed feedback with examples of understanding and improvement of basic learning concepts if students repeatedly did not respond to the questions of a particular field of study. Teachers had the opportunity to remotely monitor student progress and create statistics on student participation and learning outcomes on the topic. This delivered the coaches the chance to emphasis on the instructions and topics they need for additional improvement and support to bring students to the highest level of basic concepts. Students have taken online educational exams to be an effective learning tool able to develop their knowledge, understanding, critical thinking and comprehension skills. An association between the performance of the students and their overall performance was very strong and they were able to make more positive evaluations than those that did not. The results of the work by Ron, Jillian and Angela (Ron Fisher, 2011) show that a voluntary intervention designed using the influential evaluation features is beneficial for the results of student learning. The research was conducted on recent graduate students in the degree program and the voluntary research methodology and scientific writing i.e. review of the literature of 5 articles was proposed. The intervention is designed to be of a training nature to provide students with regular attendance activities

to address and respond to problems, limitations of words, deadlines and staff intervention.

Chapter 3

Methodology

This chapter discusses the methodology used for data collection and analysis in a learning environment. It gives an overview of the application, describes usability testing, research design, research question, research setting, participants, constraints, ethical issues and data gathering.

Project Overview

The study was designed to evaluate the effect that digital game-based learning has on learning quality and efficiency in a conventional setting using formative assessments. For that purpose, an extensive literature review was carried out to find the correct methodology, relevant techniques and data science involved.

Study Design

The study design was divided into two phases: the first phase consisted of designing and developing an application that used DGBL principles to carry out formative assessment along with its usability testing while the second phase included designing of the instruments or tools that will help measure the effects this application has on learning quality and efficiency. For that purpose UbD framework was utilized. UbD framework was first proposed

by Jay McTighe and Grant Wiggins in “Understanding by Design” (1998). Understanding by Design is fairly dependent on "backward design" as Wiggins and Mctighe like to call it (also known as "backwards planning"). Teachers, according to UbD proponents, traditionally start curriculum planning with activities and textbooks instead of identifying classroom learning goals and planning towards that goal. In backward design, the teacher starts with classroom outcomes and then plans the curriculum, choosing activities and materials that help determine student ability and foster student learning (Darling-Hammond, 2012).

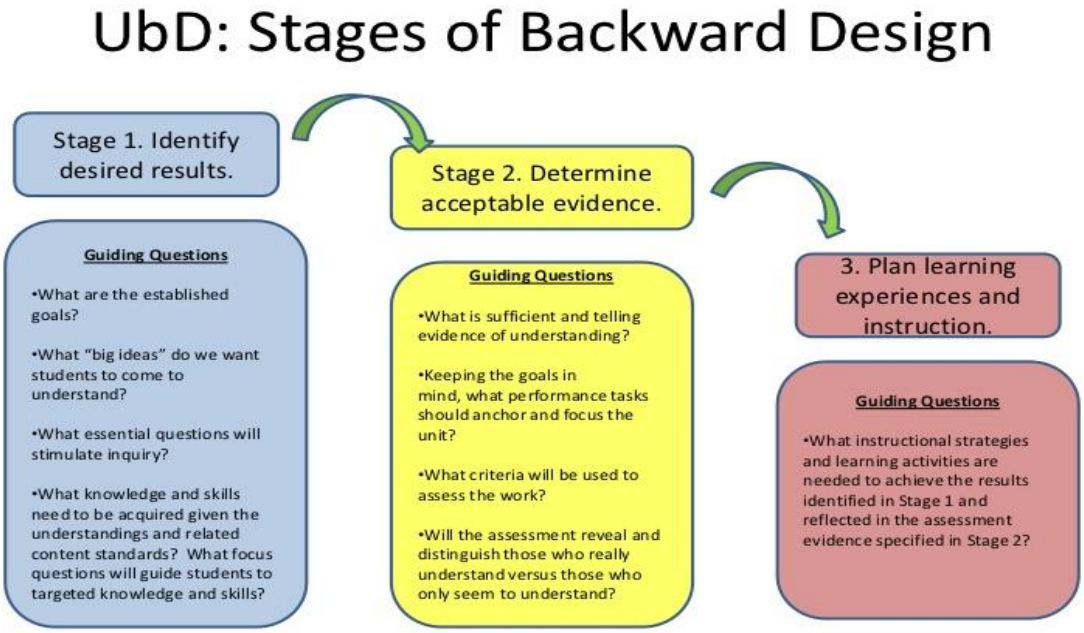


Figure 3 Stages of UbD backward design

The application of UbD stages on our study are as follows:

Stage 1: Identify desired results.

The result desired from this study was to measure whether the digital game-based learning formative assessments have an effect on students' learning quality and efficiency or not.

Stage 2: Determine acceptable evidence.

The acceptable evidence would have to be a conclusive assessment, so that it may be statistically analyzed to check the effect of digital game-based learning formative assessment on learning quality and efficiency.

Stage 3: Plan learning experience and instructions.

As the title suggested there will have to be digital game-based learning, therefore an application will be designed and to measure the effect of formative assessment, an assessment will be taken afterwards. The details are mentioned in research design.

Sampling and Participants

The purpose of the research was to measure the effect of digital game-based learning formative assessments on leaning quality and efficiency. The research was to be carried out on Grade 3 students that studied the English Book 3 published by Punjab Textbook Board under the National curriculum of Pakistan. For that purpose, Qazi Public School, Ferozepur Road, Lahore was chosen. A total of 200 participants were chosen enrolled in Boys and Girls branches of the school.

Topic Selection for DGBL formative assessment

In EFL classrooms, one of the more difficult topics to cover is the grammar section. Haudeck (1996) has reported that many learners have difficulty in internalizing grammar rules, although these have been taught intensively (European Commission, 2006). So much so that students not only hesitate to study it but also begin to avoid using it in front of the teacher. The English teacher is often portrayed as an "unattractive grammar monger whose only pleasure in life is to point out the faults of others" (Baron, 1982). This made us chose Grammar and Composition of Unit 1 of the English book 3.

Overview of the Application and Targeted Outcomes

According to (Reinders, 2012) employing technology, such as interactive games in education, allows teachers to solidify the connection between learning that occurs inside the classroom with that which occurs outside the school environment. This view is in line with sociocultural and interactionist perspectives on language learning (Khatibi & Cowie, 2013). Three types of effectiveness studies in DGBL can be distinguished based on learning goals embedded in digital games (Bleumer, et al., 2012).

DGBL is a paradigm which utilizes the game as a medium for conveying the learning contents, it is all about leveraging the power of computer games to captivate and engage end-users for a specific purpose, such as to develop new knowledge and skills (Corti, 2006). Specifically, digital games can aim at either transmission of knowledge (cognitive learning results), acquiring skills (skill-based learning outcomes) or attitudinal and behavioral change (affective learning outcomes). Typically those games are incorporated in

education that aim at knowledge transfer. For example, some studies show that digital games have a significant impact to teach mathematics (Bai & et al., 2012) and language learning (Yip & Kwan, 2006). Digital games aimed at skill acquisition are typically implemented in a training and corporate context. Several studies have observed an impact of playing games to practice managerial skills (Corsi & et al., 2006). Games aimed at behavioral change are typically implemented in the health sector. An example of this are the healthy eating games influencing the diet and physical activity of children (Baranowski & et al., 2008). Games aimed at attitudinal or behavioral change are implemented to raise awareness on a certain topic, such as poverty (Neys & et al., 2012). Our will be a DGBL targeted at knowledge transfer.

Formative assessment is concerned with how judgments about the quality of student responses (performances, pieces, or works) can be used to shape and improve the student's competence by short-circuiting the randomness and inefficiency of trial-and-error learning (Royce, 1989). Previous studies have suggested that providing online formative assessments undoubtedly improves e-learning motivation and effectiveness (Gardner, Sheridan, & White, 2002) (Henly, 2003). Feedback is an important and crucial part of formative assessments, as it is the force the drives one to re-evaluate their mistakes and improving the understanding using trial and error whilst building one’s own knowledge. Based on timing, feedback is divided into immediate feedback, instant feedback (after completing the assessment) and delayed feedback (feedback messages after a few minutes or longer) (Shute V. J., 2008).

In our work we have made a digital game-based learning English grammar formative assessment application with the purpose to enhance students' learning quality with the content and motivate them to progress further in their learning quest rather than assessing them in terms of fail or pass grade. The educational objective of such application is to help students clear their concepts and to reinforce and integrate part of the knowledge learnt in expository lecture (Thomas & Hooper, 1991).

According to Burr, assessments should be designed in a way that they meet the intended learning outcomes (S. B. , 2009). The purpose of developing this formative assessment quiz application was to move students' learning forward by providing timely feedback in the form of additional information and to engage them with the content by adding game elements. The students' learning outcomes covered in the application, according to the National Curriculum are given below:

Vocabulary:

Oxford dictionary defines vocabulary as "the body of words used in a particular language, a range of artistic or stylistic forms, techniques, or movements". (Yip & Kwan, 2006) in their study entitled "Online vocabulary games as a tool for teaching and learning English vocabulary" state that learners playing online vocabulary games tend to learn more appropriately and could retain the new words for a longer period and retrieve more words compared to those who aren't provided with vocabulary games. Basic Vocabulary words from the grade 3 English book were included and students should be able to remember and retain them.

Parts of speech:

Parts of speech are defined by Merriam Webster as “a traditional class of words (such as adjectives, adverbs, nouns, and verbs) distinguished according to the kind of idea denoted and the function performed in a sentence.”. Participants will be able to identify correct part of speech from a sentence and based on the recognition will be able to correctly apply them.

Composition:

Merriam Webster defines composition as “the way in which something is put together or arranged; the combination of parts or elements that make up something; a piece of writing; especially; a brief essay written as a school assignment”. The level of composition in the curriculum was that the student should be able to form short sentences or be able to recognize and arrange sentences in correct structure.

Other targeted outcomes of the application include but are not limited to:

- Timely feedback
- Self-assessment
- Increased motivation
- Fun learning environment

Application Design

The application was designed in Unity 3d, which is a game development platform so that it may be compatible with Android OS. It was designed for grade 3 students for English subject

according to the National curriculum being followed i.e. English textbook 3 published by Punjab Textbook Board. In particular the Unit 1 grammar section was chosen. This section was further divided into three parts, namely: parts of speech, vocabulary and composition.

To provide ease of access to participants of all learning styles; graphics, text, audio and visual aid was used. The number of visual learners is more as compared to any other type of learners (Riggs & Gholar, 2009) that is why interactive visuals and images are used in this application. Instructional design was mapped to student learning outcomes. At the end of game play with the application every student should be able to answer questions related to these grammatical concepts. Depth of these concepts covered is the same as in the English book of grade 3 students.

Main or starting screen:

The main screen of the application is shown below:

Figure 4 Screen Shot of DGBL Main Screen

This screen lets the user choose their desired field of grammar to take the quiz. These buttons take the user directly to the relevant quiz of their choice. The exit button at the bottom right let's the user exit the game. Colorful and bright theme was chosen to attract children. As the age group to be catered lied between 8-10 years old. The fonts were also chosen on purpose to make the children feel the fun and feel less formal or rather casual.

Parts of Speech:

If the user selects Parts of speech, then the next screen appears as shown in figure below. Four baskets can be seen at the bottom of the screen each titled with verb, noun, pronoun and adjective. These were the only four parts of speech that were to be taught to the students in the first chapter.

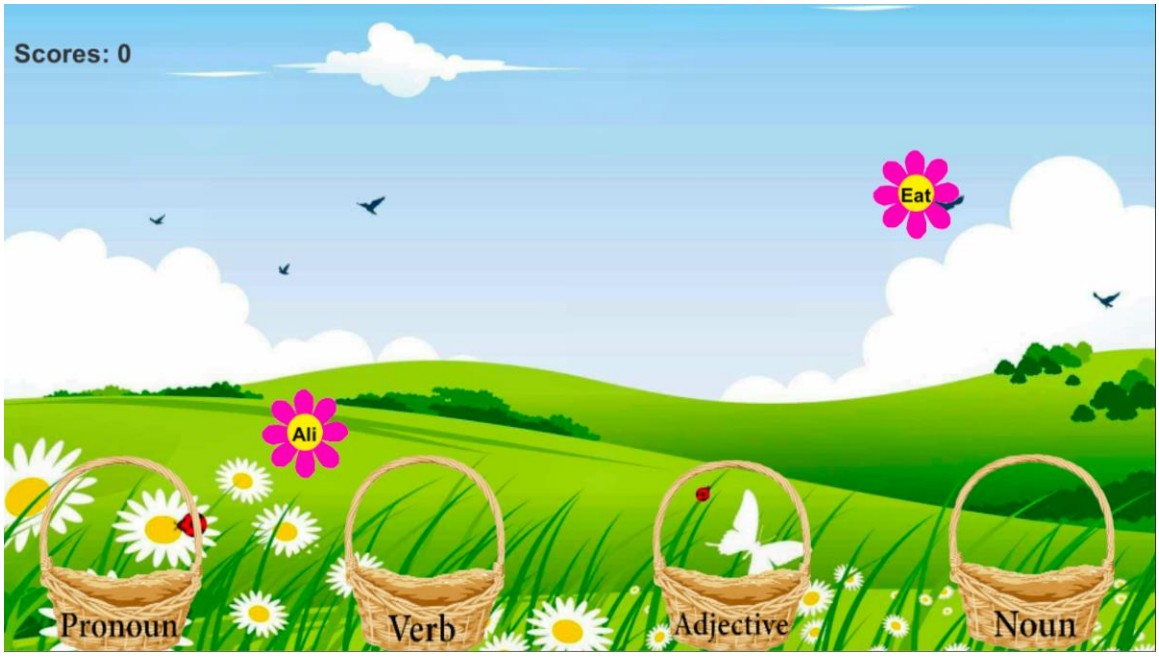


Figure 5 Screen Shot 1 of Parts of Speech

Multicolored flowers start falling downwards; the flowers actually have words written on it. The user must identify the part of speech that the flower or the word in the flower represents and have to match it with the right basket right part of speech. For instance, in the figure above flower on right has Eat written on eat, eat is a task, hence it is a verb; so, the student must drag the eat flower and drop it in the verb basket. The students must collect maximum number of correct flowers in the basket. The count is kept via scoreboard as in the figure above on the top left corner. It can be seen the score is 0 as the game has not started yet but in figure below the score is 30 as the user must have guessed three parts of speech correctly and drag dropped it in the corresponding basket.

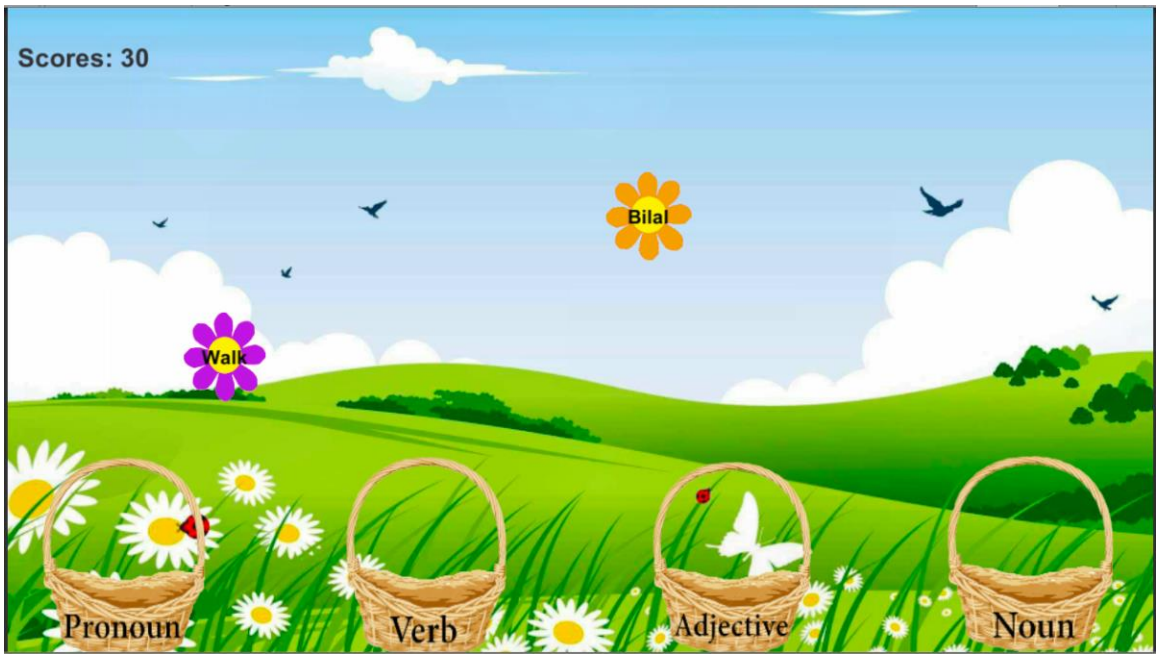


Figure 6 Screen Shot 2 of Parts of Speech

If the user cannot identify the flower in its falling down duration, then the flower escapes the screen, if ten flowers leave the screen unidentified, the game eventually stops. If the user keeps on identifying correct part of speech, the flowers keep coming but their speed of falling keeps on increasing and they start falling with high speed.

Composition:

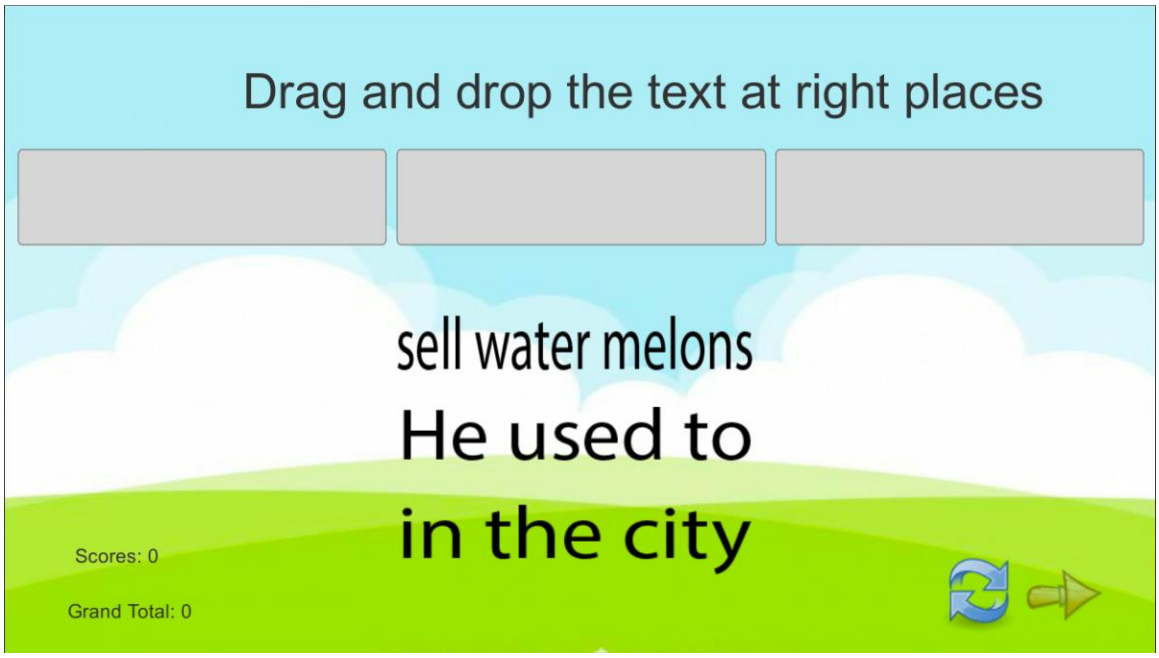


Figure 7 Screen Shot 1 of Composition

Similarly, if a user selects to play the composition section of the DGBL formative assessment application, the user will be asked to put together the phrases in correct order in order to form a sentence or rather "compose" a sentence. If the user drag drops a phrase in the right space it turns green otherwise it lightens up in red color. Hence providing timely, on the spot feedback and providing the user with the chance to correct their mistake.

Figure 8 Screen Shot 2 of Composition

Vocabulary:

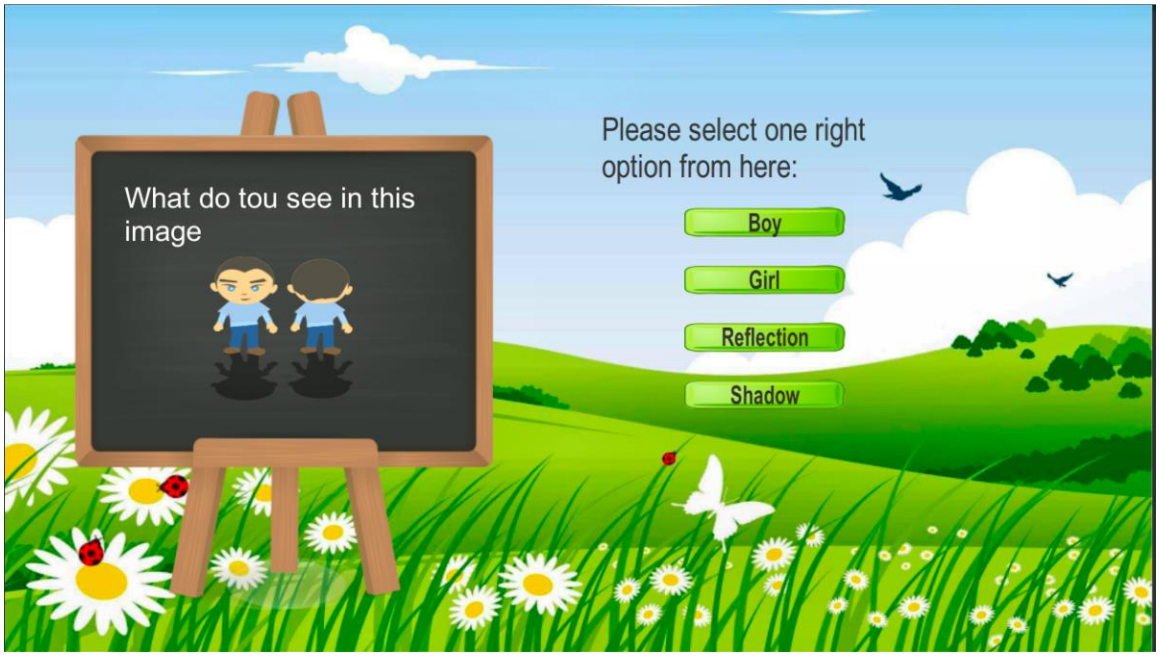


Figure 9 Screen Shot 1 of Vocabulary

When a game ends the user is taken back to the main screen. If the user then chooses vocabulary from there, then the user is taken to the vocabulary assessment and the following screens start to appear. These are simple Multiple-choice questions. The user is asked to identify the correct vocabulary word for the image displayed on the left. The vocabulary list is taken from the English textbook 3.

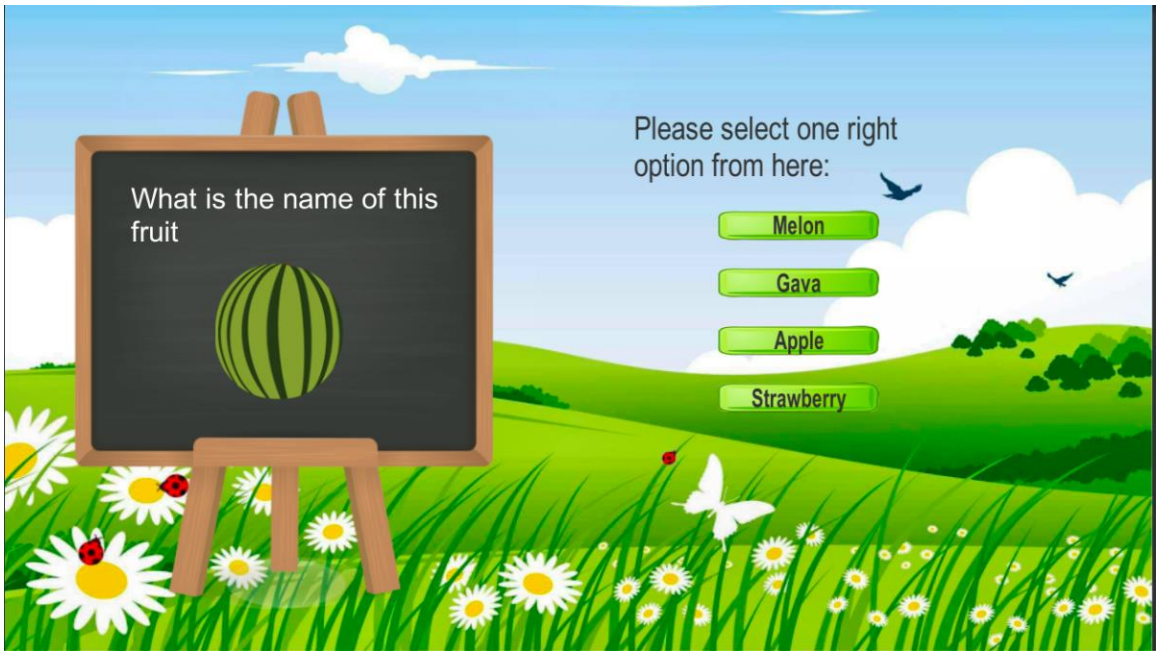


Figure 10 Screen Shot 2 of Vocabulary

Moreover, when a student selects a wrong answer, a pop-up appears that informs and motivates the students to try again and choose the correct option. The pop-up dialogue box enables the student to re-think the answer. This is an example of immediate feedback in formative assessment. An example of feedback is given below:

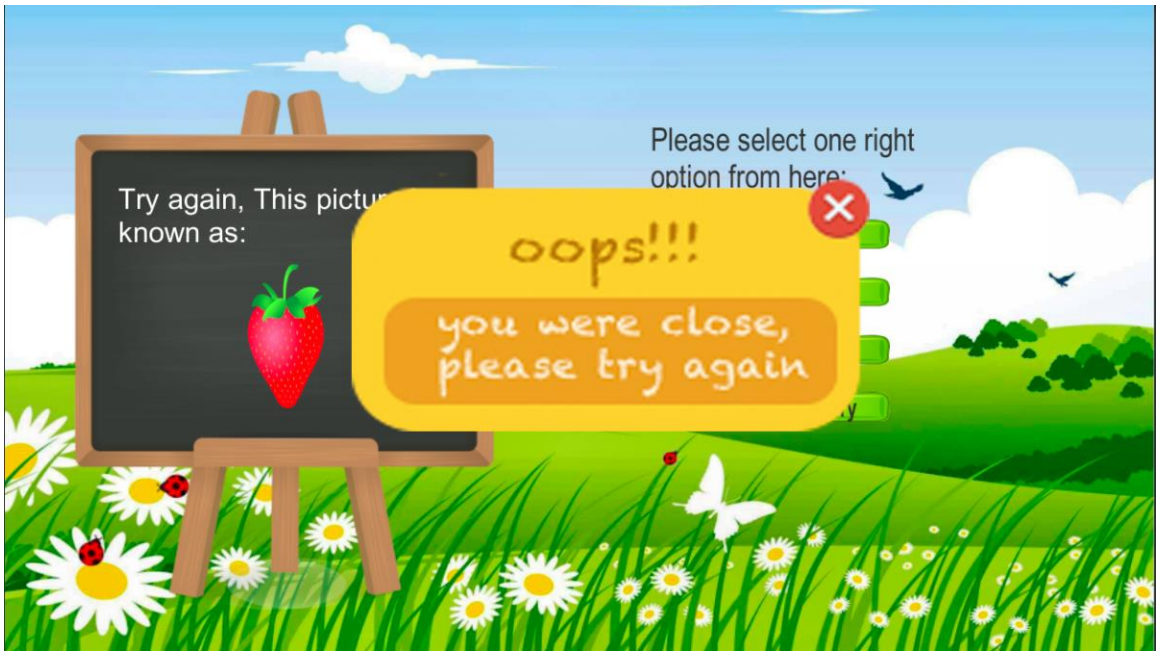


Figure 11 Screen shot 1 of appropriate feedback

The last screen of the application displays total score:

Figure 12 Screen shot 2 of appropriate feedback

Usability testing of the application

Usability testing is a way by which actual users trying a product are observed and information is collected about the ease and difficulty of that product (Redish, 1999). The formative assessment application was developed with usability in mind, tested iteratively and was improved based on usability results.

The purpose of evaluating the application for testing was to identify the usability problems that were experienced by students. The testing comprised of 15 third grade students, including boys and girls.

	Attending Government School	Attending Private School	Total
Girls	1	5	6
Boys	3	6	9
	4	11	15

Table 1 Demographics of usability testing participants

The usability assessment methods used were: observation, user feedback and questionnaire (Nielsen, 1993). The participants were observed silently against a usability checklist (see Appendix A).
After using the application, each participant was given a questionnaire regarding quiz application (see Appendix B). The questions were about:

- satisfaction with application
- clarity of the content
- ease of use

- ease of drag and drop
- understandability of the content
- icon clarity

Likert scale was used to collect their responses about the design and functionality of application.

Results of Usability testing

The information about the efficiency of quiz application was collected on a Likert scale. The usability problems found in the application are shown in table:

Sr. No.	Usability Problem
a.	There was no encouraging feedback at wrong answer.
b.	On achievement of high score, motivation lacked.

Table 2 Usability Problems

Screenshots of the application before and after usability testing

a. If a user chose a wrong option, the user was asked to try again but it lacked motivation. So, a pop-up dialogue box was inserted to encourage user.

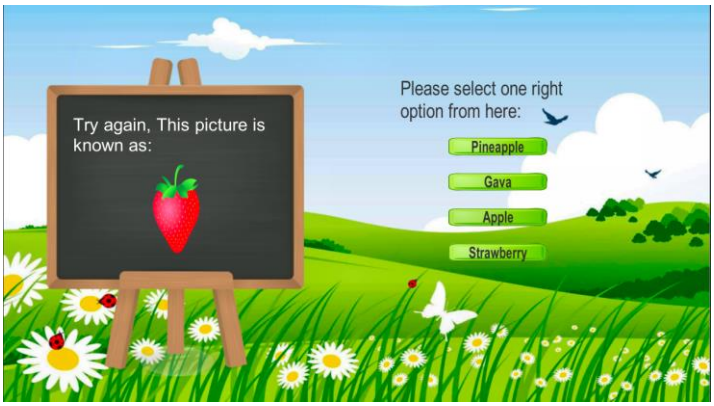
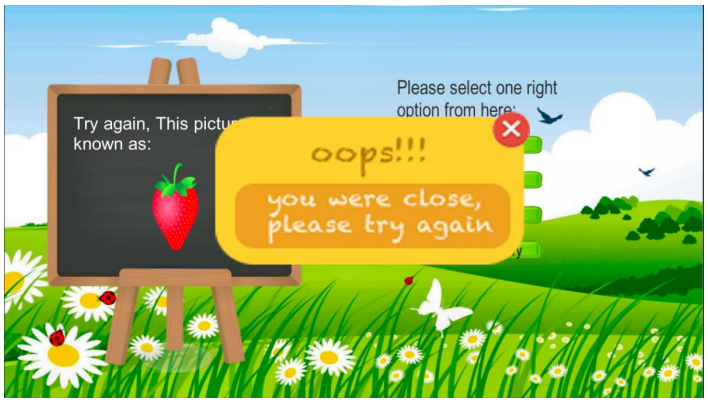
Before	After
	

Figure 13 Usability Testing screenshots 1 before and after

b- If a user attained highest score in the assessment, than congratulatory message did not appear which was fixed to increase motivation.

Before	After
	

Figure 14 Usability Testing screenshots 2 before and after

Research Methodology

The study was carried out to determine if formative assessment had an impact on learning quality of grade three students, after utilizing an interactive English quiz application. The intervention took 5 days and continued two weeks in all for both control and experiment group. Further explained in this chapter are following topics: (a) research design (b) research question directing this study (c) research site and sample (d) research constraints (e) ethical concerns (f) instruments used to collect the data (g) procedures for data analysis.

Research Design

The research design for identified for this research was experimental. According to Arikunto, Experimental study is a study which aimed to know there is or not the effect of the variable studied (Arikunto, 2010). In line with this Ary says that experimental design refers to the conceptual framework within which the experiment is conducted. The most important criteria that is the design be appropriate for testing the particular hypothesis of the study (Ary, Jacobs, & Razavich, 1985).

As mentioned above, there were a total of 200 students selected to participate in the intervention. The students were chosen by random sampling and divided in two groups control group (n = 100) and experiment group (n = 100). In random sampling each individual has an equal probability of being selected (Creswell, 2013). The research was conducted as follows:

1. Participants were tested before the experimental manipulation through pre-test.

2. The control group learned by traditional method from textbook while experiment group was subjected to experimental manipulation and were exposed to a treatment in which they were given android tablets to play and learn the digital game-based learning quiz application.
3. Participants were tested after the manipulation to see what changes occurred.

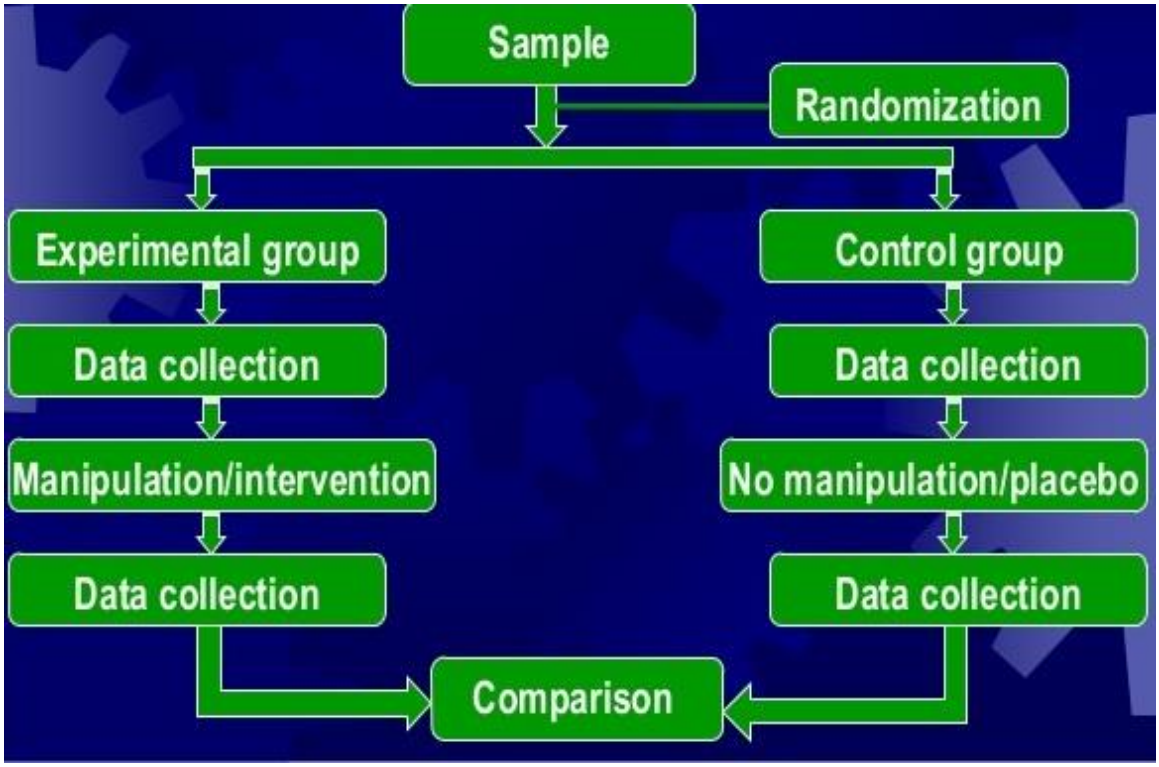


Figure 15 Experimental Research Design

The intervention took place for 5 days in four weeks. The data was collected by following ways:

Tools	Type of data
Pre-Test	Quantitative
Post-Test	Quantitative

Table 3 Data Collection tools

The quantitative data collected through the questionnaires was coded in the SPSS software and then analyzed.

Research Question

The purpose of this study was to determine the effectiveness of formative assessment on learning quality and efficiency, using digital game-based learning English quiz application. The study addressed the following research question:

What is the effect of digital game-based learning formative assessment on learning quality and efficiency?

Research Site and Sample

The sample used for this study consisted of grade 3 students. A sample is a smaller and manageable version of a large data set. Samples are used in statistical data analysis when the observation population size is too large (Investopedia, 2016). The study was conducted at boys and girls branch of Qazi Public school during the academic year 2016. The schools were chosen for convenience sampling due to ease of access. Two sections of the participants were chosen randomly as control (n = 100) and experiment group (n = 100). The gender distribution included 100 boys and 100 females.

Total number of participants was 200. The particulars of participants are given in the table below:

Gender	Group	No. of participants	School
Male	Control	50	Qazi Public School, Ferozepur Road, Boys Branch, Lahore
	Experiment	50	
	Sub-total	100	
Female	Control	50	Qazi Public School, Ferozepur Road, Girls Branch, Lahore
	Experiment	50	
	Sub-total	100	
Total		**200**	

Table 4 Demographics of Participants

Control and Experiment Group

Control Group

The control group included 50 male participants and 50 female participants. Research was carried out separately for boys and girls in separate branches of same school, located in the same vicinity. In both schools, two sections of third grade students were randomly selected as control and experiment group. The total number of control group participants was 100.

Participants in the control group were taught first unit of English Grammar by textbooks through conventional method as teachers normally instruct in schools. After teaching the first unit of English

textbook 3, post-test was conducted to measure their learning. There were no instructional changes nor were the participants provided with any treatment. The scores were recorded before and after the instruction, through pre-test and post-test respectively.

Experiment Group

The experiment group consisted of 50 males and 50 females. Research was conducted separately for boys and girls in separate branches of same school, located in the same vicinity In both schools, two sections of third grade students were randomly selected as control and experiment group. The number of participants in experiment group was 100.

These 100 participants took a pre-test. In the next step, Android tablets were given to the experiment group participants for using digital game-based learning formative assessment for English Grammar quiz. Afterwards the post-test was conducted of the same participants.

There were otherwise no major differences between control and experiment groups.

Research Constraints

There was limited time for the development of digital game-based formative assessment application. The development not merely included coding but also graphics of the application. As this application has localized content designed specifically for the students of Pakistan, which means that new scenarios and story boarding had to be done as any other material from foreign books could not be taken.

Ethical Issues

Convenience sampling chose the schools chosen, as they were easy to reach. The branch heads of both the schools were explained about the intervention and a suitable time was mutually agreed. They signed the consent letter in this regard. Also, students were asked whether they were willing for the intervention or not and those who participated were volunteers. Students will be highly encouraged to fill all the questionnaires with full honesty though they will not be forced to do that (Cohen, Manion, & Morrison, 2007). Students will have to fill questionnaires before and after the intervention and extra time for focused group discussions with the students and teachers was already informed. It was promised that results and data collected will be kept anonymous and students name will not be shared. No personal information from participants was collected. This research will propose solutions to improve school's current situation. At any time, school authority is free to withdraw themselves from this study.

Data Collection

In order to collect data about the impact of formative assessment on students' learning quality, control and experiment group were monitored before the study through pre-test and at the end of the study by post-test. The data was recorded using numerical values in Microsoft Excel Spreadsheet. The excel sheet was then used in SPSS 20. 3 days were allocated for data collection. Two days were allocated for students to fill questionnaires. One day to have focused group discussion with students. In this time "completeness", "accuracy" and "uniformity" (ibid) was to be ensured and necessary measures were to be taken to ensure them. The responses collected

from the questionnaire were further triangulated by focus group discussion with the students. The tools used for data collection are discussed below:

Tools for Data Collection

Pre-Test and Post-Test

To determine the effectiveness of digital game-based learning formative assessment, pre-test and post-test were used to collect the data before and after the intervention. Both tests comprised of 8 questions with subparts from the first unit of English textbook 3, grammar part.

The pre and post-test were designed using Jeri L. Little & Elizabeth Ligon Bjork's approach of designing assessments. Multiple-choice pretesting was better for the learning of such related information than was cued-recall pretesting or a pre-fact-study control condition (Jerri & Ligon, 2016). Cued recall is a recall in which subjects are given hints (cues) at the time of recall. The cues are supposed to help the subject recall the memorized items (Dewey, 2007). For instance, questions based on true/false responses or Likert scale answers do not require to think hard as cues are provided in questions for recalls.

To validate the pre-test and post-test, two English teachers were asked to fill a questionnaire; One from boys' branch and one from girls, branch, that confirmed that the tests were understandable to the third-grade students and were made according to the National Curriculum.

Chapter 4

Results

Introduction

To determine the statistical significance of digital game-based learning application, the data was coded and analyzed using IBM's SPSS 20 (Statistical Package for the Social Sciences) statistical software. Statistical significance means that results from a research or experiment are not likely to occur randomly or by chance and instead have a dependent variable that influences them (Investopedia, 2016).

Firstly, in SPSS the data of both control and experiment groups were analyzed for normality to find out whether the data set holds the normal distribution.

In order to determine the impact of formative assessment, the gathered data was compared:

1. Amongst groups: (a) pre-test of control and experiment group (b) post-test of control and experiment group were compared respectively.
2. Within groups: (a) pre-test & post-test of control group and (b) pre-test & post-test of experiment group were compared accordingly.

Based on the theoretical review on statistical data analysis with non-parametric methods we used Mann-Whitney U Test and Wilcoxon Signed-Rank on population samples for comparison between groups and comparison within the group respectively.

Mann-Whitney U Test

It is a non-parametric test of null hypothesis that is used to check how likely it is that a random value from one sample set is less than or greater than a random value from another sample set. The data set must fulfill the following assumptions in order to proceed:

- Sample scores in control and experiment group are independent of each other
- Distribution of scores for both groups of independent variables (control and experiment) are symmetrical in shape
- Samples scores are measured at continuous level (0-10)

The above assumptions are listed in Mann-Whitney U test in SPSS (Laerd, 2013). Our data set successfully fulfilled all of these assumptions.

Wilcoxon Signed-Rank

It is a non-parametric hypothesis test that is used to check how much the mean rank differs for two related samples or two matched samples. In our work we used it for statistical analysis of samples

within the group, our data set successfully satisfied the assumption of use as listed in Wilcoxon Signed-Rank Test using SPSS (Laerd, 2013).

- Sample score are measured at continuous level (0-10)
- Independent variable consists of two related groups (pre-test experiment and post-test experiment) to indicate same subjects are present in both groups before and after intervention.
- Distribution of scores between independent variables needs to be symmetrical in shape

Comparison amongst Groups

Pre-test

Data of pre-test was examined to determine if there was statistical significant difference between control and experiment group. The non-parametric method, Mann-Whitney U Test was used to compare the mean rank of pre-test results. The significance level or was set at $P < 0.05$.

Post-test

In order to determine the impact of digital game-based learning formative assessment, the data of post-test was examined to determine if there was statistical significant difference between control and experiment group. The non-parametric method, Mann-Whitney U Test was utilized to compare the mean rank of post-test results. The significance level was set at $P < 0.05$.

Comparison within Groups

Control Group

Although no change in teaching methodology was applied in the control group and it was taught in a traditional manner. The data collected from control group was examined to determine if there was statistical significant difference between pre-test and post-test results of students from the same group. The non-parametric method, Wilcoxon Signed-Rank Test was used to compare the median difference between pre-test and post-test scores. The significance level was set at $P < 0.05$.

Experiment Group

In order to examine the impact of digital game-based learning formative assessment, the results of pre-test and post-test of experiment group were analyzed using Wilcoxon Signed-Rank Test, to determine if there was statistical significant difference between pre-test and post-test. The significance level was set at $P < 0.05$.

Purpose of the study

The purpose of this study was to determine the impact of digital game-based learning formative assessment in grade 3 English students. This study addressed the following question:

Research Question

What is the significance of digital game-based learning formative assessment in improving learning quality?

Null Hypothesis One

H0: There is no significant difference in students' score between control and experiment group.

Null Hypothesis Two

H0: There is no significant difference in students' scores between pre and post-test of an experiment group, after using digital game-based learning formative assessment.

Analysis of Learning Efficiency Questionnaire

Average pre-test score for control group was found out to be 11.59 that rounds off to approximately 12 and 11.52 that rounds off to approximately 12 as well for experimental group respectively. These scores implied that it is justifiable to conduct quasi control experimental research on the given sample, as the difference between the means of both the groups is negligible. Descriptive analysis was performed on the pre-test scores of both control and experimental group and the conclusive results are shown in the table below:

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Pretest_Score	Exp	100	11.515	3.3420	.3342
	Control	100	11.590	2.8598	.2860

Table 5 Descriptive statistics of Pre-test of both groups

Also, an Independent t-test was performed on the pre-test scores and it showed that no significant difference ($p>0.05$) was found between the control and experiment groups.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Pretest_Score	Equal variances assumed	2.050	.154	-.171	198	.865	-.0750	.4399	-.9424	.7924
	Equal variances not assumed			-.171	193.381	.865	-.0750	.4399	-.9425	.7925

Table 6 T-test results for pre-test of control and experiment group

Descriptive analysis performed on the post-test scores of both control and experimental group and the conclusive results are shown in the table below:

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Posttest_score	Exp	100	13.075	2.8010	.2801
	Control	100	10.835	2.7561	.2756

Table 7 Descriptive statistics of Post-test of both groups

In order for the Independent t test to run there are two prerequisite tests that need to be run first. These tests are:

1. Test for normality

2. Test for homogeneity of variance

Test for normality

Tests for normality is conducted in order to check whether the sample data actually is normally distributed or not. Normally distributed data means a set of data with a central value (i.e. mean) and majority of the rest of the values are symmetrically centered about that mean, whereas the spread of the data (i.e. the data range) is determined by its standard deviation. The mean and standard deviation are also known as parameters of the distribution. In statistical data analysis we estimate the parameters of distribution assumed for the data set, such methods that use distributional assumptions are called parametric methods (Altman). In statistical analysis, normal data set is a requirement for parametric tests (Laerd, 2013). Parametric methods assume that data being analyzed follows normal distribution, which means that the population from which the data samples are collected is normally distributed with uniform variance between groups. Populations that are not normally distributed are analyzed using non-parametric methods. There are two tests that give this value namely: Kolmogorov-Smirnov test and Shapiro-Wilk test.

Shapiro Wilk test was used in this scenario, as this is the most appropriate test for data with sample sizes that range from 0 to 2000. Shapiro-Wilk test is considered to be powerful than Kolmogorov-Smirnov Test (Zahediasl, 2012). As mentioned above that in our study the sample size was of 200 students, hence, Shapiro Wilk test was used.

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Group	Statistic	df	Sig.	Statistic	df	Sig.
Pretest_Score	Control	.089	100	.049	.973	100	.041
	Exp	.084	100	.078	.976	100	.064

a. Lilliefors Significance Correction

Table 8 Test of normality for pre-test score of control and experiment group

As evident from the table above that P-values (sig) for both Control post-test scores and Experimental post-test scores were > 0.05 and W was indeed very close to 1. Hence, we can conclude that data was normally distributed for pre-test score of both control and experiment group. Similarly, Shapiro-Wilk test was run again to check the normality of data in post-test scores of control and experiment group and the results are shown below:

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Group	Statistic	df	Sig.	Statistic	df	Sig.
Posttest_score	Control	.104	100	.010	.980	100	.140
	Exp	.099	100	.016	.968	100	.015

a. Lilliefors Significance Correction

Table 9 Test of normality for post-test score of control and experiment group

Test for homogeneity of variance

The second test is conducted to check that whether the spread of scores was similar in both the groups or not. The spread of scores was reflected in the variance. This was checked by the value of Levene's test.

If the Levene's test results is statistically significant ($p<0.05$) then the data does not show homogeneity of variance, however if the Levene's test results is statistically not significant ($p>0.05$) then data shows homogeneity of variance.

		Levene's Test for Equality of Variances	
		F	Sig.
Pretest_Score	Equal variances assumed	2.050	.154
	Equal variances not assumed		

Table 10 Levene's test for pre-test score

P value (represented by sig) for Levene's Test was ($p>0.05$) hence it might be concluded that pre-test scores for both control and experimental group showed homogeneity of variance

		Levene's Test for Equality of Variances	
		F	Sig.
Posttest_score	Equal variances assumed	.001	.980
	Equal variances not assumed		

Table 11 Levene's test for post-test score

P value (represented by sig) for Levene's Test was ($p>0.05$) hence it might be concluded that post-test scores for both control and experimental group showed homogeneity of variance. The scores approximately normally distribute (as proved by the value of Shapiro-Wilk test) and by Levene's test ($p>0.05$) it was proved that spread of scores for both the control group and the experimental group had been similar hence we may proceed with the Independent t test.

Independent t-Test

Post-test scores from two different groups (control and experimental) had been taken so Independent t test was used (Greasley, 2008). Independent t-test is used to check whether the difference of average between two groups (control group and experimental group) is likely to have occurred because of some random possibility or reason in sample selection. The difference tends to be more plausible

- If the sample size is large
- If there is largely considerable difference between the

average of both the groups

- Standard deviation is less, i.e. responses from both the groups are quite close to one another.

T-test yields the value of statistical significance and the size of effect (Greasley, 2008). The statistical significance measures whether the difference between the two sample groups can be represented by the difference in the averages of both the samples. First it is checked if there lies a difference between the averages of both the samples. The differences among the averages should be large enough and should make sense practically. The effect size indicates that value for meaningfulness (ibid).

Independent samples T-test assumes that the means of post-test scores of two groups (control group and experimental group) have significant difference between amongst them. Significance values chosen is 0.05, which makes the confidence interval for mean to be 95%. Confidence interval of 95% indicates that on 95% of occasions, it goes to imply that when our sampling is repeated it would be expected that 95% of the times differences would lie within these parameters.

By looking at the significance value of (2-tailed) in the table below, it shows the value of statistical differences between the post-test results of both the groups (control group and experimental group). Now, since this value is <= 0.05 thus we can say that null hypothesis is rejected and there lies a significant difference in the learning quality and efficiency of those students that used interactive learning application in their classrooms during the course of intervention.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Posttest_score	Equal variances assumed	.001	.980	5.700	198	.000	2.2400	.3930	1.4651	3.0149
	Equal variances not assumed			5.700	197.948	.000	2.2400	.3930	1.4651	3.0149

Table 12 T-test results for post-test of control and experiment group

In order to further the claim that the Null hypothesis mentioned above holds or not, we ran further tests that are discussed below.

Null Hypothesis One

To determine if digital game-based learning formative assessment had an impact on the scores of primary school students in English, the pre-test scores of control group were compared with the pre-test scores of experiment group while the post-test scores of control group were compared with the post-test scores of experiment group. The data collected through pre-test and post-test were validated by the concerned teacher. A non-parametric test, Mann Whitney U Test, for the comparison of scores between control and experiment groups was used. The significance level was set at 0.05.

The first null hypothesis: “there is no significant difference between control and experiment group”, was rejected with a significance level set at < 0.05. The results of Mann Whitney U Test concluded that there was a significant impact of formative assessment in improving scores in formative assessment at primary level.

Demographics of the participants

To reject the null hypothesis one, the data of 100 students of control group and 100 students of experiment group were analyzed. The table below shows the demographics of the participants:

Group	No. of participants	Gender	Grade
Control	100	Male Female	3rd Grade
Experiment	100	Male Female	
Total	200		

Table 13 Demographics of Participants for first hypothesis

Test to compare scores between Control and Experiment Group

Pre-Test

A Mann-Whitney U Test was run to determine if there were differences in pre-test score between control and experiment group. The effect of independent variable (group) was measured on dependent variable (score). The results showed a significance value of 0.802 with U= 4,897.5000 and z = - 0.251. The results are shown in the table below:

Total N	200
Mann-Whitney U	4,897.500
Wilcoxon W	9,947.500
Test Statistic	4,897.500
Standard Error	408.738
Standardized Test Statistic	-.251
Asymptotic Sig. (2-sided test)	.802

Table 14 Mann Whitney U Test results for pre-test of control and experiment group

The hypothesis test summary as illustrated below, shows that the distribution of pre-test score was the same in both groups, there was a slight negligible difference between pre-test scores of control and experiment group.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Pretest_Score is the same across categories of Groups.	Independent-Samples Mann-Whitney U Test	.802	Retain the null hypothesis.

Table 15 Hypothesis test summary of pre-test

The figure below compares the mean rank of control and experiment group. The pre-test mean rank for control group was

101.52, which was higher than experiment group, which had a mean rank of 99.48, which consequently should imply a higher post-test score mean for control group.

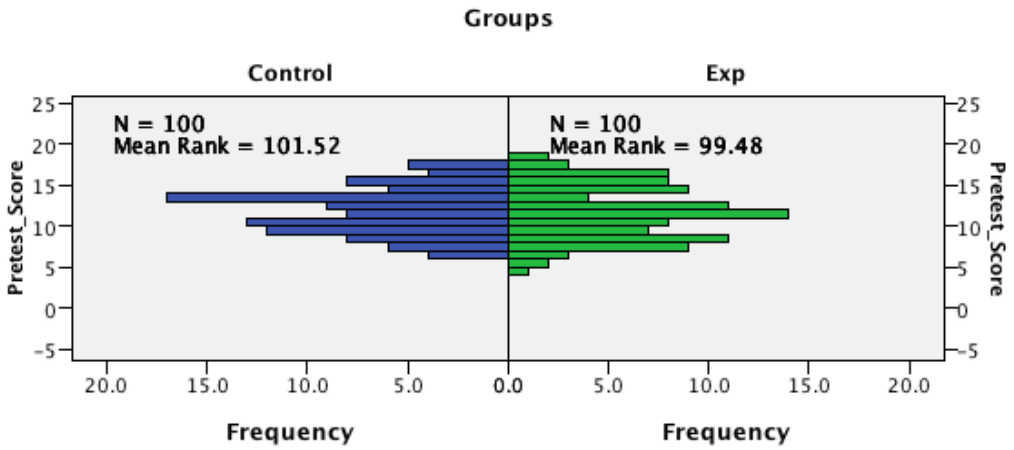


Figure 16 Comparison of control and experiment group's pre-test

Post-Test

In order to reject the null hypothesis "there is no significant difference in students' score between control and experiment group", a non-parametric test, Mann-Whitney U Test was run on post-test data. The effect of independent variable (group) was measured on dependent variable (score). The results of Mann Whitney U Test are given in the table below. The results showed a significance of 0.000 that is less than 0.05, with U= 7134.000 and z = 5.222. Hence the significance value of 0.00 rejected the null hypothesis, which implies that there is a significant difference in students' score between control and experiment group

Total N	200
Mann-Whitney U	7,134.000
Wilcoxon W	12,184.000
Test Statistic	7,134.000
Standard Error	408.620
Standardized Test Statistic	5.222
Asymptotic Sig. (2-sided test)	.000

Table 16 Mann Whitney U Test results for post-test of control and experiment group

The hypothesis test summary is illustrated below shows that the distribution of post-test score was not same in both groups, and there was a significant difference between post-test scores of control and experiment group. The significance value is less than 0.05 that rejected the null hypothesis.

2	The distribution of Posttest_score is the same across categories of Groups.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 17 Hypothesis Test Summary of post-test

The graph below compares the mean rank of control and experiment group. The post-test mean rank for experiment group was 121.84, which was statistically significantly higher than control group that had a mean rank of 79.16. This showed that there is a significant impact of digital game-based learning formative assessment. The hypothesis that there is a significant difference in students' score between post-test of control and experiment group was supported by the results.

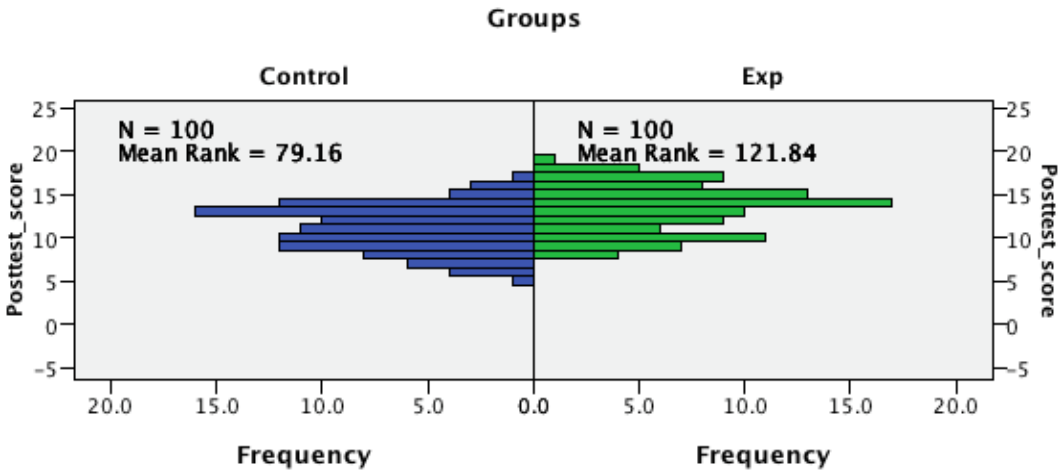


Figure 17 Comparison of control and experiment group's post-test

The following table gives an overview of the mean rank of pre-test and post-test of both control and experiment group. The mean rank of control and experiment group's pre-test was 101.53 and 99.48 respectively. While the mean rank of control and experiment group's post-test was 79.16 and 121.84 respectively. The collected data rejected the null hypothesis, which showed that there was a strong impact of formative assessment, using digital game-based learning formative assessment, in enhancing learning outcomes of students.

Ranks

	Groups	N	Mean Rank	Sum of Ranks
Pretest_Score	Control	100	101.53	10152.50
	Exp	100	99.48	9947.50
	Total	200		
Posttest_score	Control	100	79.16	7916.00
	Exp	100	121.84	12184.00
	Total	200		

Table 18 Mean rank of pre-test and post-test of control and experiment group

Null Hypothesis Two

To answer the research question, the difference of scores within control group and within experiment group was determined. The pre-test scores and post-test scores of control group were compared to prove that there was no significant improvement of scores after teaching with conventional method. While the pre-test scores of experiment group were compared with the post-test scores of experiment group to determine the impact of digital game-based learning formative assessment. The data collected through pre-test and post-test was validated by the concerned teachers. The non-parametric test, Wilcoxon Signed-Rank test was used to compare the scores within groups. The significance level was set at 0.05. The data was compared (1) within control group and (2) within experiment group.

The second null hypothesis: "there is no significant difference in students' scores between pre and post-test of an experiment group, after using a digital game-based learning formative assessment", was rejected with a significance value < 0.05. The results of Wilcoxon Signed-Rank test concluded that there was a

significant increase in students' scores after exposing them to digital game-based learning formative assessment.

Within Control Group

Demographics of the participants

Group	Gender	No. of participants	Grade
Control	**Male**	50	3rd Grade
	Female	50	
Total		100	

Table 19 Demographics of Participants for second hypothesis

The table above shows the demographics of the participants. The control group consisted of a total of 100 participants of whom 50 were male participants and 50 female students took part. To compare the control group's pre-test scores with post-test scores, data of all 100 students of control group was analyzed.

To test the null hypothesis, "there was no significant difference in students' scores between pre and post-test of control group", this test analyzed a sample of 100 participants. At the start of experiment, a pre-test was taken through which we measured our dependent variable (score). The students were taught through conventional method of teaching that is followed in their regular setting. Control group participants were not exposed to intervention. The same dependent variable was measured again for these 100 participants by taking post-test after a particular interval of time (i.e. completion of intervention with the participants of experiment group). The effect of independent variable (group) was measured on the dependent variable (score). The group had two related groups:

Related Group # 1: immediately at the start of the manipulation (pre-test)
Related Group # 2: after teaching with the conventional method (post-test)

A non-parametric test, the Wilcoxon Signed-Rank test was used to determine whether there was a median difference between pre-test and post-test scores of control group. From SPSS data the statistical significance (p-value = 0.004), showed that the median difference between the two related groups was statistically significantly different. The results of difference in pre-test scores and post-test scores of control group are shown in the table below:

Total N	100
Test Statistic	.000
Standard Error	9.579
Standardized Test Statistic	-2.871
Asymptotic Sig. (2-sided test)	.004

Table 20 Wilcoxon Signed-Rank test's results of control group's pre-test and post-test

The significance value (p = 0.004) came out to be smaller than the significance level of 0.05. The hypothesis test summary of control

group generated by Related-samples Wilcoxon Signed-rank test is given in the table below:

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Pretest_Score and Posttest_score equals 0.	Related-Samples Wilcoxon Signed Rank Test	.004	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 21 Hypothesis test summary of comparison of pre-test & post-test within control group

The median of difference between pre-test and post-test scores within control group came out to be 0, as shown in the hypothesis test summary. The results of Wilcoxon Signed-Rank test showed that there was no improvement in scores of control group after teaching the participants with traditional method of teaching. The test generated a histogram which illustrated that out of 100 participants of control group, 10 participants obtained fewer score in post-test as compared to pre-test. None of the students showed improvement in score or positive results and 90 students had same scores in post-test, as in pre-test. Hence it was concluded that there was no improvement in scores of control group after teaching them the same chapter by conventional method without any exposure to intervention. The histogram is illustrated below:

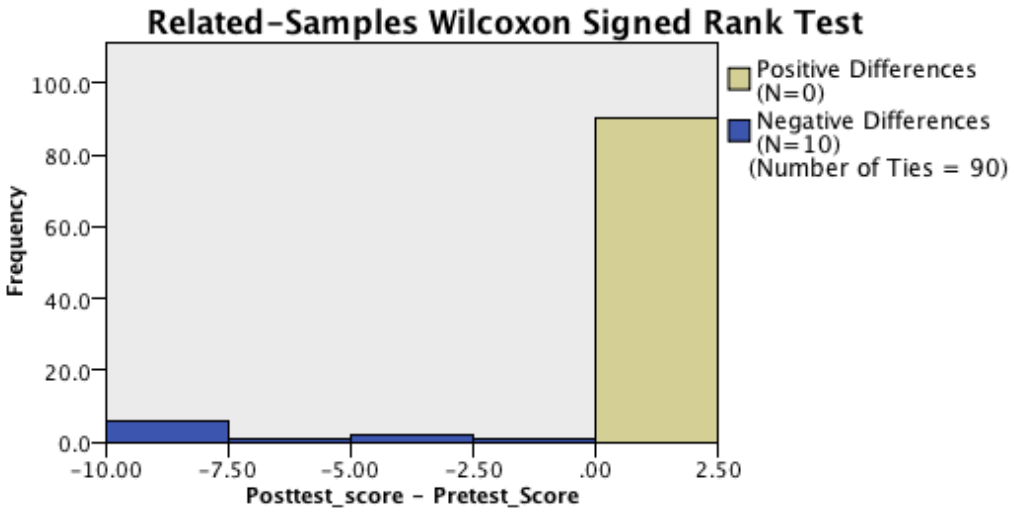


Figure 18 Histogram of difference between pre-test score and post-test score of control group

Within Experiment Group

Demographics of the participants

Group	Gender	No. of participants	Grade
Experiment	Male	50	3rd Grade
	Female	50	
Total		100	

Table 22 Demographic of participants for part two of second hypothesis

The table above shows the demographics of the participants. The experiment group consisted of 100 participants of whom 50 were male participants and 50 female students participated. To compare the experiment group's pre-test scores with their own post-test

scores, data of all the 100 participants of experiment group was analyzed.

To test the null hypothesis, "there is no significant difference in students' scores between pre and post-test of an experiment group, after using an interactive quiz application", the data of 100 participants was analyzed. At the start of experiment, the dependent variable (score) was measured by taking a pre-test. The same dependent variable was measured again for these 100 participants after a particular time period in which intervention took place by taking a post-test. The effect of independent variable (group) was measured on the dependent variable (score). The group had two related groups:

Related Group # 1: before intervention (pre-test)
Related Group # 2: after intervention (post-test)

Total N	100
Test Statistic	3,344.500
Standard Error	277.692
Standardized Test Statistic	3.486
Asymptotic Sig. (2-sided test)	.000

Table 23 Wilcoxon Signed-Rank test's result of experiment group's pre-test & post-test

A non-parametric test, the Wilcoxon Signed-Rank test was used to determine whether there was a median difference between pre-test and post-test scores of experiment group. The significance value came out to be 0.000, with z = 3.486. From SPSS data, the statistical significance (p-value = 0.000), showed that the median difference between the two related groups was statistically significantly different. The results of difference in pre-test scores and post-test scores of experiment group are shown in the table above.
The significance value (p = 0.000) came out to be less than 0.05, thus rejecting the null hypothesis. The hypothesis test summary of experiment group generated by Related-samples Wilcoxon Signed-Rank test is given in the table below:

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Pretest_Score and Posttest_score equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 24 Hypothesis test summary of comparison of pre-test & post-test within experiment group

The results of Wilcoxon Signed-Rank test showed that there was a significant improvement in scores of experiment group after exposing them to formative assessment's interactive quiz application. The test generated a histogram which illustrated that out of 100 participants of control group, 33 participants obtained fewer

score in post-test as compared to pre-test. 60 participants showed improvement in score or positive results and 3 students had same scores in post-test, as in pre-test.

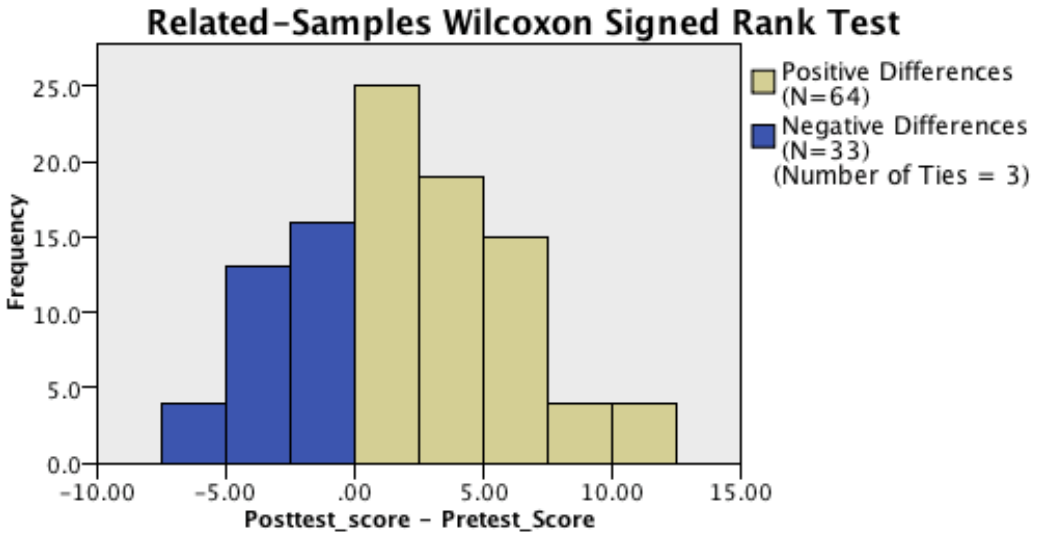


Figure 19 Histogram of difference between pre-test score and post-test score of experiment group

Since the number of participants that showed improvement is significantly higher than those who attained the same or fewer marks, it was concluded that there was a significant impact of formative assessment in enhancing learning efficiency, by using digital game-based learning formative assessment. The conclusion can be drawn that there was a statistical significant improvement in experiment group's pre-test and post-test score, after the introduction of intervention.

Chapter 5

Discussion

Present study has contributed in the field of educational research especially for a country like Pakistan where dropout rate is increasing day by day and it is becoming impossible to motivate and excite students towards learning. In urban and rural areas percentage of boys and girls willing to go to school is not at all satisfying (Pakistan National Action Plan, 2007-2008).

Conventional classroom set up is considered to be the main cause of lack of positive learning outcomes and decrease in learning efficiency. Conventional classroom setup focuses on transmission of information from teachers to students instead of focusing on student learning and better understanding. So, in spite, the student rote learns the answers or takes a guess during quizzes. This approach may help them to get higher grades and more marks, but they are unable to solve any problems in the real world or for that matter any problems outside the book.

Students act like passive receiver of the syllabus while the teacher is the sole in charge of the classroom, does not have enough resources to access whether each student has understood the concept or not. In some cases, they are either too afraid or too shy to ask the teacher to either repeat themselves or to let the teacher know in case they didn't quite get the gist of what's being taught. There is no way by which students can construct knowledge on their own that is why the process of going to school and getting

education seems boring to almost all the students. Most of the students do not find it exciting to come to school and those who come are only taking notes without understanding the concepts well. As every day there will be a typical boring lecture, teacher will call either one or two students to solve few problems on the black board and then lecture ends. There are no means to check that whether each student has understood the concept or not.

Similarly, the assessments taken are only a measure of how much a student remembers rather than focusing on students understanding and students' ability to apply the knowledge in real life. The current methods of assessments being practiced are mostly summative rather than formative. Formative assessment is an ongoing assessment that improves learning through efficient feedback and giving ownership of their work to students (Abiodullah, 2009). Literature review done as part of the current study shows similar notions about the use of formative assessment practices in the classroom environment.

In addition to this, use of digital game-based learning formative assessments as an enabler has also proved its significance in today's ICT dominated world. Studies have demonstrated that appropriate implementation of ICT in language classroom boosts learning, promote communication, enhance targeted outcomes and motivates students. Academia can use more powerful tools for effective learning than the traditional conventional teaching methods. While the past research work has indeed established formative assessment and its associated attributes to be really effective in improving the learning quality and learning efficacy, we have examined its efficacy in the context of Pakistan's primary education context in conjunction with digital game-based learning

formative assessments for grade 3 English class.

Purpose of this study

The purpose of the study was to determine whether digital game-based formative assessments has an impact on learning quality and efficiency of grade 3 English students. The application and research methodology was designed in accordance with material included in grade 3 English Book conferred by National curriculum followed in school provided by Punjab Textbook Board. The study shed light on an overview of current state of affairs of Pakistan's educational system and how it lags behind in fulfilling the learning needs of students. The research design and methodology used formative assessment attributes in order to investigate the research question. The study results showed that formative assessment enhances student's learning, understanding, quality and efficiency.

Impact on learning Quality

Null Hypothesis 1: There is no significant difference in students' score between control and experiment group.

The first null hypothesis was analyzed by applying the non-parametric Mann-Whitney U at significance level < 0.05. The results of Mann-Whitney U test showed that post-test data of control group and post-test data of experiment group not only differed in score distribution but also the high mean score values of experiment group lead to the conclusion that the null hypothesis is rejected. Hence, the result established that the use of digital game-based leaning formative assessments does have a positive impact on learning quality as a significant difference in students' score between control and experiment group was observed.

Null Hypothesis 2: There is no significant difference in students' scores between pre and post-test of an experiment group, after using digital game-based learning formative assessment.

The second null hypothesis was analyzed by using Wilcoxon Signed-Rank test at significance level < 0.05. The results of Wilcoxon Signed-Rank test showed that pre-test and post-test scores of experiment group were statistically significantly different and large improvement in experiment group's scores established that the null hypothesis is rejected. As evident from the results, use of digital game-based leaning formative assessments does have a positive impact on learning quality as a significant difference in students' score between control and experiment group was observed.

It can be argued that after a certain interval of exposure to the topic and a foreign assessment tool, the students of control group could also have shown significant difference in their pre-test and post-test scores. In order to resolve the ambiguity, the scores of pre-test and post-test of control group were also run using the same test i.e. Wilcoxon Signed-Rank test at significance level < 0.05. The results of Wilcoxon Signed-Rank test showed that there was no improvement in scores of control group after teaching the participants with traditional method of teaching and no significant difference between their pre-test and post-test scores was observed. Which further strengthens our claim that the improvement in performance and learning quality of students was because of the introduction of digital game-based learning formative assessment as evident from the difference in scores of within group pre-test and post-test.

The findings from the above two hypotheses suggest that student performance significantly improved in comparison to the conventional classroom setting. While assessing the results of overall pre-test and post-test results of control as well as experiment group, one must bear in mind that this data was gathered from students who are not native English speakers and English is being taught to them as secondary or foreign language. Second language educationists have argued that students bring with them to the classroom a complicated set of "attitudes, experiences, expectations, beliefs, and learning strategies" (Benson, 2001) (Nyikos & Oxford, 1993) (Oxford, 1992). These "attitudes toward learning, and the perceptions and beliefs that determine them", may have an insightful effect on learning performance (Bandura & Schunk, 1981) (Como, 1986) (Cotterall, 1995) (McCombs, 1984) and "learning outcomes" (Martin & Ramsden, 1987) (Van Rossum & Schenk, 1984) (Weinert & Kluwe, 1987).

The use of formative attributes in learning proved to be an important factor in increasing the understanding and cognitive thinking of students. A culture of success can be developed using formative assessment practices in the classroom by concentrating on specific learning needs (Wiliam & Black, 2010). Games have the ability to sustain engagement and motivation over the period of time, which is specifically useful when the learning tasks are challenging (Katie Larsen McClarty, 2012). The use of rewards to gauge the learning progress in the application design supported self-assessment. Self-assessment helps students develop understanding of learning goals and where they stand on the learning continuum (Wiliam & Black, 2010). The use of immediate descriptive feedback in the assessment application design helped the students to identify

the areas that needed more understanding and attention in order to well develop their thought process and resolve disparities as Black & William highlighted in their findings. Formative feedback and peer assessment results in co-ownership in the learning and development of students (Ahmed, 2015). Further the quiz application designed didn't limit the assessment time for students, which enabled the students to think and then answer the questions without the stress of being timed out. There are possibilities for continued practice and encouragement as games allow replay and advancement to next levels (Katie Larsen McClarty, 2012). The use of multiple choice questions and drag and drop options in the assessment also invalidated rote learning that is in common practice in the schools in Pakistan. Learners and learning outcomes are not considered in designing the assessments, instead rote learning and memorization of manuscripts is encouraged in Pakistan's educational system (Khattak, Assessment in schools in Pakistan, 2012). The application design goal was instead to instigate thinking and reasoning that proved to enhance learning and student performance. It is not the technology alone but how teachers and students utilize it makes it an important learning tool (Norbert Pachler C. D., 2009). Our study further proves that the use of technology in the assessment has the potential to diversify the assessment methods being used in Pakistan, which are mainly summative in nature. The positive results from this digital assessment show that using digital game-based learning formative assessments can enhance learning quality, efficiency and teaching strategy.

From the research design, methodology and corresponding results of this study we can emphasize the use of formative assessment in

primary classroom setting. Our data result establishes that in the current learning scenario of primary schools in Pakistan, use of formative assessment is indeed valuable and can improve learning outcomes of students. The student performance can foster in comparison to the traditional ways of teaching and learning that prevail presently. In an interactive learning environment, students felt more motivated and involved in the learning process and their ownership towards their learning goals enhanced. Teachers and instructors in Pakistan can utilize these attributes of formative assessment in contrast to summative assessment, in their classroom to foster learning and develop an inherent interest in education. As pointed out in the problem statement the educational scenario in Pakistan is considerably lagging as compared to the internationally set standards of minimum education. Summative assessment is the most commonly practiced mode of assessment in which learning targets and achievements of students are in clear disparity. Thus, there is a need to introduce reforms and make education friendly policies on the national level to address the current deficiencies. The work done in this study is not only an eye opener but brings a lot of possibilities and opportunities for educators and learners to improve their learning skills, attain achievement, accelerate knowledge and develop true understanding of their learning targets. The literature review of the formative ways of learning and the formative characteristics used in the research design of this study clearly establish formative assessment and its attributes as a gatekeeper in education. Engagement, interaction, collaboration, self-assessment, descriptive feedback, motivation and self-esteem are the colors of formative assessment that are vivid in research and practice of education. Our study added to the existing findings and augmented it further with the use of gamified quiz application for assessment in

the primary schools of Pakistan. Specifically, for Pakistan where technology is not as widely used as in the developed countries, the use of game-based application for assessment is one example of how much potential the information and communication technology possess to improve the standard of education, which is in complete agreement with earlier research work. The use of tablets and handheld devices is getting popular day by day and so is their availability and affordability. From our results, we believe that if game-based assessments become readily available for different courses, students will actively participate in achieving their learning objectives. Also, teachers will be able to monitor the student progress and optimize the course work and assessment design accordingly. This can bring a shift in the pedagogy and assessment practices, which today are dominated by summative assessment. Primary education is compulsory in Pakistan and it is therefore really important to engage students well into the educational philosophy instead of demotivating them or proving them as failure. From our study, we establish that formative assessment is imperative to the learning improvement, self-efficacy and motivation of students. It considerably fosters learning outcomes and understanding of students, therefore it should be practiced in the education sector of Pakistan.

Chapter 6

Conclusion

The study intended to measure the effect of digital game-based learning formative assessment on learning quality and efficiency and it concluded that Gamification of formative assessments does have a significant impact on the learning of students.

The first phase of the study involved the design and development of digital game-based leaning formative assessment in accordance with the National curriculum of Pakistan. For that purpose, similar applications, relevant literature and the textbook were reviewed in depth to better understand the prerequisites of game development.

The second phase of the study involved collecting and analyzing data with the help of pre-test and post-test along with the usage of digital game-based learning formative assessment to compare results for control group and experiment group.

The results of the study show that all the null hypotheses were not supported and thus it showed that student learning quality and efficiency were seen to have improved by the use of digital game-based leaning formative assessment.

In view of the fact of the results of this research of our digital game-based learning formative assessment, the effect of DGBL is

considerably significant than that of conventional method i.e. summative assessments.

Limitations of the study

The study contained some limitations that must be taken into consideration before applying it to other research settings. They are as follows:

Research Time Period

As the research in school was to be carried out before their first term exams that too with in a period of two weeks, this may affect the results of long term intervention. A longer time period can be more helpful in the evaluation of real results of student's learning as well as the extension of curriculum being covered in the application.

Availability of Single Observer

During the research, a single researcher was responsible for observing the students, helping them, resolving issues and conducting FGD. For improvement of the reliability of the ratings, multiple researchers must be appointed to establish a measure of inner rating reliability.

Availability of Tablets

There were not enough Android tablets available to conduct the assessment on large population. The data collected in this study utilized 20 tablets and used them iteratively to record assessment response from 200 primary school students. All the students were not alike in the use of Android tablets. The non-uniformity in the ease of use raises the question whether this study was able to

effectively measure the qualitative and quantitative data, which may change over time.

Availability of Graphic Designers

Due to limited availability of graphic designers, the graphic design of the digital game-based learning application was not up to the standard of good usability design. Lack in the quality of graphics and animations was experienced that could have been achieved if proper graphic designer help was available.

Chapter 7

Recommendations

General Recommendations

The current study tries to give a glimpse of effective digital formative assessments on students learning and has demonstrated that these Interactive learning environments and formative assessments are the future and much needed innovation in Public sector education.

In order for that to happen the Government should do the needful and provide ample resources to the concerned department. Some of the suggested resources include but are not limited to:

- Financial resources
- Physical resources (building better equipped computer labs, sufficient tablets so that all students may work simultaneously and other form of DGBL playable devices)
- Intellectual resources (digital teachable material, games, online assessments etc. according to or related to their curriculum)
- Human resources (teachers, lab assistants or qualified instructor familiar with use of technology)
- Relevant training programs

English being the foreign language itself is a challenge for the students therefore extra classes of English spoken or short courses should be offered in their own campuses so that they might be able to improve their skills. Lastly but most importantly, there is a need to introduce reforms and make education friendly policies on the national level to address the current deficiencies.

Recommendations for Future Research

The blend of education, interactivity and technology is a very appealing field of research and the only way forward is in adapting these changes and applying these on large scale. Although this research provides an insight in doing the needful, it also contains a lot of scope for further research in the field. Some recommendations for future research are as follows:

- This research only addresses the impact of digital game-based learning formative assessment in grade 3 students. It should be expanded to measure efficiency of other grades as well.
- Similarly, it only adheres to the effect in English Subject that too of grammar unit one only, it should be further completed for all units as well as subjects.
- The study data was collected from a limited number of students from different branches of same school, we recommend extending the study to further schools.
- The intervention took place for a limited time; a long-term study must be conducted in which students become more familiar with the interactive learning process.
- The intervention may be repeated with more number of tablets so that all participants can work simultaneously.
- We recommend the use of video cameras to figure out the contribution and impact of different external factors on the results.
- There should be more than one observer conducting the research in order to get the right gist of observatory factors.

Chapter 8

References

Bibliography

(2017, November 8). Retrieved from Associated Press of Pakistan Corporation: http://www.app.com.pk/govt-taking-initiatives-to-improve-education-standard-57-million-illiterate-in-country/

Abiodullah, H. M. (2009). Using Assessment for Improving Students Learning: an analysis of University Teachers' Practices. *Bulletin of Education and Research*, 47-59.

Ahmed, S. (2015). *Formative assessment and productive pedagogy in Finnish Classroom assessment:In the lens of curriculum materials.* University of Oslo.

Akhtar, R. N. (2011). Pakistani Students' and Teachers' Beliefs about English Language Learning at School Level: An Analytical Study. *Journal of Education and Practice*.

Altman, D. G. (n.d.). *Parametric v non-parametric methods for data analysis.* From the BMJ: doi: http://dx.doi.org/10.1136/bmj.a3167

Anwar, M., Tahir, T., & Batool, S. (2012). *HIGH DROPOUT CONTRIBUTES TO LOW LITERACY RATE IN PAKSITAN:COMPARISON OF DROPOUT AT H.S.S.C AND COLLEGE LEVEL IN PAKISTAN.* Academic Research International.

Aravinthan, T., & L., W. (2011). Effectiveness of formative online quizzes in learning and teaching a structural engineering course. *AAEE Conference.* Fremantle, Australia.

Arikunto, S. (2010). *Research Procedure: Practical Approach.* Jakarta: Rineka Cipta.

Ary, D., Jacobs, L., & Razavich, A. (1985). *Introduction to Research in Education.* New York: Holt, Rinchart and Winston.

Bai, H., & et al. (2012). Assessing the effectiveness of a 3-D instructional game on improving mathematics achievement and motivation of middle school students. *British Journal of Educational Technology*, 993-1003.

Bandura, A. &. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personal and Social Psychology*, 586-598.

Bandura, A., & Schunk, D. H. (1981). Cultivating Competence, Self-Efficacy and Intrinsic Interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 586-598.

Baranowski, T., & et al. (2008). Playing for real: video games and stories for health-related behavior change. *American journal of preventive medicine*, 74.

Baron, D. (1982). *Grammar and good taste: Reforming the American language. .* New Haven, NJ: Yale University Press.

Benson, P. (2001). Teaching and Researching Autonomy in Language Learning. *Harlow: Longman/Pearson Education*.

Biggs, J. (1999). *What the student does: Teaching for enhanced learning* (Vol. 18).

Bleumer, L., All, A., Marien, I., Schurmans, D., Looy, J. V., Jacobs, A., . . . D., G. F. (2012). State of Play of Digital Games for Empowerment and Inclusion: A Review of the Literature and Empirical Cases.

Buckless, F. A. (2014). Using virtual worlds to simulate real-world audit procedures. *Issues in Accounting Education*, 389-417.

Campbell, D. S. (1963). *Campbell, D.T., Stanley,Experimental and quasi-experimental designs for research.* Boston: Houghton Mifflin.

Chiang, Y. T. (2011). Exploring Online Game Players' Flow Experiences and Positive Affect. *The Turkish Online Journal of Educational Technology*, 106-114.

Cohen, L., Manion, L., & Morrison, K. (2007). Research Methods in Education. *British Journal of Educational Studies*, 347-485.

Como, L. (1986). The metacognitive control components of self-regulated learning. *Contemporary Educational Psychology*, 333-346.

Corsi, T., & et al. (2006). The real-time global supply chain game: New educational tool for developing supply chain management professionals. *Transportation Journal*, 61-73.

Corti, K. (2006). *Games-based Learning: a serious business application*. Retrieved from www.pixelearning.com/docs/seriousgamesbusinessapplications.pdf

Cotterall, L. (1995). Readiness for autonomy: Investigating learner beliefs. *System*, 195-205.

Creswell, J. W. (2013). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches.* Washington DC: Sage.

Darling-Hammond, L. (2012). *Powerful Teacher Education: Lessons from examplary programs.* Hoboken, NJ: John Wiley & sons, Inc.

European Commission. (2006). *The Main Pedagogical Principles Underlying the Teaching of Languages to Very Young Learners.*

Gardner, L., Sheridan, D., & White, D. (2002). A Web-based learning and assessment system to support flexible education. *Journal of Computer Assisted Learning*, 125-136.

Gómez, M. (2014). Ciencias Sociales y gamificación. 257-262.

Greasley, P. (2008). *Quantitative data analysis using SPSS – an introduction for health and social science.* Berkshire,England: Open University Press.

Hainey, T. (2010). Using Games-Based Learning to Teach Requirements Collection and Analysis at Tertiary Education Level. *PhD thesis, University of the West of Scotland*.

Henly, D. C. (2003). Use of web-based formative assessment to support student learning in a metabolism/nutrition unit. *European Journal of Dental Education.*, 116-122.

Investopedia, L. (2016). *Investopedia - Sharper insight. Smarter Investing.* Retrieved 2016 йил December from Investopedia: http://www.investopedia.com/

Jacoby, J. C. (2014). Enhancing learning through formative assessment. *Innovations in Education and Teaching International*, 72-83.

Jerri, L., & Ligon, E. B. (2016). On the learning benefits of confidence-weighted testing. *Cognitive Research: Principals and Implications*.
Katie Larsen McClarty, A. O. (2012). *A Literature Review of Gaming in Education .* PEARSON.
Khatibi, E., & Cowie, E. (2013). *Language Learning through Interactive Games.* Malmo: Malmo Hogskola University.
Khattak, S. G. (2012). Assessment in schools in Pakistan. *SA-eDUC JOURNAL*.
Khattak, S. G. (2012). Assessment in Schools in Pakistan. *SA-eDUC JOURNAL*.
Kraiger, K., Ford, J., & Salas E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of applied psychology*, 311.
Laerd. (2013). *SPSS Statistics Tutorial and Statistical Guide*. From Laerd Statistics: https://statistics.laerd.com/
Malik, Amin, & Ahmad. (2015). *Education for All 2015 National Review Report : Pakistan.* UNESCO.
Martin, E., & Ramsden, P. (1987). Learning skills, or skill in learning? In M. &. J.T.E. Richardson, *Student Learning* (pp. 155-167). Milton Keynes, England: Open University Press.
McCombs, B. (1984). Process and skills and underlying continuing intrinsic motivation skills training interventions. *Educational Psychologist*, 199-218.
Neys, J., & et al. (2012). Poverty is not a game: behavioral changes and long term effects after playing PING. *International Speech Communication Association.* Portland: 13th annual conference .
Norbert Pachler, C. D. (2009). Formative e-assessment: practitioner cases. *Computers and Education.*
Norbert Pachler, C. D. (2010). Formative e-assessment: Practitioner cases. *Computers & Education*, 715-721.
Nyikos, M., & Oxford, R. (1993). A factor analytic study of language learning strategy use: Interpretations from information processing theory and social psychology. *Modern Language Journal*, 11-22.
Oxford, R. (1992). Who are our students? A synthesis of foreign and second language research on individual differences. *TESL Canada Journal*, 30-49.
(2007-2008). *Pakistan National Action Plan.*
Pakistan, E. M. (2013-2016). *Pakistan National Plan of Action.*
Rehmani, A. (2003). Impact of Public Examination System on Teaching and Learning in Pakistan. *International Biannual Newsletter ANTRIEP*, 3-7.
Rehmani, A. (2012). Changing assessment practices in Pakistani schools: A case of AKU-EB middle school assessment framework. *In Search of Relevance and Sustainability of Educational Change : An International Conference at Aga Khan University Institute for Educational Development*, (pp. 285-295).
Reinders, H. (2012). *Digital Games in Language Learning and Teaching.* Kondon: Palgrave Macmillan.
Review, N. E. (2006).
Riggs, E. G., & Gholar, C. R. (2009). *Strategies that promote student engagement: Unleashing the desire to learn.* Thousand Oaks: CA: Corwin Press.

Ron Fisher, J. C. (2011). Assisting transition to university: using assessment as a formative learning tool, Assessment & Evaluation in Higher Education. *Assessment and Evaluation in Higher Education* , 225-237.

Royce, S. (1989). Formative asssessment and the design of Instructional system. *Instructional Science*, 119-144.

Ruland, J. W. (2011). The Impact of Using Formative Assessment Attributes in Daily Instruction on Student Affect.

S., A. (2004, July). Designs for learning: Studying science museums exhibits that do more than entertain. *Wiley Periodicals, Inc.*

S., B. (2009). *Active Learning Literature Survey.* Madison: University of Wisconsin.

Shute, V. (2011). Stealth assessment in computer-based games to support learning. Computer games and instruction. 503-524.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research.*, 153-189.

Squire, K. D. (2008). Video game–based learning: An emerging paradigm for instruction. *Performace Improvement Quarterly* (pp. 7-36). International Society for Performance Improvement.

StatsDirect. (n.d.). *StatsDirect Statistical Analysis Software*. From Stats Direct: http://statsdirect.com/

Thomas, R., & Hooper, E. (1991). Simulations: An Opportunity We Are Missing. *Journal of Research on Computing in Education*, 497-513.

Van Rossum, E., & Schenk, S. (1984). The relationship between learning conception, study strategy and learning outcomes. *British Journal of Educational Psychology*, 73-83.

Weinert, F., & Kluwe, R. (. (1987). *Metacognition, motivation and understanding.* Hillsdale, NJ: Lawrence Erlbaum.

Wiliam, D., & Black, P. (2010). Inside the Black Box: Raising Standards Through Classroom Assessment. 139 - 148.

Willms, J. (2003). *Student engagement at school: A sense of belonging and participation results from PISA 2000.* pisa: Paris:OECD.

Yang, C. &. (2010). Integrating Video-Capture Virtual Reality Technology into a Physically Interactive Learning Environment for English Learning. *Computers and Education*, 1346-1356.

Yip, F. W., & Kwan, A. C. (2006). Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational Media International*, 233-249.

Zahediasl, A. G. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology and Metabolism*, 4.

Chapter 9

Appendices

Appendix A: Usability Checklist

User Friendly	
Easy to Use	
The icons are clear	
The content of the app is clear	
The language used is easy to understand	
Drag & Drop clear (Expected Result) The label stops on the tagline	
Back button (Expected Result) Back button should navigate the user to previous screen.	

Replay (Expected Result) Replay button should reload that scene	
Application Start (Expected Result) Application must not take more than 1-2s to start.	
Exit Button (Expected Result) Exit button should exit the scene/game	

1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree

Appendix B: Usability Form

(To be filled by the students)

Usability Test

Name: _____

1. **I'm satisfied with this DGBL application.**

Strongly Disagree — Strongly Agree

1	2	3	4	5

]

2. **The content of application is clear to me.**

Strongly Disagree — Strongly Agree

1	2	3	4	5

3. **The application is easy to use.**

Strongly Disagree — Strongly Agree

1	2	3	4	5

4. **It was easy to drag and drop the labels.**

Strongly Disagree — Strongly Agree

1	2	3	4	5

5. **The content written is easy to understand.**

Strongly Disagree — Strongly Agree

1	2	3	4	5

6. **The icons are easy to understand.**

Strongly Disagree — Strongly Agree

1	2	3	4	5

Appendix C: Pre-Test

Pre-Test

Name: _____ Age: _____

Class: _____ School: _____

1- Encircle the **verb/s** in following sentence:
My sister follows me because my shadow saves her from the sun.

2- Encircle the **pronoun/s** in following sentence:
My sister follows me because my shadow saves her from the sun.

3- Encircle the **noun/s** in following sentence:
My sister follows me because my shadow saves her from the sun.

4- Complete the sentences with appropriate adjective from given list:
Tasty Deep Old Younger

A) A _____ food
B) The _____ man
C) A _____ boy
D) A _____book

5- A **cottage** is a:

A) Fish
B) Big House
C) Small House
D) Car

6- Warmly means:

A) Kindly
B) Angrily
C) Sadly
D) Coldly

7- Pick the correct sentence

A) I eating cake
B) I am eat cake
C) I am eating cake
D) I eats cake

8- Put the words in appropriate column

Friday **His** **Sit** **Talk** **Beautiful** **She**
Came **Kind** **Fatima** **My** **Moon** **Dark**

NOUN	VERB	ADJECTIVE	PRONOUN

Appendix D: Post-Test

Post Test

Name: _____ Age: _____

Class: _____ School: _____

1- Encircle the **verb/s** in following sentence:
He was walking and his sister was following him as they sold water-melons.

2- Encircle the **pronoun/s** in following sentence:
He was walking and his sister was following him as they sold water-melons.

3- Encircle the **noun/s** in following sentence:
He was walking and his sister was following him as they sold water-melons.

4- Complete the sentences with appropriate adjective from given list:

Tall Kind Small Dirty

E) A _____ teacher
F) A _____ cottage
G) A _____ tree
H) The _____shoes

5- Follow means to:

C) Empty
D) Chase

C) Welcome
D) Arrange

6- Afraid means:

C) Kind
D) Angry

C) Sad
D) Scared

7- Pick the correct sentence

C) I like chocolate
D) I am like chocolate

C) I am liking chocolate
D) I likes chocolate

8- Put the words in appropriate column

June **Her** **Stand** **Walk** **Strong** **He**
Went **Brave** **Ali** **Her** **Sun** **Bright**

NOUN	VERB	ADJECTIVE	PRONOUN

Appendix E: Validity form for pre/post –test

Name: _____ School: _____

1. Do you think the content written is appropriate for the students of Grade 3?

   _____

   _____

2. The pre/post tests are made keeping in mind the following Learning Outcomes:
   - **Identification and correct usage of parts of speech, vocabulary and composition w.r.t grammar section of chapter 1**

Do they fulfill these learning outcomes?

_____

_____

3. Is the vocabulary used appropriate?

   _____

   _____

4. Any comments:

   _____