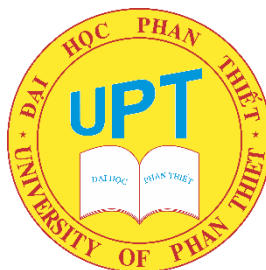


**MINISTRY OF EDUCATION AND TRAINING
UNIVERSITY OF PHAN THIET**



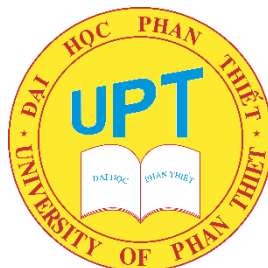
TRAN THI KIM NGAN

**THE EFFECTS OF THE ONLINE GAME BAMBOOZLE
ON PRIMARY SCHOOL STUDENTS' MOTIVATION IN
LEARNING ENGLISH: A CASE STUDY AT LE QUY DON
SCHOOL**

**MASTER'S GRADUATION PROJECT
MAJORED IN ENGLISH LANGUAGE**

Lam Dong Province - 2026

**MINISTRY OF EDUCATION AND TRAINING
UNIVERSITY OF PHAN THIET**



TRAN THI KIM NGAN

**THE EFFECTS OF THE ONLINE GAME BAMBOOZLE
ON PRIMARY SCHOOL STUDENTS' MOTIVATION IN
LEARNING ENGLISH: A CASE STUDY AT LE QUY DON
SCHOOL**

MAJOR IN ENGLISH LANGUAGE

CODE: 8220201

MASTER'S GRADUATION PROJECT

SCIENCE INSTRUCTOR'S NAME(s):

- 1. Assoc. Prof. Ph.D. LE DINH TUONG**
- 2. Ph.D. TRAN THI QUYNH LE**

Lam Dong Province – 2026

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Tran Thi Quynh Le, for her invaluable guidance, constructive feedback, and continuous encouragement throughout the process of conducting this research. Without her patience, insightful comments, and expertise, this thesis would not have been possible.

My sincere thanks also go to the board of administrators and English teachers at Le Quy Don Primary School for granting me permission to conduct the study and for their warm cooperation. I am especially grateful to the students who participated enthusiastically in the research activities. Their energy, curiosity, and willingness to engage in learning made this study both meaningful and enjoyable.

I would also like to extend my appreciation to my lecturers in the Faculty of English at Phan Thiet University who have inspired me with their dedication and knowledge during my years of study.

A heartfelt thank you goes to my family and friends, whose unconditional love, encouragement, and understanding have been a constant source of strength. Their support has sustained me not only in this thesis but also throughout my academic journey.

Finally, I am truly grateful to all who, in one way or another, contributed to the successful completion of this research.

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis entitled: “The Effects of the Online Game Baamboozle on Primary School Students’ Motivation in Learning English: A Case Study at Le Quy Don School” is the result of my own independent research and study. Except where proper acknowledgment is made, this thesis does not contain any materials that have been previously published, written, or submitted by another person for the award of any academic degree or diploma at this or any other tertiary institution.

I confirm that all sources of information, data, and ideas from published or unpublished works have been fully and properly cited in the text and included in the list of references. Any contributions made by others to this research, including but not limited to guidance, data collection support, or proofreading, have been duly acknowledged.

I also understand that plagiarism in any form constitutes a serious violation of academic integrity and the regulations of Phan Thiet University. I therefore accept full responsibility for the originality and authenticity of this thesis.

Thank you for considering this Certificate of Originality.

Lam Dong, June 2026

Tran Thi Kim Ngan

RETENTION AND USE OF THE GRADUATION PROJECT

I, Tran Thi Kim Ngan, a graduate student of the Master's Program in English Language Teaching at Phan Thiet University, fully understand the regulations of the university regarding the retention and use of master's theses.

In accordance with these regulations, I hereby agree that my thesis entitled:

“The Effects of the Online Game Baamboozle on Primary School Students’ Motivation in Learning English: A Case Study at Le Quy Don School” shall be retained in the library of Phan Thiet University and its affiliated repositories. The university is authorized to file, preserve, and provide access to this thesis in print or electronic form for the purposes of study, teaching, and academic reference.

The university has the right to reproduce this thesis, in whole or in part, for inclusion in databases, digital archives, or other academic platforms, provided that such reproduction is strictly for educational and research purposes, and proper acknowledgment of the author is made.

This retention and use policy does not extend to granting any party the right to use this work for commercial purposes. Any reproduction, distribution, or adaptation of this thesis for profit or non-academic purposes requires my prior written permission.

By signing below, I acknowledge and accept the university's rights and responsibilities in relation to the retention and use of my graduation thesis.

Sincerely,

Lam Dong, September 2025

Tran Thi Kim Ngan

ABSTRACT

This study investigated whether Baamboozle-integrated English review during the review stage of Grade 2 lessons was associated with stronger classroom motivation than teacher-centered review, and how students and their teacher experienced that format in a Vietnamese public primary school. The study addressed the problem of uneven and sometimes psychologically unsafe participation in ordinary review, where more confident pupils tend to dominate while hesitant pupils remain peripheral. Grounded in Self-Determination Theory, the research used a mixed-methods case study with an embedded pre-test/post-test comparison between two intact Grade 2 classes ($N = 60$) taught by the same teacher over six weeks. Data were collected through a 12-item emoji-based motivation questionnaire, 24 observed review lessons, 14 student interviews, and one teacher interview. The two classes were equivalent at pre-test, but at post-test the Baamboozle class reported significantly higher motivation than the comparison class ($M = 28.67$, $SD = 2.95$ vs. $M = 25.27$, $SD = 2.83$), $t(58) = 4.56$, $p < .001$, $d = 1.18$, with stronger gains also evident across autonomy, competence, and relatedness. Qualitative findings showed that Baamboozle made the review stage more exciting, easier to enter, less intimidating, and more socially supportive, while also revealing conditions related to noise, pacing, fairness, and technology. Within this bounded case, Baamboozle appears to support broader and more psychologically safe participation during English review in Vietnamese lower-primary classrooms when carefully mediated by the teacher.

Keywords: *Baamboozle, gamification, young EFL learners, classroom motivation, Self-Determination Theory, Vietnamese primary education.*

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	I
CERTIFICATE OF ORIGINALITY	II
RETENTION AND USE OF THE GRADUATION PROJECT	III
ABSTRACT	IV
TABLE OF CONTENTS	V
LIST OF TABLES	XII
LIST OF FIGURES	XIII
ABBREVIATIONS	XIV
CHAPTER 1: INTRODUCTION.....	1
1.1 Background of the Study	1
1.2 Statement of the Problem.....	4
1.3 Purpose of the Study	6
1.4 Research Questions.....	6
1.5 Research Hypotheses	7
1.6 Research Objectives	7
1.7 Significance of the Study.....	8
1.8 Scope and Delimitations	9
1.9 Limitations	10
1.10 Definition of Key Terms.....	10
1.11 Organization of the Thesis	11

CHAPTER 2: LITERATURE REVIEW	12
2.1 Theoretical Framework: Self-Determination Theory	12
2.1.1 Core propositions of SDT	12
2.1.2 SDT in second language motivation research.....	13
2.1.3 Why SDT fits young learner EFL classrooms	14
2.1.4 SDT and the design logic of gamified review.....	15
2.2 Motivation in Second Language Acquisition	16
2.3 Young Learners in EFL Classrooms	18
2.3.1 Developmental characteristics and classroom consequences.....	18
2.3.2 Participation as socially mediated and emotionally vulnerable	19
2.4 Measuring Motivation in Young Learners.....	20
2.5 Gamification and Game-Based Learning in TESOL	21
2.5.1 Conceptual distinctions	21
2.5.2 Why gamified review may support motivation.....	22
2.5.3 Critical debates and contradictory evidence	23
2.6 Baamboozle as a Classroom Review Tool	24
2.7 Empirical Studies on Digital Games, Motivation, and Language Learning.....	26
2.8 Vietnamese and Asian EFL Contexts	27

2.9 Research Gap	28
2.10 Conceptual Framework.....	30
2.11 Chapter Summary	33
CHAPTER 3: METHODOLOGY	34
3.1 Research Paradigm and Case Study Design Logic.....	34
3.2 Alignment Chain of the Study	35
3.3 Research Site and Participants	35
3.4 Variables and Operationalization	36
3.5 Instruments	39
3.5.1 Motivation questionnaire.....	39
3.5.2 Classroom observation checklist.....	40
3.5.3 Student interview protocol	41
3.5.4 Teacher interview protocol	42
3.6 Validity, Reliability, and Trustworthiness	43
3.6.1 Construct alignment	43
3.6.2 Developmental appropriateness	44
3.6.3 Reliability evidence.....	44
3.6.4 Assumption testing.....	45
3.6.5 Qualitative trustworthiness.....	46

3.7 Ethical Considerations	48
3.8 Data Collection Procedures	48
3.9 Intervention Protocol	49
3.10 Data Analysis	51
3.10.1 Quantitative analysis	51
3.10.2 Qualitative analysis	52
3.10.3 Mixed-methods integration	53
3.11 Threats to Validity and Their Management	54
3.12 Chapter Summary	55
CHAPTER 4: FINDINGS	55
4.1 Quantitative Findings.....	55
4.1.1 Reliability and assumption testing	56
4.1.2 Descriptive statistics.....	59
4.1.3 Baseline equivalence and post-test comparison	61
4.1.4 Within-group change.....	63
4.1.5 Supplementary gain-score comparison	65
4.1.6 SDT subdimension comparisons	67
4.1.7 Individual gain-pattern summary from raw data.....	69
4.1.8 Non-uniform and contradictory quantitative cases	70
4.1.9 Quantitative answer to RQ1	71

4.2 Qualitative Findings	72
4.2.1 Overview of the qualitative corpus	73
4.2.2 Observation summary	73
4.2.3 Theme 1: Joyful engagement and sustained attention.....	74
4.2.4 Theme 2: Perceived competence and confidence	76
4.2.5 Theme 3: Relatedness through team participation and peer support	77
4.2.6 Theme 4: Motivation with conditions: noise, pacing, fairness, and technological constraints	79
4.2.7 Interview report summaries.....	80
4.2.8 Qualitative answer to RQ2	81
4.3 Descriptive Mixed-Methods Joint Display	82
4.4 Chapter Summary	84
CHAPTER 5: DISCUSSION AND CONCLUSION.....	84
5.1 Discussion of RQ1: Was Baamboozle Associated with Stronger Classroom Motivation Within the Case?	85
5.2 Discussion of RQ2: How Was Baamboozle Experienced? ...	88
5.3 Interpretation Through Self-Determination Theory	91
5.3.1 Autonomy.....	91
5.3.2 Competence.....	92
5.3.3 Relatedness.....	93

5.3.4 An Integrated SDT Interpretation.....	94
5.4 Comparison With Previous Studies	95
5.5 Triangulation and Strength of Evidence	97
5.6 Pedagogical Implications.....	98
5.7 Contributions of the Study.....	101
5.7.1 Theoretical Contribution	101
5.7.2 Methodological Contribution	102
5.7.3 Pedagogical Contribution.....	103
5.7.4 Contextual Contribution.....	104
5.8 Limitations of the Study.....	104
5.9 Recommendations for Future Research	105
5.10 Chapter Conclusion	107
REFERENCES	112
APPENDICES	119
APPENDIX A. MOTIVATION QUESTIONNAIRE.....	119
Autonomy Items.....	119
Competence Items.....	119
Relatedness Items.....	119
APPENDIX B. CLASSROOM OBSERVATION CHECKLIST	

.....	120
-------	-----

**APPENDIX C. STUDENT AND TEACHER INTERVIEW
PROTOCOLS.....121**

C1. Student Interview Protocol	121
C2. Teacher Interview Protocol	121

APPENDIX D. PARENTAL CONSENT FORM.....122

APPENDIX E. STUDENT ASSENT FORM.....123

**APPENDIX F. BAAMBOOZLE INTERVENTION LESSON
PLANS124**

Lesson Plan 1	124
Lesson Plan 2	124
Lesson Plan 3	124
Lesson Plan 4.....	125
Lesson Plan 5.....	125
Lesson Plan 6.....	126

APPENDIX G. RAW QUANTITATIVE DATA.....127

Appendix Table G1. Baamboozle Group.....	127
Appendix Table G2. Comparison Group	128

APPENDIX H. QUALITATIVE CODING MATRIX.....130

Appendix Table H1. Sample Qualitative Coding Matrix.....	131
--	-----

APPENDIX I. INTERVIEW REPORT SUMMARIES..... 131

Baamboozle Group	131
Comparison Group	131
Teacher	131

LIST OF TABLES

Table 3.1 Alignment chain of the study	36
Table 3.2 Variables and operationalization	37
Table 3.3 Item-total statistics for the motivation questionnaire	45
Table 3.4 Eight-week data collection schedule	50
Table 4.1 Internal consistency of the motivation questionnaire	58
Table 4.2 Tests of Normality	59
Table 4.3 Tests of Homogeneity of Variances	60
Table 4.4 Descriptive statistics summary	61
Table 4.5 Baseline equivalence and post-test comparison	63
Table 4.6 Within-group change from pre-test to post-test	65
Table 4.7 Gain-score comparison between groups	67
Table 4.8 Post-test SDT subdimension comparisons	69
Table 4.9 Individual gain-pattern summary from raw data	71
Table 4.10 Observation summary across the 24 lessons	76
Table 4.11 Descriptive mixed-methods joint display	84

LIST OF FIGURES

Figure 2.1. Baamboozle interface.....	24
Figure 2.2. Conceptual Framework of the Study	32
Figure 4.1. Motivation scores of the Baamboozle and comparison groups	62

ABBREVIATIONS

AUT	—	Autonomy
COM	—	Competence
df	—	Degrees of Freedom
EFL	—	English as a Foreign Language
H0	—	Null Hypothesis
H1	—	Alternative Hypothesis
L2	—	Second Language
M	—	Mean
REL	—	Relatedness
RQ	—	Research Question
SD	—	Standard Deviation
SDT	—	Self-Determination Theory
SLA	—	Second Language Acquisition
TESOL	—	Teaching English to Speakers of Other Languages

CHAPTER 1: INTRODUCTION

1.1 Background of the Study

English has become a central curricular language in contemporary education because it is associated with international mobility, economic participation, digital access, and access to global knowledge (Dörnyei & Ushioda, 2009; Lamb, 2017; Pinter, 2017). In the Vietnamese context, the growing institutional emphasis on English has extended downward into the primary years, where early instruction is increasingly expected to lay the foundation for later language development (Bui et al., 2025; Nguyen, 2021; Phan et al., 2024; Phuong, 2025; Thi & Thinh, 2023). Yet starting English earlier does not, by itself, guarantee better outcomes, since young learners require age-appropriate pedagogy, meaningful exposure, and classroom conditions that support attention, confidence, and participation (Butler, 2019; Cameron, 2001; Pinter, 2017; Krashen, 1982). The effectiveness of early English education therefore depends heavily on the quality of classroom experience, particularly the extent to which lessons are able to sustain learners' willingness to participate (Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017; Reeve, 2015; Ryan & Deci, 2018).

This issue is especially important in lower-primary classrooms. Grade 2 learners are typically seven or eight years old, and at this stage they rarely participate because of long-term instrumental goals, future professional ambitions, or fully formed L2 self-concepts (Cameron, 2001; Pinter, 2017; Dörnyei & Ushioda, 2009; Lamb, 2017). Their engagement is more immediate, situated, and interactional, shaped by the emotional tone of the lesson, the availability of peer support, the visibility of success, and the extent to which classroom participation feels safe rather than threatening (Deci & Ryan, 1985, 2000; Hidi & Renninger, 2006; Oga-Baldwin et al., 2017; Reeve, 2015; Ryan & Deci, 2018). For such learners, motivation is not an abstract background factor but a condition of moment-to-moment classroom participation, especially in foreign-language settings where confidence, interest, and perceived competence strongly influence

whether children are willing to listen, respond, repeat, guess, and take part in communicative activity (Gardner et al., 1985; Noels et al., 2000; Lou & Noels, 2017; Reinders & Wattana, 2015; Bui et al., 2025).

Research in second language acquisition has consistently shown that motivation is a major predictor of effort, persistence, and willingness to engage with language tasks (Dörnyei & Ushioda, 2009; Gardner et al., 1985). However, much of the most influential motivation research has been developed with adolescents or adults in mind. Frameworks such as the socio-educational model or the L2 Motivational Self System have offered major insights into attitudes, goals, and future self-guides, but they do not always map neatly onto the experiential realities of very young learners. In early primary EFL classrooms, motivation is more plausibly understood as the learner's readiness to join, answer, try again, and remain involved in classroom activity.

For that reason, Self-Determination Theory (SDT) offers a particularly strong theoretical lens for the present study, especially because it explains motivation not simply as the amount of effort learners display, but as the quality of learners' engagement with classroom activity (Deci & Ryan, 1985, 2000; Ryan & Deci, 2018; Noels et al., 2000). SDT proposes that motivation becomes more robust, internalized, and self-endorsed when three basic psychological needs are supported: autonomy, competence, and relatedness (Deci & Ryan, 2000; Ryan & Deci, 2018). In classroom terms, autonomy concerns whether learners feel involved rather than merely controlled; competence concerns whether they feel capable and able to succeed; and relatedness concerns whether they feel socially included and supported (Reeve, 2015; Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017; Ryan & Deci, 2018). The relevance of this framework to young learner classrooms is substantial, since children's language-learning motivation is strongly shaped by classroom structure, emotional support, perceived success, and social participation rather than by distant future goals alone (Cameron, 2001; Pinter, 2017; Dörnyei & Ushioda, 2009; Lamb, 2017; Oga-

Baldwin et al., 2017). A Grade 2 child does not need full curricular independence in order to experience autonomy; rather, small forms of turn ownership, visible participation, meaningful choice, and non-threatening involvement may already support a sense of agency within highly structured classroom routines (Reeve, 2015; Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017). Likewise, competence does not require easy tasks, but it does require feedback structures that make success attainable, visible, and emotionally safe rather than humiliating, particularly for young foreign-language learners whose willingness to respond may depend on confidence and perceived ability (Butler, 2019; Lou & Noels, 2017; Reinders & Wattana, 2015; Ryan & Deci, 2018). Relatedness is equally central because young children often experience classroom participation through peer presence, team belonging, teacher warmth, and supportive interaction, making social inclusion a major condition for sustained engagement in early EFL classrooms (Cameron, 2001; Pinter, 2017; Oga-Baldwin et al., 2017; Bui et al., 2025; Carello & Butler, 2015).

These issues become especially visible during the review stage of English lessons. In many Vietnamese primary classrooms, review is conducted through teacher-fronted questioning, repetition, board work, and worksheet completion. Such practices are not inherently deficient; indeed, they can be efficient and curriculum-aligned. Yet they also have identifiable motivational limitations. Rapid questioning tends to reward already confident students. More hesitant learners may remain silent. Repetition may ensure exposure, but it does not necessarily maintain attention. As a result, classroom motivation can become unevenly distributed, with a small group of active participants and a larger group of partial observers.

This limitation helps explain the growing interest in game-like digital tools. Teachers often turn to such tools not because they assume technology is automatically transformative, but because they need practical ways to reorganize participation without abandoning the curriculum. Among these tools, Baamboozle

has become attractive because it is simple, projector-friendly, adaptable to textbook-based review, and usable even when students do not have individual devices. It allows teachers to turn ordinary review content into team-based rounds in which learners select tiles, anticipate turns, answer questions, and earn points. What appears pedagogically modest may nevertheless be motivationally significant. Tile selection can introduce bounded choice. Immediate answer confirmation can make competence more visible. Team play can convert isolated response into socially shared participation.

At the same time, gamification cannot be treated naively. Not all digital games are equally educational. Not all forms of competition are beneficial. Not all increases in excitement should be mistaken for meaningful motivational improvement. A classroom activity can be noisy, novel, and emotionally stimulating without being psychologically supportive in the SDT sense. It may also produce uneven effects, benefiting confident learners more than hesitant ones. A critical study is therefore needed, especially in lower-primary Vietnamese settings where device access is uneven, class sizes are fixed, and teachers must work within curricular pacing demands.

The present thesis addresses that need by examining the association between Baamboozle-integrated English review and the classroom motivation of Grade 2 learners at Le Quy Don Primary School, and by exploring how those learners and their teacher experienced that bounded intervention in practice. Framed as a mixed-methods case study with an embedded comparison between two intact classes, the study is significant not because it introduces an unfamiliar technological platform, but because it subjects a familiar classroom tool to theoretically grounded, methodologically aligned, and context-sensitive analysis.

1.2 Statement of the Problem

The central problem addressed in this thesis is that routine English review in lower-primary classrooms often fails to generate broad, sustained, and psychologically safe participation. In many Grade 2 classes, review is formally

completed but motivationally uneven: confident pupils answer quickly, weaker pupils remain peripheral, and the lesson phase that should consolidate learning instead reproduces participation gaps. In this sense, the problem is not simply whether children like English in general, but whether ordinary review structures allow enough children to enter, remain in, and benefit from classroom interaction.

This pedagogical problem has theoretical consequences. In research on classroom technology, motivation is often reported through broad labels such as enjoyment, interest, or engagement (Hidi & Renninger, 2006). Those descriptors are useful at a surface level, but they do not identify the motivational processes through which participation becomes more or less likely. For young learners especially, the distinction is critical. A lesson may be lively without supporting agency; it may be enjoyable for stronger pupils while remaining intimidating for hesitant ones; and it may generate short-term excitement without creating the conditions for repeated, self-endorsed participation. A sharper conceptual account is therefore needed.

Self-Determination Theory provides that sharper account because it directs attention to whether classroom activity supports autonomy, competence, and relatedness. In lower-primary EFL settings, these needs are expressed in developmentally specific ways: whether a child gets a meaningful turn, whether success feels attainable, whether error is survivable, and whether peers function as support rather than as an audience of judgment. The practical problem of uneven participation and the theoretical problem of under-specified motivation therefore converge. What appears to be a simple review-format decision may in fact reshape the need-support conditions under which young learners participate.

The contextual dimension of the problem is equally important. Vietnamese public primary schools remain underrepresented in the literature on gamified English teaching, despite the fact that implementation is heavily conditioned by local realities: whole-class projection, fixed class groupings, curriculum pacing, uneven technology reliability, and classrooms in which teacher guidance remains

central. A tool such as Baamboozle cannot be evaluated meaningfully in abstraction from those conditions, because its motivational effects depend not only on its interface but also on how it fits actual classroom ecology.

A further problem is methodological. Much of the available literature on classroom games and digital review tools is promising but evidentially uneven.

Positive claims are often based on single-group designs, brief interventions, or post-hoc self-report that documents favorable impressions without establishing whether change exceeds what ordinary teaching would already produce. Studies also tend to foreground positive outcomes while giving less analytic attention to contradictory cases, uneven uptake, or implementation constraints. That leaves an important evidential gap for researchers who need to know not only whether a tool can work, but how robustly, under what conditions, and with what qualifications.

This thesis therefore addresses a compound problem at once pedagogical, theoretical, contextual, and methodological: Grade 2 learners need review structures that support wider and more equitable participation; existing discussions of gamified motivation require a more explicit psychological account; Vietnamese lower-primary classrooms remain insufficiently represented; and stronger mixed-methods evidence is needed to explain how a familiar tool such as Baamboozle may alter motivation in practice. The study is designed to resolve these linked problems through a theory-driven, comparison-based, and context-sensitive investigation.

1.3 Purpose of the Study

The purpose of this study was to examine whether Baamboozle-integrated English review was associated with stronger classroom motivation than teacher-centered review among Grade 2 students at Le Quy Don Primary School. The study also explored how students and their English teacher experienced Baamboozle-integrated review in relation to participation, confidence, peer support, and the SDT dimensions of autonomy, competence, and relatedness.

1.4 Research Questions

This thesis is guided by two research questions

RQ1 (Quantitative): To what extent was Baamboozle-integrated English review associated with differences in the classroom motivation of Grade 2 students at Le Quy Don Primary School, compared with teacher-centered review during the review stage of Grade 2 English lessons?

RQ2 (Qualitative): How do Grade 2 students and their English teacher perceive and experience Baamboozle-integrated English review during the review stage of Grade 2 English lessons, particularly in relation to participation, confidence, and peer support?

1.5 Research Hypotheses

For the quantitative strand, the following hypotheses were tested:

Null hypothesis (H0): There was no statistically significant difference in post-test classroom motivation scores between Grade 2 students who received Baamboozle-integrated English review and those who received teacher-centered review during the review stage of Grade 2 English lessons.

Alternative hypothesis (H1): Grade 2 students who received Baamboozle-integrated English review demonstrated statistically significantly higher post-test classroom motivation scores than those who received teacher-centered review during the review stage of Grade 2 English lessons.

Because RQ2 is exploratory and interpretive, it is not linked to a null-alternative hypothesis structure. Instead, it is addressed through thematic analysis of observation and interview data.

1.6 Research Objectives

The study pursued three interrelated objectives aligned with the two research questions:

- 1. To compare classroom motivation in the Baamboozle class and the teacher-centered comparison class before and after the intervention period.*
- 2. To examine whether motivational differences were visible across the SDT dimensions of autonomy, competence, and relatedness.*

3. *To explain how Grade 2 students and their English teacher perceived and experienced Baamboozle-integrated English review during the review stage.*

1.7 Significance of the Study

The significance of the study lies first in its theoretical contribution to the analysis of young learner motivation. Rather than treating motivation as a diffuse emotional reaction to classroom games, the thesis applies Self-Determination Theory to specify how autonomy, competence, and relatedness operate in a lower-primary EFL setting. This matters because much work on children's classroom engagement remains descriptively rich but theoretically thin. By demonstrating how immediate classroom structures shape need support among Grade 2 learners, the study moves motivation analysis from generic enjoyment language toward a more explanatory account.

Second, the study is significant methodologically because it demonstrates a tighter alignment chain than is common in small-scale classroom gamification studies while keeping the work appropriately framed as a mixed-methods case study. The research problem, questions, variables, instruments, analyses, findings, and discussion are organized around the same construct logic. The study also embeds a pre-test/post-test comparison between two intact classes within one bounded school context, thereby allowing the thesis to move beyond a single-source account of motivational difference without overstating causal reach. This matters at thesis level because explanatory force depends not only on whether findings are positive, but on whether the evidential architecture can sustain interpretation.

Third, the study is pedagogically significant because it investigates a realistic instructional practice rather than an idealized technological intervention. Baamboozle is neither a costly innovation nor a fully individualized digital ecosystem. It is a low-barrier, teacher-mediated tool that can be integrated into ordinary review in projector-based classrooms. The value of the thesis therefore

lies in clarifying the specific pedagogical work the tool appears to do: widening turn entry, making success more visible, and embedding participation in peer-supported teams, while also identifying the management conditions necessary for those effects to occur.

Fourth, the study is contextually significant because it centers a Vietnamese public primary school environment that is often absent from high-visibility literature on digital pedagogy. Many published discussions of educational technology implicitly assume richer infrastructure, older learners, or one-device-per-student models. This thesis instead addresses a more typical local reality in which teachers work with intact classes, limited time, curricular obligations, and modest technological affordances. By doing so, it contributes evidence that is more transferable to similar classrooms than findings derived from highly resourced settings.

1.8 Scope and Delimitations

This study is limited to one Vietnamese public primary school and to two intact Grade 2 classes taught by the same teacher. The sample comprises 60 learners, with 30 students in the Baamboozle class and 30 in the comparison class. The study examines classroom motivation during the review stage of Grade 2 English lessons rather than broader language achievement. More specifically, it focuses on textbook-aligned review conducted through teacher-mediated, projector-based whole-class instruction. It does not claim to measure long-term proficiency development, retention across semesters, or generalized speaking ability. The intervention lasts six instructional weeks within an eight-week research window.

The study therefore investigates a bounded school-based case: one school, one grade level, two intact classes, one teacher, and one review-stage intervention period. Within that case, the same Grade 2 English content was taught in both classes, but the review stage was organized differently: through Baamboozle-

integrated English review in one class and through teacher-centered review in the other.

1.9 Limitations

Several limitations delimit the scope of the study and should be kept in view when evaluating its claims.

The first concerns design. The study was framed as a case study with an embedded comparison between two intact classes rather than as a randomized experiment. Baseline equivalence reduces the most obvious concern about pre-existing difference, but class-level influences cannot be eliminated with the same confidence as in a fully experimental design.

The second concerns context. The research was conducted in one public primary school with one teacher. This supports procedural consistency, but outcomes may vary in classrooms with different institutional cultures, technological conditions, learner profiles, or teaching styles.

The third concerns duration. A six-week intervention is sufficient for examining short-term motivational change during the review phase, but it is not long enough to establish whether the pattern would remain stable over a semester or after novelty diminishes.

The fourth concerns measurement. Motivation is complex and multidimensional. Although the questionnaire, observations, and interviews provide triangulated evidence, they capture classroom readiness to participate during review lessons rather than an exhaustive account of all motivational processes or of subsequent language achievement.

The fifth concerns qualitative compression. Observation records and interview recordings were condensed into lesson summaries and structured interview summaries for analysis and reporting. This supported manageability and age-appropriate fieldwork, but it also means that some interactional detail is necessarily represented in summarized rather than fully transcribed form.

1.10 Definition of Key Terms

Baamboozle refers to a web-based classroom game platform that allows teachers to present review questions through team-based tile selection and scoring.

Gamification refers to the use of game-like elements such as points, turns, teams, and immediate feedback in activities that are not originally games (Deterding et al., 2011).

Classroom motivation refers to learners' readiness and willingness to participate during English review lessons, operationalized through the SDT dimensions of autonomy, competence, and relatedness.

Autonomy refers to the learner's sense of volition, involvement, and meaningful participation.

Competence refers to the learner's sense of capability, success, and recoverability after error.

Relatedness refers to the learner's sense of belonging, peer connection, and social support in the classroom.

Young learners refers here specifically to Grade 2 primary school students, approximately seven to eight years old.

1.11 Organization of the Thesis

Chapter 1 introduced the study, its rationale, research problem, questions, objectives, significance, scope, and limitations. Chapter 2 reviewed the literature and developed the theoretical and conceptual framework. Chapter 3 explained the mixed-methods case study methodology, including participants, instruments, procedures, analyses, and ethical safeguards. Chapter 4 presented the quantitative, qualitative, and mixed-methods findings. Chapter 5 interpreted the findings through SDT, compared them with prior literature, discussed contributions and implications, acknowledged limitations, and concluded the thesis. The appendices contain the questionnaire, observation checklist, interview protocols, consent materials, lesson plans, raw data tables, coding matrix, and interview summaries.

CHAPTER 2: LITERATURE REVIEW

2.1 Theoretical Framework: Self-Determination Theory

2.1.1 Core propositions of SDT

Self-Determination Theory is especially useful for this thesis because it distinguishes not only how much motivation learners show, but also what kind of motivation is being supported in classroom activity (Deci & Ryan, 1985, 2000; Noels et al., 2000; Ryan & Deci, 2018). From an SDT perspective, visible participation alone is not sufficient evidence of high-quality motivation, since learners may comply because of pressure, fear, routine, or external control, or they may participate because the activity is experienced as meaningful, manageable, and socially supportive (Deci & Ryan, 2000; Reeve, 2015; Ryan & Deci, 2018). This distinction is crucial in primary EFL classrooms, where outward busyness can easily be mistaken for genuine motivational support, particularly when young learners' engagement is strongly shaped by classroom affect, perceived competence, teacher support, and peer interaction (Cameron, 2001; Pinter, 2017; Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017; Bui et al., 2025).

At the center of SDT are three basic psychological needs: autonomy, competence, and relatedness (Deci & Ryan, 1985, 2000; Ryan & Deci, 2018). Autonomy refers to experiencing one's participation as volitional rather than wholly controlled; competence refers to feeling capable of acting effectively and making progress; and relatedness refers to feeling socially connected, respected, and supported by significant others in the learning environment (Reeve, 2015; Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017; Ryan & Deci, 2018). These needs are analytically powerful because they move the discussion beyond diffuse labels such as enjoyment, fun, or interest and direct attention to the classroom processes that make sustained participation more or less likely, including meaningful choice, visible success, supportive feedback, and socially safe participation (Hidi & Renninger, 2006; Lou & Noels, 2017; Reeve, 2015; Ryan & Deci, 2018).

For the present thesis, SDT functioned as the primary theoretical framework rather than as a decorative interpretive label. It defined classroom motivation as the quality of learners' participation in relation to autonomy, competence, and relatedness, and it guided the questionnaire design, observation indicators, interview prompts, and interpretation of Baamboozle-integrated review (Deci & Ryan, 2000; Reeve, 2015; Ryan & Deci, 2018).

2.1.2 SDT in second language motivation research

Within second language learning, SDT has been especially valuable because it clarifies how learners internalize reasons for engaging in language study and how classroom conditions shape the quality of that engagement (Noels et al., 2000). In contrast to models that emphasize broad attitudes or future self-guides, SDT examines whether the learner's immediate experience of the classroom supports volition, efficacy, and social belonging. That concern is especially relevant in language learning, where participation often involves public performance, uncertainty, and the risk of visible error.

Noels et al. (2000) showed that SDT can explain why some forms of language learning motivation are more self-determined than others, and later work has continued to argue that the theory offers a productive lens for understanding autonomy in language education without reducing autonomy to simple independence or isolated self-study (Lou & Noels, 2017). For the present study, that distinction is important. The autonomy relevant to Grade 2 learners is not the autonomy of independently planning a curriculum or self-directing long-term language goals. It is the more immediate classroom experience of having a meaningful turn, a bounded choice, a chance to contribute, and a sense that participation is invited rather than imposed.

Because the present study defines motivation as classroom readiness and willingness to participate, Self-Determination Theory (SDT) fits more closely than broader trait-like or future-oriented approaches that emphasize stable dispositions or long-term identity development (Dörnyei & Ushioda, 2009; Lamb, 2017; Ryan

& Deci, 2018). SDT captures motivation as something situated, relational, and dynamically shaped by the immediate conditions of classroom activity, including task structure, feedback, social interaction, and perceived success (Deci & Ryan, 2000; Reeve, 2015; Oga-Baldwin et al., 2017; Ryan & Deci, 2018). This perspective is particularly appropriate for young learners, whose engagement is strongly influenced by moment-to-moment classroom experiences rather than by abstract future goals or fully developed L2 self-concepts (Cameron, 2001; Pinter, 2017; Oga-Baldwin & Nakata, 2015).

2.1.3 Why SDT fits young learner EFL classrooms

SDT is particularly appropriate for young learners because it does not depend on sophisticated introspection about distant future goals. Research on language motivation has often been shaped by constructs developed for adolescents and adults, including orientations toward the target language community and imagined future selves (Gardner et al., 1985; Dörnyei & Ushioda, 2009). These frameworks remain influential and useful, but their explanatory power becomes less direct when the participants are seven- or eight-year-old children in routine primary classrooms. Young learners are less likely to organize their engagement around stable identity narratives and more likely to respond to immediate classroom conditions: whether the task is inviting, whether they can answer, whether mistakes feel survivable, and whether classmates function as support or threat.

This developmental point is strongly supported by work on young language learners. Cameron (2001), Butler (2019), Pinter (2017), Carello & Butler (2015) all show that children's language learning is shaped by age-appropriate pedagogy, task design, interactional support, and classroom climate. In other words, motivational analysis among young learners must remain close to what children actually experience in the classroom rather than importing adult-centered constructs without adaptation. SDT is therefore especially suitable because autonomy, competence, and relatedness can be operationalized in developmentally

appropriate ways. A child does not need to articulate an “ideal L2 self” to communicate whether it feels good to have a turn, whether answering is frightening, or whether teammates help.

Research with elementary foreign language learners further strengthens this point. Oga-Baldwin and Nakata (2015) showed that autonomy support in elementary foreign language classes should not be understood as the absence of structure; rather, structure can itself support autonomy when it helps children understand expectations and enter participation successfully. Oga-Baldwin et al. (2017) likewise found that self-determined motivation in elementary foreign language classes is closely tied to classroom experience over time. These studies matter for the present thesis because they help explain two features of the findings simultaneously: first, why Baamboozle could support motivation by increasing turn ownership, visible success, and team support; and second, why the comparison group still improved under a structured teacher-led condition. The relevant contrast is therefore not between freedom and structure, but between different kinds of structure and the different motivational conditions they create.

2.1.4 SDT and the design logic of gamified review

A further reason SDT is central to this thesis is that it provides a disciplined way to evaluate gamification in educational settings (Deci & Ryan, 2000; Ryan & Deci, 2018; Sailer et al., 2017; Dichev & Dicheva, 2017). In everyday educational talk, digital games are often assumed to be motivating because they are fun, competitive, visually attractive, or novel, but research on gamification shows that game elements do not automatically produce meaningful learning motivation (Deterding et al., 2011; Dichev & Dicheva, 2017; Ratinho & Martins, 2023; Wang, 2015). SDT complicates that assumption by shifting attention from surface-level excitement to the psychological quality of participation: a game-like activity is not motivationally beneficial merely because it produces energy or enjoyment, but because its design supports autonomy, competence, and relatedness rather than frustrating them (Deci & Ryan, 2000; Reeve, 2015; Sailer et al., 2017;

Ryan & Deci, 2018). This distinction is particularly important in young learner EFL classrooms, where game-based tools may increase visible participation, yet their educational value depends on whether learners feel able to succeed, safe to answer, included by peers, and meaningfully involved in the activity (Cameron, 2001; Pinter, 2017; Reinders & Wattana, 2015; Oga-Baldwin et al., 2017; Bui et al., 2025).

This insight is important because specific game elements can have different psychological effects depending on how they are used. Sailer et al. (2017) demonstrated that game design elements do not operate as a single motivational package; their effects depend on the psychological needs they support. Likewise, Dichev and Dicheva (2017) argued that the educational literature on gamification often overstates benefits while underexamining mechanism and context. In SDT terms, competition may energize participation for some learners but become controlling or threatening for others; point systems may make progress visible but may also intensify social comparison; choice may appear to support autonomy but remain superficial if all meaningful control stays with the teacher.

This theoretical caution maps directly onto the present thesis's findings. The qualitative strand did not portray Baamboozle as uniformly and unconditionally motivating. Instead, it showed that motivation rose together with constraints involving noise, fairness, pacing, and technological reliability. SDT helps make sense of this pattern by showing that gamification is not effective *per se*. What matters is whether the tool reorganizes participation in need-supportive ways.

2.2 Motivation in Second Language Acquisition

Motivation has long been recognized as one of the most consequential variables in second language learning, but the field has conceptualized it in different ways. Gardner's (1985) socio-educational model highlighted attitudes, orientations, and the learner's relationship to the target language community. Later, Dörnyei (2009) emphasized the future self, proposing that learners are partly

motivated by the person they imagine becoming in the target language. Lamb (2017) similarly showed that motivation research increasingly attends to the motivational dimension of classroom teaching rather than treating motivation as a purely internal trait.

These traditions remain important, but they are not equally useful for every research problem (Gardner et al., 1985; Dörnyei & Ushioda, 2009; Lamb, 2017). For the present thesis, two limitations are especially relevant. First, many major SLA motivation models were developed primarily with older learners or more cognitively mature language learners in mind, whereas Grade 2 pupils' motivation is more immediate, affective, interactional, and classroom-bound (Cameron, 2001; Pinter, 2017; Butler, 2019; Oga-Baldwin et al., 2017). Second, these models often operate at a more global level than the one examined in this study, focusing on orientations, attitudes, language identity, or future self-concepts rather than the micro-conditions of classroom participation (Gardner et al., 1985; Dörnyei & Ushioda, 2009; Lamb, 2017; Noels et al., 2000). The present research did not investigate long-term identity formation, language learning orientations, or broad attitudes toward English; instead, it investigated whether a particular review structure altered students' immediate willingness to join, answer, persist, and participate socially in ordinary lessons. For that reason, SDT provides a closer theoretical match, since it explains how classroom activity can support or frustrate autonomy, competence, and relatedness in ways that directly affect situated participation (Deci & Ryan, 2000; Reeve, 2015; Oga-Baldwin & Nakata, 2015; Ryan & Deci, 2018).

For that reason, the most useful way to conceptualize motivation here is as dynamic, task-sensitive, and interactional. This position does not reject broader SLA motivation theories; rather, it narrows the construct to the level at which the current study actually collected data. Such a narrowing is also methodologically necessary. A short, emoji-based questionnaire for Grade 2 learners cannot validly measure the full complexity of long-range language motives. It can, however,

measure classroom-participation readiness across theoretically meaningful dimensions, especially when triangulated with observation and interview evidence.

2.3 Young Learners in EFL Classrooms

2.3.1 Developmental characteristics and classroom consequences

Young learners differ from adolescent and adult learners in ways that are pedagogically and theoretically consequential. Cameron (2001) and Pinter (2017) argue that children tend to respond more strongly to concrete, visible, rhythmic, and socially mediated activity than to abstract explanation alone. They are also still developing self-regulation, sustained attention, and the ability to verbalize internal states. This does not mean they are less motivated learners; rather, it means their motivation is expressed differently and must be studied differently.

Butler's (2019) review of young learner English education in East Asia similarly shows that classroom motivation among children is closely linked to developmental fit, classroom ecology, and the practical organization of lessons. For primary learners, participation often depends on whether tasks are manageable, emotionally safe, and interactionally clear, since young learners benefit from concrete routines, visible support, and age-appropriate opportunities for successful language use (Cameron, 2001; Butler, 2019; Pinter, 2017). This has direct implications for the present study. The findings of joyful engagement, reduced fear of answering, and heightened peer support are plausible not because children naturally like games, but because the intervention restructured the lesson in ways that matched young learners' need for visible turns, quick feedback, and socially shared activity (Deci & Ryan, 2000; Reeve, 2015; Oga-Baldwin et al., 2017; Ryan & Deci, 2018). In this sense, the motivational value of the intervention lies less in game novelty alone and more in its capacity to create developmentally suitable participation conditions in which Grade 2 learners could feel capable, included, and willing to respond (Butler, 2019; Pinter, 2017; Bui et al., 2025; Sailer et al., 2017).

A further developmental implication concerns repetition. In early language learning, review and recycling are indispensable because young learners require repeated exposure, retrieval, and use before vocabulary and language patterns become stable (Cameron, 2001; Butler, 2019; Pinter, 2017; Yip & Kwan, 2006). Yet repetitive review can become demotivating if it is experienced as monotonous, overly controlled, or unequally distributed, especially when only a few confident pupils repeatedly participate while others remain passive (Deci & Ryan, 2000; Reeve, 2015; Wang, 2015; Ryan & Deci, 2018). This matters because the present study targeted the review phase rather than the entire lesson. The literature suggests that, for young learners, the emotional design of routine practice is not peripheral to learning; it is part of the learning environment itself, shaping whether repetition becomes boring compliance or supportive, confidence-building participation (Hidi & Renninger, 2006; Oga-Baldwin et al., 2017; Reinders & Wattana, 2015; Bui et al., 2025).

2.3.2 Participation as socially mediated and emotionally vulnerable

Oga-Baldwin and Nakata (2015) showed that children's experience of autonomy is intertwined with classroom structure, while Oga-Baldwin et al. (2017) highlighted the importance of engagement over time in elementary foreign language classes. These studies help explain why relatedness emerged so clearly in the present findings.

This issue also explains why the thesis does not define motivation as private liking or enjoyment alone. In a primary classroom, a child may enjoy the lesson and still avoid answering if the moment feels risky, especially when participation is associated with possible error, embarrassment, or loss of face (Cameron, 2001; Pinter, 2017; Carello & Butler, 2015; Lou & Noels, 2017). Conversely, a child may attempt an answer with only partial certainty if peers and teacher feedback make the risk manageable and if the activity provides enough structure for success to feel attainable (Deci & Ryan, 2000; Reeve, 2015; Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017). Motivation must therefore be understood as a condition

of participation, not just as a favorable attitude, because young learners' willingness to respond is shaped by perceived competence, emotional safety, classroom relationships, and the immediate social organization of the task (Ryan & Deci, 2018; Butler, 2019; Bui et al., 2025; Reinders & Wattana, 2015).

2.4 Measuring Motivation in Young Learners

Studying motivation in young learners requires methodological adaptation, especially because children's responses may be shaped by developmental, linguistic, and classroom factors that differ from those of older learners (Cameron, 2001; Pinter, 2017; Butler, 2019). Standard verbal Likert scales assume a level of abstract discrimination and linguistic processing that may exceed what seven- or eight-year-olds can do reliably in school-based data collection, particularly when they are asked to report internal states such as confidence, enjoyment, willingness, or perceived support (Mellor & Moore, 2014; Cohen et al., 2013; Creswell & Creswell, 2018). Mellor and Moore (2014) demonstrated that children's use of response scales requires careful design attention, especially when the target construct is subjective rather than objectively verifiable. This point supports the present study's use of an emoji-based three-point questionnaire with oral support, as the simplified visual format reduces unnecessary linguistic burden while still allowing Grade 2 learners to indicate motivational responses in a developmentally accessible way (Mellor & Moore, 2014; Dörnyei, 2007; Cohen et al., 2013).

Such adaptation, however, must not be confused with theoretical looseness. Simplified instruments are methodologically appropriate only when the underlying construct is defined clearly and when the instrument is not asked to do more than it can reasonably do (Cohen et al., 2013; Dörnyei, 2007; Creswell & Creswell, 2018; Mellor & Moore, 2014). In the present study, that meant two things. First, motivation was operationalized narrowly as classroom readiness and willingness to participate through autonomy, competence, and relatedness, rather than as a broad or fully internal psychological state (Deci & Ryan, 2000; Reeve, 2015; Ryan & Deci, 2018). Second, questionnaire data were triangulated with classroom

observations and interviews rather than treated as sufficient in isolation, which is especially important when researching young learners whose self-report responses may need contextual interpretation (Cohen et al., 2013; Dörnyei, 2007; Creswell & Creswell, 2018; Braun & Clarke, 2006).

This methodological position is consistent with the wider literature. When research concerns very young learners, construct validity depends not only on psychometric coherence but also on developmental accessibility and evidential triangulation (Mellor & Moore, 2014; Cohen et al., 2013; Creswell & Creswell, 2018). The present thesis's measurement design therefore aligns with both the age group and the study's explanatory aims. It allows the thesis to detect broad motivational movement while using qualitative data to clarify how that movement was actually produced in practice, especially through classroom interaction, perceived success, peer support, and teacher-mediated participation (Dörnyei, 2007; Braun & Clarke, 2006; Oga-Baldwin et al., 2017; Bui et al., 2025).

2.5 Gamification and Game-Based Learning in TESOL

2.5.1 Conceptual distinctions

Gamification and game-based learning are related but not identical. Deterding et al. (2011) define gamification as the use of game design elements in non-game contexts. Game-based learning, by contrast, refers more broadly to learning through games or through structured participation in game environments (Prensky, 2001; Reinders & Wattana, 2015; Yip & Kwan, 2006). This distinction is important because the educational claims made for each are not the same. Full digital games may provide immersion, narrative, and complex interaction, whereas gamified classroom activities often insert selected game elements such as points, teams, timed response, visible progress, and feedback into otherwise conventional lessons (Deterding et al., 2011; Dichev & Dicheva, 2017; Sailer et al., 2017; Ratinho & Martins, 2023).

Baamboozle fits more closely within gamified classroom review than within fully immersive game-based learning. It does not replace the curriculum,

nor does it create an autonomous virtual world. Rather, it reorganizes ordinary vocabulary and grammar review through tile selection, team turns, scoring, and immediate feedback, which are the kinds of game design features commonly associated with classroom gamification rather than full game-based learning (Deterding et al., 2011; Sailer et al., 2017; Wang, 2015; Saud et al., 2022). This conceptual clarification matters because the present thesis is not testing whether game worlds improve language learning in general. It is testing whether a light-touch, teacher-mediated review format can alter motivation under ordinary classroom conditions, especially by making participation more visible, socially shared, and feedback-rich for young EFL learners (Amanah et al., 2025; Azzahra & Kembaren, 2025; Vasya et al., 2025; Zulbani & Syafryadin, 2025).

2.5.2 Why gamified review may support motivation

There are several reasons why gamified review can be motivationally effective. First, it can make routine retrieval more eventful by converting repeated questioning into a sequence of anticipated turns (Yip & Kwan, 2006; Wang, 2015; Saud et al., 2022). Second, it can increase perceived competence through immediate answer confirmation, visible scoring, and rapid movement to the next opportunity (Sailer et al., 2017; Reeve, 2015; Ryan & Deci, 2018). Third, it can widen relatedness when team structures make responses socially shared rather than individually exposed (Deci & Ryan, 2000; Oga-Baldwin et al., 2017; Reinders & Wattana, 2015). Fourth, it can support autonomy in limited but meaningful ways by giving learners visible entry points, turn ownership, and bounded choice within a structured classroom task (Reeve, 2015; Oga-Baldwin & Nakata, 2015; Sailer et al., 2017; Ryan & Deci, 2018).

These mechanisms are consistent with both SDT and language-learning research. Yip and Kwan (2006) found that online vocabulary games could make otherwise repetitive lexical work more attractive, while Reinders and Wattana (2015) showed that digital game-based environments could alter learners' affective relationship to participation and willingness to communicate. Sailer et al. (2017)

further argued that specific game elements can support psychological need satisfaction rather than functioning as a single undifferentiated motivational package. Taken together, these studies suggest that the relevant educational value of gamified review lies less in entertainment than in the way it restructures attention, risk, competence feedback, and social interaction (Deterding et al., 2011; Dichev & Dicheva, 2017; Ratinho & Martins, 2023; Ryan & Deci, 2018).

2.5.3 Critical debates and contradictory evidence

At the same time, the literature on gamification is far from uniformly celebratory. Dichev and Dicheva (2017) argue that claims about educational gamification often move faster than the evidence, with many studies emphasizing short-term engagement without adequately demonstrating deeper or more durable motivational change. Wang (2015) similarly showed that positive reactions to game-based classroom response systems may weaken over time, which means that early enthusiasm cannot automatically be read as sustained motivational transformation. More recently, Ratinho and Martins (2023) concluded from a systematic review that gamified learning strategies often have positive effects on motivation, but that results vary by design, context, and duration.

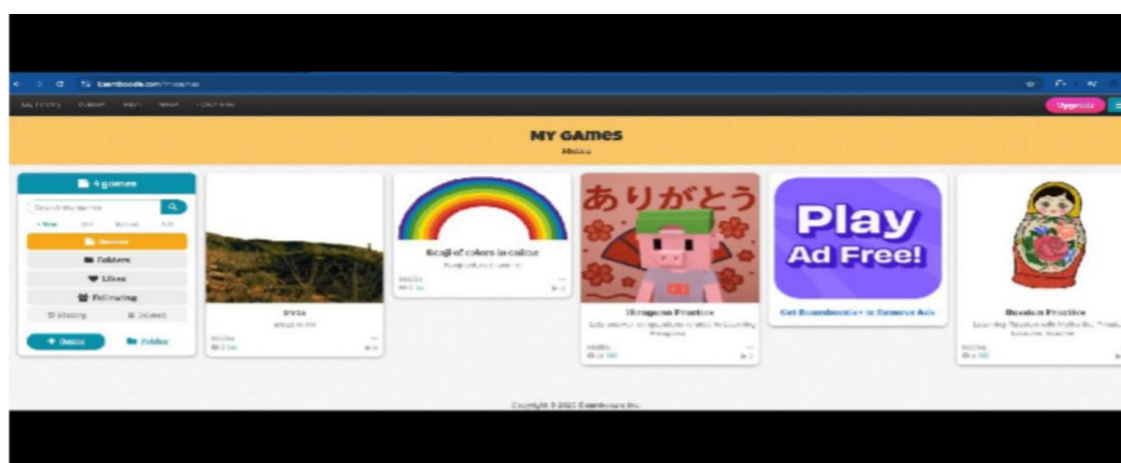
These critiques are especially relevant to TESOL, where participation, language use, confidence, and classroom interaction matter as much as visible activity (Cameron, 2001; Pinter, 2017; Butler, 2019; Lamb, 2017). A gamified activity may generate energy while still privileging already confident learners, especially if turn-taking, feedback, and peer support are not carefully managed (Dichev & Dicheva, 2017; Wang, 2015; Reeve, 2015; Oga-Baldwin et al., 2017). It may increase speed at the cost of reflection, depend heavily on novelty or teachers' labor-intensive preparation, or shift students' attention toward winning rather than language use if poorly designed (Deterding et al., 2011; Dichev & Dicheva, 2017; Wang, 2015; Ratinho & Martins, 2023). For the present study, these debates are not side issues. They are precisely what make the mixed findings meaningful. The Baamboozle condition was more motivating overall, yet its

benefits remained conditional and teacher-mediated. The literature therefore supports a critical reading of the findings: Baamboozle appears pedagogically promising, but its value depends on how it is implemented and on which motivational processes are actually being supported (Sailer et al., 2017; Deci & Ryan, 2000; Reeve, 2015; Ryan & Deci, 2018; Azzahra & Kembaren, 2025; Saud et al., 2022).

2.6 Baamboozle as a Classroom Review Tool

Baamboozle is pedagogically interesting because it occupies a practical middle ground (Safrizal et al., 2026). It is simple enough to be used in projector-based whole-class teaching and flexible enough to incorporate textbook-linked review content. That matters in contexts where one-device-per-student assumptions are unrealistic or where teachers must integrate digital activity into established lesson routines rather than redesign the curriculum completely. In other words, Baamboozle is not educationally significant because it is technologically sophisticated. It is significant because it can modify participation structures without demanding a highly resourced classroom. Figure 2.1 illustrates the tile-based team review format that makes Baamboozle suitable for teacher-mediated, whole-class review in lower-primary classrooms.

Figure 2.1. Baamboozle interface



This feature aligns closely with the present study's design and findings. The intervention did not change what learners studied; it changed how the review

stage was organized. That distinction is theoretically important because it narrows the claim. Any motivational difference found in the study is more plausibly explained by the structure of participation than by differences in syllabus content (Deterding et al., 2011; Sailer et al., 2017; Deci & Ryan, 2000; Ryan & Deci, 2018). This is one reason the platform fits the thesis well: it allows the research to focus on review as a pedagogical phase and on motivation as a classroom-process variable, rather than treating gamification as a general replacement for ordinary instruction (Dichev & Dicheva, 2017; Wang, 2015; Ratinho & Martins, 2023).

At the same time, Baamboozle should not be treated as self-explanatory. Its potential advantages are inseparable from teacher orchestration. Turn distribution, pacing, question difficulty, response time, error handling, and team composition all determine whether the tool widens participation or simply accelerates competition (Reeve, 2015; Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017; Sailer et al., 2017). This point anticipates the present study's qualitative finding that motivation increased together with new management demands. The platform is therefore best understood not as an independent causal agent but as a structured affordance whose effects depend on pedagogical use, classroom management, and the extent to which autonomy, competence, and relatedness are actively supported during play (Deci & Ryan, 2000; Ryan & Deci, 2018; Azzahra & Kembaren, 2025; Saud et al., 2022; Vasya et al., 2025).

Recent classroom studies have generally reported favorable outcomes for Baamboozle in vocabulary review, grammar practice, classroom participation, and learner perceptions in EFL and school-based contexts (Saud et al., 2022; Sakdiyah et al., 2024; Walidaina et al., 2024; Azzahra & Kembaren, 2025; Zulbani & Syafryadin, 2025). However, this literature remains uneven because many studies focus on achievement or general perceptions and give less analytic attention to younger learners, theoretically specified motivation, and classroom implementation constraints. For the present thesis, these studies justify Baamboozle as a relevant review tool while also requiring a cautious design that

examines how the platform reshapes participation, confidence, and peer support rather than merely whether students report enjoyment.

2.7 Empirical Studies on Digital Games, Motivation, and Language Learning

The wider empirical literature on digital games and game-like platforms suggests a broadly positive but methodologically uneven pattern. Studies across language and educational contexts often report increased attention, participation, enjoyment, or willingness to communicate when game-like structures are introduced (Yip & Kwan, 2006; Reinders & Wattana, 2015). Reviews of widely used classroom response tools, especially Kahoot!, also report positive effects on classroom dynamics, perceptions of learning, and learner attitudes. These findings make it plausible that a tool like Baamboozle could enhance motivation during repetitive review work.

However, the literature is less consistent when the question becomes more specific. First, many studies collapse motivation, engagement, enjoyment, and satisfaction into a single positive narrative, which makes explanation difficult (Dichev & Dicheva, 2017; Ratinho & Martins, 2023; Sailer et al., 2017). Second, many interventions are brief, which makes it hard to distinguish sustained motivational support from novelty effects (Wang, 2015; Dichev & Dicheva, 2017; Ratinho & Martins, 2023). Third, many studies rely heavily on post-use perceptions or on one-group designs without active comparison conditions, limiting the strength of causal or explanatory claims (Cohen et al., 2013; Creswell & Creswell, 2018; Dörnyei, 2007). Fourth, language-learning studies often focus on older learners, especially secondary students, university students, or adult learners, leaving younger primary EFL learners comparatively less represented (Butler, 2019; Cameron, 2001; Pinter, 2017; Oga-Baldwin et al., 2017).

This pattern is methodologically important for the present thesis. It means that the strongest unanswered questions are not whether digital classroom games can feel motivating in general, but under what conditions, for which learners, and

through which mechanisms (Dichev & Dicheva, 2017; Sailer et al., 2017; Ratinho & Martins, 2023; Ryan & Deci, 2018). The present study addresses exactly these questions by combining a comparison-class design with qualitative evidence and by interpreting the outcome through SDT rather than through generic positive-attitude language (Deci & Ryan, 2000; Braun & Clarke, 2006; Creswell & Creswell, 2018; Reeve, 2015; Ryan & Deci, 2018). In that sense, the study extends the empirical literature by narrowing and strengthening the explanatory claim: Baamboozle is not treated as motivational simply because it is enjoyable or digital, but because its teacher-mediated review structure may support autonomy, competence, and relatedness under ordinary Grade 2 EFL classroom conditions (Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017; Sailer et al., 2017; Azzahra & Kembaren, 2025; Bui et al., 2025).

2.8 Vietnamese and Asian EFL Contexts

The Asian EFL literature provides a useful but incomplete regional backdrop for the present study. Butler (2019) shows that English education among young learners in East Asia must be interpreted with sensitivity to policy, pedagogy, and age-appropriate classroom practice rather than through imported assumptions alone. This regional caution is important because digital pedagogy does not operate in a vacuum. Classroom authority, participation norms, access to technology, and curricular expectations all shape how game-like activities function.

Vietnamese research on English education for young learners generally presents an encouraging but uneven picture: primary pupils tend to show positive attitudes toward learning English, and classroom innovations such as cartoons, stories, and digital storytelling have been reported to support speaking, listening, grammar learning, and classroom engagement (Nhun & Dieu, 2020; Nguyen, 2021; Tham & Thuy, 2023). Studies in Vietnam also suggest that exposure to English beyond school is associated with stronger listening and reading performance among primary learners (Thi & Thinh, 2023). At the same time,

research repeatedly notes that the success of young learner English education depends less on learner interest than on implementation conditions, since teachers and schools still face constraints related to policy enactment, professional support, materials, infrastructure, and uneven rural–urban conditions (Bui et al., 2025; Phan et al., 2024; Phuong, 2025). Overall, the Vietnamese literature suggests that young learner English education has strong potential and is becoming more interactive and learner-centered, but sustained improvement requires better teacher preparation, stronger institutional support, and more realistic responses to local classroom conditions (Bui et al., 2025; Phan et al., 2024; Phuong, 2025).

These contextual studies are directly relevant to the present thesis for two reasons. First, they justify the focus on a low-barrier, projector-based tool rather than on individualized device-rich learning, since Vietnamese primary EFL classrooms may still face uneven ICT access, teacher-mediated lesson structures, and practical constraints in technology integration (Phuong, 2025; Bui et al., 2025; Phan et al., 2024; Tham & Thuy, 2023). Second, they support the thesis refusal to treat technological intervention as context-free. In Vietnamese primary classrooms, the pedagogical value of a digital tool depends not only on its design but also on whether it fits teacher-led instruction, fixed class groupings, limited time, curriculum routines, and uneven technological reliability (Phuong, 2025; Nguyen, 2021; Nhung & Dieu, 2020; Thi & Thinh, 2023). The study’s discussion of noise, pacing, turn equity, and technology constraints therefore emerges not as an accidental by-product but as a contextually expected part of implementation, especially when gamified review is used with young learners in real classroom conditions rather than in idealized digital environments (Cameron, 2001; Pinter, 2017; Dichev & Dicheva, 2017; Wang, 2015; Reeve, 2015).

2.9 Research Gap

The reviewed literature indicates that motivation is central to second-language learning, but it also shows that motivation among Grade 2 EFL learners must be conceptualized as an immediate classroom-participation construct rather

than as a distant future-oriented disposition. Major L2 motivation frameworks have clarified attitudes, effort, and self-related orientations (Gardner et al., 1985; Dörnyei & Ushioda, 2009; Lamb, 2017), yet these constructs do not fully explain how seven- or eight-year-old learners decide whether to answer, take risks, and remain involved during routine review lessons (Cameron, 2001; Pinter, 2017; Butler, 2019).

The first gap is therefore developmental and theoretical. Existing young-learner EFL studies acknowledge that children need age-appropriate, concrete, and emotionally safe pedagogy, but fewer studies explain motivation through a mechanism that links task design to autonomy, competence, and relatedness. SDT addresses this limitation by specifying why bounded choice, visible success, and peer-supported participation may make learners more willing to enter and sustain classroom activity (Deci & Ryan, 2000; Oga-Baldwin & Nakata, 2015; Oga-Baldwin et al., 2017; Ryan & Deci, 2018).

The second gap is methodological. Studies of gamification and digital games often report positive enjoyment or engagement, but many rely on brief interventions, single-group designs, or post-use perceptions that do not clearly separate novelty from motivational support (Dichev & Dicheva, 2017; Wang, 2015; Ratinho & Martins, 2023). For classroom TESOL research, stronger evidence requires comparison, developmentally appropriate measurement, observation of classroom process, and attention to contradictory or conditional cases.

The third gap is contextual and tool-specific. Baamboozle has been increasingly reported as useful for vocabulary, grammar, and review activities, but it remains underexamined as a theory-linked motivational review structure for very young EFL learners in Vietnamese public primary classrooms. Existing studies often focus on older learners, achievement outcomes, or general perceptions, whereas the present study examined how a low-barrier, projector-based tool reorganized the review stage under ordinary classroom constraints.

The present study addressed these gaps by investigating Baamboozle-integrated English review in two intact Grade 2 classes at Le Quy Don Primary School. Its contribution is not a broad claim that digital games are inherently motivating, but a bounded explanation of whether and how a gamified review format may support classroom motivation through autonomy, competence, and relatedness in a Vietnamese lower-primary EFL setting.

2.10 Conceptual Framework

Based on the reviewed literature, the present study is guided by a conceptual framework that connects Baamboozle-integrated English review with young learners' classroom motivation through the three psychological needs proposed by Self-Determination Theory: autonomy, competence, and relatedness. The framework begins from the assumption that Baamboozle does not improve motivation simply because it is digital, colorful, competitive, or entertaining. Rather, its motivational value depends on how it changes the structure of classroom participation during the review stage of English lessons.

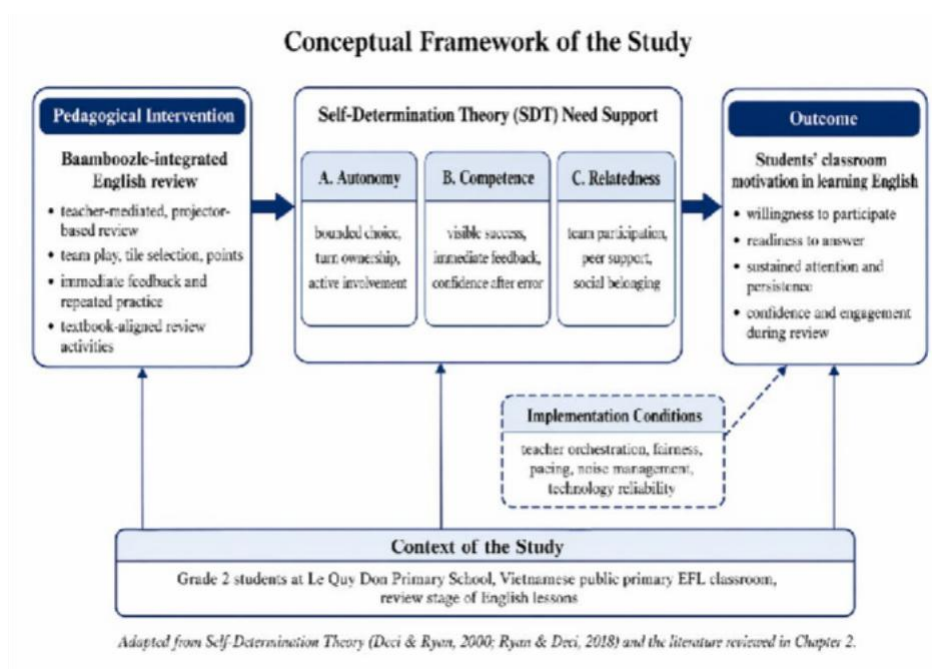
In the study, Baamboozle-integrated review refers to the use of Baamboozle during the review stage of Grade 2 English lessons. The review content remains aligned with the textbook and lesson objectives, but the format of review changes. Instead of relying mainly on teacher-fronted questioning, repetition, board work, or worksheet-based consolidation, the Baamboozle condition organizes review through team participation, tile selection, points, immediate feedback, and turn-based answering. These features are expected to influence learners' classroom motivation indirectly by supporting autonomy, competence, and relatedness. First, Baamboozle may support autonomy by giving students bounded but meaningful participation opportunities. Grade 2 learners are not expected to control the lesson independently, but they may experience autonomy when they are invited to choose tiles, wait for their team's turn, contribute an answer, and feel that they have an active role in the lesson. In this sense, autonomy is understood developmentally. It refers not to full independence but to meaningful involvement within a structured

classroom activity. Second, Baamboozle may support competence by making success more visible and feedback more immediate. During ordinary review, hesitant learners may avoid answering because mistakes are public and emotionally risky. In Baamboozle, however, the game structure may make answering feel more manageable because learners respond as part of a team, receive quick confirmation, and move rapidly to the next opportunity. Points, correct-answer feedback, and repeated review questions can help learners feel that English is something they can attempt and succeed in, rather than something only confident classmates can perform. Third, Baamboozle may support relatedness by transforming review into shared participation. Team play allows learners to answer with peer support rather than as isolated individuals. This may reduce fear of mistakes and increase the sense that classmates are partners in the activity. For young learners, this social dimension is especially important because classroom motivation is strongly shaped by peer presence, teacher support, and the emotional safety of participation.

The expected outcome of this framework is stronger classroom motivation in learning English. In this study, motivation refers specifically to students' readiness and willingness to participate during English review lessons. It includes attention, willingness to answer, confidence, persistence after mistakes, enjoyment of participation, and involvement with peers. The framework also recognizes that the effect of Baamboozle is not automatic. Teacher mediation remains essential. The motivational value of the activity may be strengthened or weakened by classroom management, fairness in turn-taking, question difficulty, pacing, noise control, team organization, and technology reliability. Therefore, the conceptual framework positions Baamboozle as a classroom participation structure rather than as a direct technological cause of motivation. Baamboozle-integrated review is expected to support motivation when its game elements are pedagogically managed in ways that support autonomy, competence, and relatedness. This framework guides both the quantitative and qualitative strands of the study. RQ1

examines whether the Baamboozle group shows stronger motivation than the teacher-centered comparison group. RQ2 examines how students and the teacher experience the Baamboozle review format, especially in relation to participation, confidence, peer support, and classroom conditions. Figure 2.2 presents the conceptual framework of the study.

Figure 2.2. Conceptual Framework of the Study



The conceptual framework guiding the study is deliberately narrow and aligned with the thesis variables, instruments, and findings. The independent variable is review format, operationalized as Baamboozle-integrated team-based review in the focal class and teacher-centered review in the comparison class. The dependent variable is student motivation during English review lessons, operationalized quantitatively as total motivation score and qualitatively as classroom participation readiness, confidence, and peer-supported involvement. The mediating classroom logic of the framework is grounded in SDT. First, Baamboozle may support autonomy by increasing bounded choice, turn visibility, and perceived involvement. In this study, autonomy does not mean unrestricted freedom; it refers to learners' sense that they can enter the activity meaningfully. Second, Baamboozle may support competence by providing immediate feedback,

repeated response opportunities, and a format in which mistakes are more recoverable. Third, Baamboozle may support relatedness by embedding response in teams, encouraging peer prompting, and reducing solitary exposure during answer attempts.

The framework also includes conditions that may weaken or distort these advantages. If pacing becomes too rapid, competence may suffer for slower learners. If a few pupils dominate turns, autonomy and fairness may decline for others. If competition overwhelms cooperation, relatedness may weaken. If technical interruption breaks activity flow, attentional benefits may be reduced. The framework therefore anticipates the possibility of qualified rather than absolute positive effects, which aligns closely with the present findings and strengthens the coherence between theory, design, and interpretation.

2.11 Chapter Summary

This chapter has established the theoretical and empirical foundation for the thesis in a way that is tightly aligned with the study's actual findings and discussion. It has argued that SDT provides the most appropriate framework for analyzing motivation in this study because the intervention operated through classroom participation structures and because the observed differences were strongest in autonomy, competence, and relatedness. It has shown that broader SLA motivation models remain valuable but are less directly suited to Grade 2 learners and to the present study's focus on immediate classroom participation. It has also demonstrated that young learner motivation must be treated as developmental, social, and methodologically sensitive.

The review of gamification and digital language-learning research suggests that game-like review can support motivation, but that such support is neither automatic nor uniform. Existing literature points to benefits in attention, participation, and confidence, while also warning against novelty effects, overgeneralized claims, and poorly designed competition. Vietnamese and regional studies further show that pedagogical value is inseparable from local

classroom ecology. These lines of literature converge on the need for exactly the kind of study conducted in this thesis: a mixed-methods, SDT-informed case study of a low-barrier gamified review tool with embedded comparative quantitative evidence from a Vietnamese primary school context.

CHAPTER 3: METHODOLOGY

3.1 Research Paradigm and Case Study Design Logic

The thesis is positioned within a pragmatic mixed-methods paradigm because its central aim is explanatory adequacy rather than methodological exclusivity. The study asks two kinds of question that no single method can answer fully on its own: whether motivation changed differentially across two naturally occurring classroom conditions, and how that pattern was experienced and produced in classroom practice. A pragmatic stance is therefore appropriate because it authorizes the use of quantitative and qualitative evidence in relation to the demands of the research problem rather than to an a priori commitment to one epistemological tradition (Cohen et al., 2013).

The study is framed as a mixed-methods case study. The case boundary was defined as one public primary school, one grade level (Grade 2), two intact classes taught by the same English teacher, and a six-week period in which the review stage of English lessons was organized in two different ways. This bounded framing is central to the thesis because the purpose was to understand how Baamboozle functioned in one real school setting rather than to claim universal effectiveness across schools.

Within that bounded case, a quantitative pre-test/post-test comparison between the two intact classes was embedded to provide structured comparative evidence. The classes remained in their normal school organization, and in consultation with the school and classroom teacher one intact class was designated as the Baamboozle class while the other served as the teacher-centered comparison class. No student-level random assignment occurred. The qualitative strand was incorporated as an explanatory necessity: statistical differences could establish pattern and magnitude, but they could not by themselves show how children entered the activity, how error was experienced, how peers functioned during review, or why some cases remained unchanged or contradictory. The design is therefore best understood as a convergent mixed-methods case study with

embedded quantitative comparison and qualitative process explanation (Creswell & Creswell, 2018).

3.2 Alignment Chain of the Study

Table 3.1 Alignment chain of the study

Element	Specification
Problem	Low and uneven motivation during conventional Grade 2 English review
Theory	Self-Determination Theory
Independent variable	Type of review instruction
Baamboozle condition	Baamboozle-integrated review
Comparison condition	Teacher-centered review
Dependent variable	Student motivation
Subdimensions	Autonomy, competence, relatedness
Quantitative instrument	12-item emoji-based motivation questionnaire
Qualitative instruments	Observation checklist, student interviews, teacher interview
Quantitative analyses	Descriptive statistics, paired t-tests, independent t-tests, effect sizes
Qualitative analysis	Thematic analysis
Integration	Mixed-methods triangulation interpreted through SDT

This alignment is central to the thesis rigor. Each chapter is built around the same construct definitions and research logic.

3.3 Research Site and Participants

The study was conducted at Le Quy Don Primary School in Lam Dong Province, Vietnam. The site was suitable because it reflected a plausible Vietnamese public-school environment for lower-primary English teaching and because projector-based classroom technology was available. The case boundary therefore combined place, grade level, classroom organization, teacher, and

intervention period rather than treating the school only as a convenient access point.

The participant pool consisted of 60 Grade 2 students divided into two intact classes. The Baamboozle class included 30 students, and the teacher-centered comparison class included 30 students. The learners were approximately seven to eight years old. The same English teacher taught both classes and participated in a post-intervention interview. Retaining the same teacher across the two classes reduced teacher-effect variability and strengthened the comparability of the embedded quantitative strand.

For the qualitative strand, 14 students were interviewed: eight from the Baamboozle class and six from the comparison class. Purposive sampling was used in order to represent different participation profiles rather than only the most vocal learners. Sampling considered visible participation patterns across the observed lessons so that more active, moderately active, and quieter pupils were represented. The teacher interview added a cross-class perspective that the student interviews alone could not provide.

3.4 Variables and Operationalization

To ensure conceptual and methodological alignment, the key variables of the study were operationally defined before data collection and analysis. The independent variable was Baamboozle-integrated English review, while the main dependent variable was students' classroom motivation in learning English. In line with Self-Determination Theory, motivation was examined through three dimensions: autonomy, competence, and relatedness. Table 3.2 presents the operational definitions, indicators, instruments, and data sources used to measure or explore each variable in the study.

Table 3.2 Variables and operationalization

Variable / construct	Type	Operational definition	Indicators	Data source
Baamboozle-	Independent	Review	Tile	Lesson plans,

integrated English review	variable	lessons using Baamboozle through teacher-mediated, projector-based, team activities.	selection, team play, points, immediate feedback, repeated practice.	observation, interviews.
Teacher-centered review	Comparison condition	Regular review led mainly by the teacher without Baamboozle.	Teacher questioning, repetition, board work, oral review.	Observation, lesson records.
Classroom motivation	Dependent variable	Students' readiness and willingness to participate in English review lessons.	Attention, participation, confidence, persistence, willingness to answer.	Motivation questionnaire, observation, interviews.
Autonomy	SDT dimension	Students' sense of involvement and meaningful participation during review.	Turn ownership, bounded choice, active involvement.	Questionnaire, observation, interviews.

Competence	SDT dimension	Students' sense of ability and confidence when answering review questions.	Visible success, confidence, feedback, persistence after mistakes.	Questionnaire, observation, interviews.
Relatedness	SDT dimension	Students' sense of peer support and belonging during review.	Team support, cooperation, encouragement, reduced embarrassment.	Questionnaire, observation, interviews.
Implementation conditions	Contextual factor	Classroom factors affecting how Baamboozle was used.	Teacher control, fairness, pacing, noise, technology reliability.	Observation, teacher interview.

As shown in Table 3.2, classroom motivation was operationalized through autonomy, competence, and relatedness. This helped the study examine not only whether Baamboozle increased motivation, but also how it supported participation, confidence, and peer connection during English review lessons.

Total questionnaire scores ranged from 12 to 36. Higher scores represented stronger reported motivation. Subdimension scores were computed by summing the four items assigned to each SDT need: Items 1–4 for autonomy, Items 5–8 for

competence, and Items 9–12 for relatedness. Each subdimension therefore ranged from 4 to 12, while the total score represented the sum of all 12 items.

3.5 Instruments

3.5.1 Motivation questionnaire

The primary quantitative instrument was a 12-item emoji-based motivation questionnaire developed specifically for this study and aligned directly with the Self-Determination Theory dimensions of autonomy, competence, and relatedness. The instrument was constructed to capture classroom motivation as immediate participation readiness during English review lessons rather than as long-term language attitude. Four items addressed autonomy, four addressed competence, and four addressed relatedness, thereby ensuring that the theoretical framework and the operational measure referred to the same construct architecture.

Students responded through a three-point visual format: 1 = unhappy face / not true for me, 2 = neutral face / sometimes true for me, and 3 = happy face / true for me. This simplified response structure was methodologically appropriate for Grade 2 learners because it reduced linguistic and cognitive burden while preserving enough discrimination to identify directional differences across learners and groups. The instrument was administered with oral support in Vietnamese so that the task measured children’s classroom feelings more than their reading ability. In this respect, the questionnaire was designed to be developmentally accessible without abandoning theoretical precision.

The full instrument appears in Appendix A. Its role in the study was not to provide an exhaustive psychometric representation of motivation, but to generate a construct-aligned quantitative index suitable for repeated measurement with very young participants. That bounded purpose is important: the questionnaire was intended to work alongside observation and interview data, not in isolation from them.

3.5.2 Classroom observation checklist

The classroom observation checklist was designed to capture visible behavioral indicators of motivation during the review phase of lessons. Because young children do not always verbalize their classroom states reliably or at length, observational evidence was needed to document how motivation appeared in action. All 24 lesson observations were conducted by the researcher while the class teacher taught. A single-observer approach was used throughout so that the same operational definitions and recording routines would be applied consistently across the case.

- *Volunteer or hand-raise episodes during the review phase*
- *Peer-help utterances or prompts*
- *Visible positive-affect episodes such as smiling, laughter, or excited orientation to the task*
- *Off-task incidents during the review phase*
- *Refusal or hesitation episodes when a response opportunity emerged*
- *Estimated on-task participation rate across the class*
- *Field notes on pacing, turn distribution, fairness, noise, teacher support, and technological interruption*

This combination of count-based and field-note evidence allowed the observation schedule to capture both frequency patterns and interpretive context. Counts were recorded at lesson level for each observed review segment and later averaged across the 12 observed lessons in each class. The on-task participation rate was estimated through repeated whole-class scans during the review segment and converted into a single lesson-level percentage. Refusal or hesitation episodes were retained on the checklist to sensitize observation to reduced entry into participation, but they were not reported as a separate aggregate indicator in the main summary table because brief hesitation was often difficult to distinguish reliably from ordinary wait-time among seven- and eight-year-old learners.

Instead, those episodes informed the qualitative analysis of confidence, fear of answering, and turn accessibility.

3.5.3 Student interview protocol

The student interview protocol was semi-structured, brief, and closely anchored in concrete classroom experience (Dörnyei, 2007). This format was selected because it provided enough flexibility to follow children's responses while still keeping the interviews focused on the study's central concerns, namely participation, confidence, peer support, and feelings during the review stage of English lessons. The questions invited students to talk about what they liked or did not like during review, how they felt when answering in front of others, what their teammates or classmates did during the activity, what usually happened when they did not know an answer, and whether review through Baamboozle felt different from more teacher-centered review. In this way, the interview protocol was designed not to elicit abstract opinions about motivation in general, but to draw out how motivation was experienced through specific and familiar classroom events.

This design choice was deliberate and developmentally appropriate. Because the participants were Grade 2 learners, the use of abstract metacognitive language or highly reflective questioning would likely have exceeded the linguistic and cognitive level of many children and could have reduced the clarity of their responses. For that reason, the protocol approached perception indirectly, using simple prompts linked to recognizable moments in the lesson rather than theoretical or evaluative language. This allowed children to express their feelings and experiences in terms that were accessible to them, while still yielding relevant evidence about autonomy, competence, and relatedness during classroom review. The interviews were conducted individually in Vietnamese, usually in a quiet area near the classroom after the lesson, and typically lasted about five to seven

minutes, which was considered sufficient for obtaining focused responses without causing fatigue or loss of attention.

3.5.4 Teacher interview protocol

The teacher interview explored the teacher's observations of differences between the two classes across several dimensions central to the study, including student participation, attentiveness, confidence, peer interaction, workload, classroom management, and the practical role of Baamboozle in lower-primary English teaching. The interview was designed to elicit not only general impressions of the intervention but also the teacher's comparative judgment about how the two review formats functioned in actual classroom practice. Particular attention was given to whether students appeared more willing to answer, more engaged during the review stage, more supportive of one another, or more hesitant under different instructional conditions. The protocol also invited reflection on implementation issues such as pacing, turn distribution, noise level, preparation demands, and the extent to which Baamboozle was manageable within ordinary school routines.

Because the same teacher taught both classes throughout the intervention period, this interview was analytically important. It provided a cross-class practitioner perspective grounded in direct experience of both instructional conditions, thereby strengthening the study's comparative logic. This perspective was especially valuable because it could speak to patterns that were not fully captured by questionnaire scores alone, such as changes in classroom atmosphere, the practical handling of weaker or quieter learners, and the day-to-day demands involved in sustaining the review activities. The teacher interview therefore served as a key source for triangulating both supportive and qualifying evidence, helping to confirm where the quantitative and observational findings aligned while also identifying conditions, constraints, and reservations that complicated an overly simple positive interpretation. The interview was conducted in Vietnamese after the intervention in a quiet school setting and lasted approximately twenty to

twenty-five minutes, and allowed sufficient time for detailed reflection without making the session unnecessarily prolonged.

3.6 Validity, Reliability, and Trustworthiness

3.6.1 Construct alignment

Construct alignment was established by ensuring that the major instruments in the study were derived from the same conceptual framework that later guided the interpretation of the findings. More specifically, both the motivation questionnaire and the qualitative observation categories were developed in direct relation to Self-Determination Theory, which defines motivation in terms of autonomy, competence, and relatedness. This meant that the study did not treat motivation as a vague or loosely defined sense of enjoyment, excitement, or general classroom positivity. Instead, it identified in advance the particular dimensions of motivation that were theoretically relevant to the research questions and then represented those dimensions systematically in the design of the instruments. In practical terms, the questionnaire items were grouped to reflect the three SDT needs, while the observation schedule was constructed to capture visible classroom behaviors and interactional patterns that could meaningfully indicate those same motivational dimensions during the review phase.

This alignment was important because it reduced the risk of conceptual drift between theory, measurement, and interpretation. If the study had used one understanding of motivation to frame the literature review, another to design the instruments, and a third to discuss the findings, the overall argument would have become less coherent and less convincing. By contrast, in the present study the same construct logic was maintained throughout. Autonomy was reflected in indicators related to meaningful participation and involvement, competence in indicators related to confidence and success, and relatedness in indicators connected to peer support and classroom belonging. As a result, the evidence reported in Chapter 4 was not merely descriptive but was already structured in a way that corresponded to the explanatory claims later developed in Chapter 5. This

strengthened content validity because the study measured what it claimed to investigate, and it also improved interpretive rigor by ensuring that the discussion of findings remained grounded in the same conceptual definitions that shaped the research design from the outset.

3.6.2 Developmental appropriateness

Developmental appropriateness was treated as a validity issue rather than a cosmetic one. With Grade 2 learners, an instrument that is linguistically too dense or cognitively too abstract risks capturing confusion, acquiescence, or task fatigue instead of motivation. For that reason, the questionnaire used short clauses, familiar vocabulary, oral administration support, and visual emoji anchors. The interviews likewise relied on concrete classroom scenarios rather than on abstract self-analysis. These design choices increased the likelihood that children were responding to the intended construct.

3.6.3 Reliability evidence

The scale demonstrated acceptable internal consistency in the present dataset, which supports its use as a short classroom measure of motivation:

Table 3.3 Item-total statistics for the motivation questionnaire

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item - Total Correlation	Squared Multiple Correlation	Cronbach's alpha if Item Deleted
AUT1	26.31	7.842	.612	.421	.848
AUT2	26.24	7.956	.584	.396	.851
AUT3	26.38	8.117	.531	.341	.856
AUT4	26.29	8.264	.497	.298	.859
COM1	26.17	7.733	.645	.458	.846
COM2	26.22	7.895	.601	.404	.849
COM3	26.35	8.208	.522	.317	.857
COM4	26.27	8.346	.481	.276	.861

REL1	26.19	7.624	.668	.473	.844
REL2	26.25	7.781	.629	.439	.847
REL3	26.41	8.029	.553	.356	.854
REL4	26.33	8.187	.508	.309	.858

For a 12-item instrument administered to very young learners, these coefficients are adequate and analytically meaningful. They suggest that the questionnaire achieved a reasonable balance between brevity, accessibility, and internal coherence, which justifies the reporting of both total scores and SDT-related subdimension scores. Table 3.3 presents the item-total statistics for the 12-item motivation questionnaire. Overall, the results indicate that all items performed acceptably and contributed meaningfully to the internal consistency of the scale. The corrected item-total correlations ranged from .481 to .668, which suggests that each item was positively and adequately associated with the overall construct being measured. In reliability analysis, values above .30 are generally considered acceptable, so none of the items appeared weak enough to warrant removal. The strongest item-total relationships were found for REL1 (.668), COM1 (.645), REL2 (.629), and AUT1 (.612), indicating that these items were especially well aligned with the overall scale. Even the lowest values, such as COM4 (.481) and AUT4 (.497), remained within an acceptable range, suggesting that these items still contributed usefully to the instrument. The “Cronbach’s alpha if Item Deleted” column further supports the retention of all 12 items. The alpha values ranged from .844 to .861, showing that deleting any individual item would not lead to a substantial improvement in reliability. In other words, no single item appears to reduce the consistency of the questionnaire in a problematic way. The squared multiple correlations, which ranged from .276 to .473, also indicate that the items shared a reasonable amount of variance with the remaining items in the scale. Taken together, these results suggest that the questionnaire functioned as a coherent and dependable measure of student motivation. For a short instrument

administered to very young learners, the item-level reliability evidence is satisfactory and supports the use of the full 12-item scale in subsequent analysis.

3.6.4 Assumption testing

Before inferential testing, the quantitative data were screened for normality and homogeneity of variance. The non-significant Shapiro-Wilk results across the four score distributions and the non-significant Levene's tests for pre-test and post-test comparisons indicated that the data did not show assumption violations severe enough to undermine parametric analysis (Field, 2018). This matters because the inferential claims of the study depend not only on statistical significance but also on the defensibility of the analytic conditions under which significance was tested.

3.6.5 Qualitative trustworthiness

Multiple complementary procedures supporting trustworthiness in the qualitative strand helped bolster the credibility, consistency, and interpretative balance of findings. Instead of only looking at an isolated or impressionistic snapshot, we used classroom observations across both classes for a repeated measure across lessons to better identify similarities over time, rather than extrapolating too much out of an outlier moment. With the variance in young learners' behaviours from lesson to lesson based on the level of difficulty and their energy and other interactive factors within the classroom, having repeated chances to observe both classes during the intervention allowed the qualitative analyses to be based on sustained observation of the context rather than being based on selective observations.

Second, trustworthiness was strengthened through triangulation across multiple qualitative sources, namely classroom observations, student interviews, and the teacher interview. Each source illuminated the phenomenon from a different angle. Observations made it possible to document visible participation patterns, turn-taking behavior, peer support, hesitation, and classroom atmosphere during the review stage. Student interviews provided access to how children interpreted those experiences in their own age-appropriate terms, while the teacher

interview added a broader comparative perspective across the two classes and across the full intervention period. The value of this triangulation lay not only in confirming recurring patterns, but also in clarifying how certain behaviors or impressions should be understood. When similar themes appeared across these different forms of evidence, confidence in the interpretation increased; when the sources diverged, that divergence itself became analytically meaningful rather than being treated as a problem to be hidden.

A further aspect of trustworthiness involved the coding process itself. The analysis was not restricted to evidence that supported a uniformly positive view of Baamboozle, but instead deliberately retained both confirming and disconfirming evidence. This meant that the coding framework remained open to tensions, exceptions, and reservations that complicated a simple success narrative. For example, although many data points indicated higher excitement, broader participation, and stronger peer support in the Baamboozle class, the analysis also preserved evidence concerning noise, turn imbalance, pace-related pressure, and the continuing value of calm teacher explanation in the comparison condition. Retaining such tensions was important because it prevented the qualitative strand from becoming overly celebratory and allowed the findings to reflect the bounded and conditional nature of the intervention more accurately.

In addition, interview summaries were checked against the audio recordings before final theme writing took place. This step helped reduce the risk that the summaries would drift away from what participants had actually said, especially given that the study relied on concise, manageable interview reporting rather than full transcript reproduction. Brief quoted excerpts later included in Chapter 4 were translated into English only after the Vietnamese wording had been checked carefully for fit with the original recordings. This translation procedure was necessary because the trustworthiness of reported quotations depends not only on literal wording, but also on whether the translated version preserves the intended meaning and tone of the original response. By checking the Vietnamese phrasing

first and translating only after that verification, the study aimed to maintain accuracy in the representation of participants' voices.

Overall, these procedures supported qualitative rigor by ensuring that the analysis remained grounded, balanced, and responsive to complexity. In mixed-methods research of this kind, trustworthiness does not mean forcing all evidence into a single harmonious pattern. Rather, it means demonstrating that the interpretation is capable of accommodating complexity, qualification, and even limited contradiction without losing coherence. In the present study, that standard was especially important because the purpose was not merely to show that Baamboozle produced positive reactions, but to explain how it functioned in practice, for whom it appeared beneficial, and under what classroom conditions its motivational advantages were strengthened or constrained.

3.7 Ethical Considerations

Because the study involved young children within regular school instruction, ethical protection extended beyond formal consent. Parental consent and child assent were obtained through age-appropriate procedures, and students were told explicitly that the questionnaire was not a test, that there were no right or wrong answers, and that non-participation would not affect their schooling. Interview excerpts were anonymized through participant codes, and the teacher was likewise de-identified as T1.

A further ethical priority was proportionality of intervention. Both classes received the same curricular content, were taught by the same teacher, and differed only in review format, which reduced the risk that one class would be educationally disadvantaged. Data collection was scheduled so as to minimize disruption to ordinary lessons, and the reporting of results avoided identifying individuals through recognizable personal detail. In a school-based study with children, these protections are not peripheral; they are part of the methodological integrity of the project.

3.8 Data Collection Procedures

The data collection procedure followed a fixed eight-week sequence designed to keep the two classes comparable. In Week 1, consent and assent were obtained, the study was explained in age-appropriate language, and both classes completed the pre-test motivation questionnaire under the same administration conditions.

Weeks 2–7 formed the six-week intervention period. Both classes studied the same Grade 2 English content with the same teacher, but the review stage differed: the Baamboozle class used team-based Baamboozle review, whereas the comparison class used teacher-centered review. Across the intervention, 24 review lessons were observed, comprising 12 in each class.

In Week 8, both classes completed the post-test questionnaire. Student interviews and the teacher interview were then conducted to clarify how participants experienced the review formats. Quantitative data were analyzed first, qualitative data were analyzed thematically, and both strands were integrated through a joint display.

Table 3.4 Eight-week data collection schedule

Week	Activity
1	Orientation, consent/assent, pre-test questionnaire, and classroom familiarization
2	Observed lessons 1–2 in both classes (opening observation point embedded in Lesson 1)
3	Observed lessons 3–4 in both classes
4	Observed lessons 5–6 in both classes
5	Observed lessons 7–8 in both classes
6	Observed lessons 9–10 in both classes
7	Observed lessons 11–12 in both classes (closing observation point embedded in Lesson 12)
8	Post-test questionnaire, student interviews, and teacher

	interview
--	-----------

This sequence ensured that pre-test measurement, intervention exposure, observations, post-test measurement, interviews, and mixed-methods integration were clearly separated while maintaining curricular equivalence across the two classes.

3.9 Intervention Protocol

Both groups studied the same Grade 2 English content throughout the intervention period, including topics such as colors, animals, numbers, classroom objects, family, and food. This point is methodologically important because it clarifies that the study did not compare two different curricular programs or two different sets of language targets. In both classes, the core instructional content remained aligned with the same grade-level English syllabus, and both groups covered the same general lesson themes during the research period. The major difference lay in the final 12 to 15 minutes of each lesson, which functioned as the focal review segment in both classes and served as the specific site of comparison for the intervention and the classroom observations. In this sense, the study isolated the review phase as the key area of pedagogical variation rather than treating the entire lesson as wholly different across the two conditions.

In the Baamboozle class, review was conducted through Baamboozle as a team-based, teacher-mediated activity. Students worked in teams, selected numbered tiles, viewed or listened to prompts prepared in advance, discussed possible answers briefly with peers, and then responded orally when called upon. The teacher remained central to the process by moderating turns, verifying the correctness of answers, allocating points, encouraging team support, and controlling the pace of the activity so that it remained manageable for Grade 2 learners. This arrangement is important to note because the intervention did not position students as independent users of technology. Since the class remained intact and the review was conducted in an ordinary primary-school setting, Baamboozle functioned as a whole-class instructional tool projected and managed

by the teacher rather than as an individualized student-device activity. The platform therefore reorganized participation within the review stage without altering the basic teacher-led structure of the classroom.

For the purpose of comparing the two instructional methods, review for the comparison class was conducted as per regular classroom methods, which are already known in the teaching of English at primary school level. Examples of these types of review are through the use of teacher questioning, choral repetition of words and sentences, short writing or using the blackboard to practice using words and sentences, and the use of worksheets. It was important for the logic within this study that the condition for comparison was not an artificial or weak comparison to the use of a Baamboozle-based method of review. Rather than looking to compare the use of the Baamboozle method against no review or ineffective forms of classroom instruction, the study was looking to compare Baamboozle-integrated English review with already established forms of review that are widely accepted as educationally appropriate and currently utilised in primary classroom environments. This is particularly important, since any differences found between both methods are placed in a more meaningful context of evaluation as both instructional methods have been found to be effective alternative pedagogies, rather than in the context of a false or contrived comparison to an alternative pedagogical method.

The lesson plans presented in Appendix F make clear that the intervention did not introduce different language content or a different sequence of topics. What changed was the organization of the review phase and, more specifically, the way students were invited to participate in that phase. Across the six intervention weeks, 12 review lessons in each class were observed and recorded using the same checklist. Applying the same observation schedule to both classes supported comparability across conditions and allowed the study to document not only whether participation differed, but also how those differences appeared in practice over time. This repeated observation across equivalent review segments

strengthened the internal coherence of the intervention design by keeping the content stable while varying only the review format under investigation.

3.10 Data Analysis

3.10.1 Quantitative analysis

The quantitative data were analyzed in a staged sequence designed to move from measurement defensibility to outcome interpretation:

1. reliability and assumption testing
2. descriptive statistics for pre-test and post-test scores
3. independent-samples t-test for baseline equivalence
4. paired-samples t-tests for within-group change
5. independent-samples t-test for post-test difference
6. effect-size calculation using Cohen's d and d_z
7. post-test comparison across the three SDT subdimensions
8. supplementary gain-pattern analysis using the individual raw-score tables

This sequence was methodologically important because it prevented the interpretation from relying on a single inferential result. Baseline equivalence addressed comparability, within-group tests documented change over time, post-test comparison estimated intervention difference, effect sizes clarified magnitude, and gain-pattern analysis ensured that individual variation and contradictory cases remained visible rather than being absorbed into group means.

3.10.2 Qualitative analysis

The qualitative data were analyzed through thematic analysis following the recursive logic described by Braun and Clarke (2006). Observation notes, structured interview summaries, and teacher commentary were read repeatedly, coded for recurrent motivational processes, clustered into candidate categories, reviewed against the whole dataset, and then refined into higher-order themes. Student and teacher interviews were audio-recorded with permission. Because the study involved very young children and was designed as a school-based case study rather than a discourse-transcription project, the recordings were converted into

structured summary sheets in Vietnamese that retained key content, notable wording, and comparison points relevant to the research questions. Those summaries were then translated into English for reporting, with representative quotations checked back against the recordings before inclusion in Chapter 4. Particular analytic attention was given to anticipatory attention, answer confidence, peer support, fairness, pacing, noise, and technological interruption because these were the classroom processes most relevant to the study's SDT-informed interpretation.

3.10.3 Mixed-methods integration

Mixed-methods integration occurred only after the quantitative and qualitative strands had first been analyzed separately according to their own procedures and evidential standards. This sequencing was important because it allowed each strand to generate its own findings before any attempt was made to connect them. The quantitative strand identified whether there were measurable differences between the two classes in overall motivation and across the SDT-related subdimensions, while the qualitative strand examined how the review phase was actually experienced and enacted in classroom practice. Integration therefore did not begin from the assumption that the two strands must necessarily confirm each other in a simple or uniform way. Instead, it proceeded through comparison, explanation, and tension mapping in order to determine how the findings from each strand related to one another and what kind of overall interpretation they could support.

The main instrument used in this integrated study was the joint display. By using a joint display, data were laid out in a predetermined format. This allowed for statistical patterns to be placed alongside observational and interview data and allowed direct examination of cross-strand relationships. The joint display provided insight into whether the strands converge at the all the same underlying mechanisms; whether the strands provided complementary evidence by describing different aspects of an overall pattern; or whether the strands provided

qualifications that complicated otherwise straightforward interpretations. For instance, a higher post-test score from the Baamboozle class becomes more analytically meaningful when viewed together with the observational data which demonstrated greater anticipatory attention, more wide entry to turns, more visible peer support, plus the interviews confirmed that students had less anxiety and were more willing to participate. However, the joint display also accommodated tension; in that positive motivation patterns and concerns about noise and fairness coexisted.

This integrative step was essential because the strongest claims of the study do not rest on numerical difference alone, nor on qualitative impression alone, but on the correspondence between measurable change and classroom process. Quantitative results could show that motivation differed between groups and that gains were stronger in one condition, but they could not by themselves explain how those differences emerged in practice or what kind of classroom experience lay behind them. Conversely, qualitative evidence could illuminate how students participated, how the teacher managed the activity, and what kinds of support or tension were present, but on its own it could not establish the same degree of patterned contrast across groups. Integration was therefore necessary to build an interpretation that was both empirically grounded and explanatorily coherent. In the present study, that coherence depended on showing that the direction of statistical change was meaningfully related to observable classroom processes and participant perceptions, while still preserving the qualifications that prevented the final account from becoming overly simplified.

3.11 Threats to Validity and Their Management

Because the study used two intact classes within one bounded school case rather than random assignment, several threats to validity required explicit management.

Selection threat was reduced, though not eliminated, through pre-testing and the near-null baseline difference between groups. This does not create

randomization after the fact, but it does strengthen the argument that later divergence is unlikely to be explained simply by unequal starting levels.

Teacher effect was minimized by using the same teacher across both classes. This was important because enthusiasm, management style, and feedback patterns can themselves influence motivation; holding the teacher constant reduced one major source of between-group variation.

History and maturation were partly controlled by running both classes over the same instructional period and with the same syllabus content. Any broad changes associated with school routine, growing familiarity with classroom expectations, or passage of time were therefore more likely to affect both groups than only one.

Instrumentation threat was managed by administering the same motivation questionnaire at pre-test and post-test under the same age-appropriate conditions. Oral support and consistent response format reduced the likelihood that score differences reflected changes in task comprehension rather than changes in classroom motivation.

Novelty threat remained a plausible limitation, particularly for the digital-game condition. The thesis therefore interprets the findings as evidence of stronger short-term motivational support within the case, not as proof of permanent motivational transformation.

Implementation variability was reduced through recurring lesson structure, shared syllabus content, and lesson planning that kept the manipulated difference focused on review format. Even so, classroom interaction remained dynamic, so validity is best understood as strengthened rather than guaranteed.

3.12 Chapter Summary

This chapter has shown that the study used a mixed-methods case study design suited to the research questions and the school context. The case study frame remained central even as the study embedded a pre-test/post-test comparison between two intact classes. The design aligned closely with the problem statement,

the theoretical framework, and the operationalization of motivation. The next chapter reports the findings neutrally and systematically, ensuring that all provided data, including contradictory and weak patterns, are represented.

CHAPTER 4: FINDINGS

4.1 Quantitative Findings

4.1.1 Reliability and assumption testing

Prior to conducting the main inferential analyses, the questionnaire's internal consistency and the relevant statistical assumptions were assessed to ensure that the quantitative strand rested on an adequate analytic foundation. Table 4.1 indicates that the instrument achieved satisfactory reliability at both administration points, with Cronbach's alpha coefficients of .81 for the pre-test and .86 for the post-test. These coefficients suggest that the questionnaire items functioned with acceptable coherence in the present dataset. Given the brevity of the instrument and the young age of the participants, such values provide reasonable support for the reliability of the scale as a classroom measure of motivation. The item-level reliability evidence reported in Chapter 3 further supports the retention of all 12 items and the use of the fixed SDT item groupings for subdimension analysis.

Table 4.1 Internal consistency of the motivation questionnaire

Measure	N	Items	Cronbach's α
Pre-test	60	12	0.81
Post-test	60	12	0.86

The assumption checks reported in Tables 4.2 and 4.3 support the use of parametric analysis and also provide a clearer picture of the statistical shape of the dataset before the main comparisons were conducted. Because each score distribution involved 30 students, the Shapiro-Wilk test is the more informative normality indicator in the present analysis. As shown in Table 4.2, none of the four Shapiro-Wilk tests reached statistical significance. The Baamboozle group produced Shapiro-Wilk values of .961 at pre-test ($p = .338$) and .954 at post-test ($p = .212$), while the comparison group produced values of .975 at pre-test ($p = .686$) and .947 at post-test ($p = .140$). In all four cases, the p values remained above the .05 threshold, which indicates that the observed score distributions did

not differ significantly from a normal distribution. On that basis, there was no statistical indication that either group's pre-test or post-test scores showed a problematic non-normal pattern likely to distort the subsequent t-test results.

Table 4.2 Tests of Normality

Group	Kolmogorov–Smirnov Statistic	df	Sig.	Shapiro- Wilk Statistic	df	Sig.
Baamboozle (Pre-test)	.174	30	.200*	.961	30	.338
Baamboozle (Post-test)	.168	30	.200*	.954	30	.212
Comparison (Pre-test)	.149	30	.200*	.975	30	.686
Comparison (Post-test)	.158	30	.200*	.947	30	.140

The Shapiro-Wilk values also show that the four distributions were reasonably similar in form. The comparison-class pre-test distribution was the closest to normal in statistical terms, with the highest Shapiro-Wilk value (.975) and the largest p value (.686). By contrast, the comparison-class post-test distribution produced the lowest Shapiro-Wilk value (.947) and the smallest p value (.140), which suggests that this distribution was somewhat less regular than the others. However, this result still remained clearly non-significant and therefore did not indicate a serious deviation from normality. The two Baamboozle-class distributions fell between these two points, with both values comfortably within an acceptable range. In practical terms, the normality results suggest that the four distributions were not perfectly identical in shape, but none showed a departure strong enough to undermine the later parametric comparisons. The Kolmogorov–Smirnov results point in the same direction. All four significance values were reported as .200*, which in SPSS indicates that the obtained p values were at or above the program's upper reporting bound rather than statistically significant. The corresponding statistics ranged from .149 to .174 across the four distributions. Although the Shapiro-Wilk test is more relevant here, the

Kolmogorov–Smirnov results still provide supporting evidence that the score distributions were acceptably regular. The convergence of the two normality tests strengthens confidence that the quantitative data were sufficiently well behaved for inferential analysis.

Table 4.3 Tests of Homogeneity of Variances

Measure	Method	Levene Statistic	df1	df2	Sig.
Pre-test	Based on Mean	.250	1	58	.619
Post-test	Based on Mean	.660	1	58	.421

Table 4.3 provides the second assumption check by testing homogeneity of variance for the between-group comparisons. Here again, the results were non-significant at both testing points. For the pre-test comparison, Levene’s statistic was .250 with $p = .619$, and for the post-test comparison, Levene’s statistic was .660 with $p = .421$. These results indicate that the variability of scores in the Baamboozle and comparison groups did not differ significantly at either stage of measurement. This is important because the later independent-samples t-tests compare group means under the assumption that the spread of scores is broadly comparable across groups. In the present dataset, there is no evidence from Table 4.3 that unequal variance posed a serious problem for either the baseline comparison or the post-test comparison.

Taken together, Tables 4.2 and 4.3 show that the main assumptions for the study’s parametric procedures were adequately satisfied. The four score distributions were sufficiently close to normal, and the Baamboozle and comparison groups showed sufficiently similar variances, to justify the use of paired-samples and independent-samples t-tests in the analyses that follow. These checks do not imply that the data were perfectly ideal in every respect, but they do show that there was no evident assumption violation strong enough to call the later quantitative findings into question. Within the logic of this chapter, the assumption-testing stage therefore functions as more than a routine statistical

formality: it establishes that the comparisons reported in the next subsections rest on data that are suitably stable for inferential interpretation.

4.1.2 Descriptive statistics

The descriptive statistics showed that the two groups began from almost identical overall motivation levels. The Baamboozle group recorded a pre-test mean of 22.30 (SD = 3.04), while the comparison group recorded 22.33 (SD = 3.15). The difference at that point was only 0.03 points, which is descriptively negligible and indicates that the later outcome pattern did not arise from an obviously higher starting point in the intervention class.

At post-test, both groups recorded higher means than at pre-test, but the size of increase differed noticeably. The Baamboozle group rose to 28.67 (SD = 2.95), whereas the comparison group rose to 25.27 (SD = 2.83). The mean gain was therefore 6.37 points in the Baamboozle group and 2.93 points in the comparison group. The Baamboozle group's increase was thus more than twice as large as that of the comparison group.

The raw score tables in Appendix G showed the same general pattern at the individual level. In the Baamboozle group, pre-test scores ranged from 17 to 28 and post-test scores ranged from 21 to 33. In the comparison group, the range moved from 16-29 at pre-test to 19-33 at post-test. Thirteen students in the Baamboozle group reached a post-test score of 30 or above, compared with only two students in the comparison group. The descriptive picture therefore involved not only a higher mean, but also a denser concentration of high post-test scores in the Baamboozle class.

Table 4.4 Descriptive statistics summary

Group	n	Pre-test M	Pre-test SD	Post- test M	Post- test SD	Mean gain
Baamboozle	30	22.30	3.04	28.67	2.95	6.37
Comparison	30	22.33	3.15	25.27	2.83	2.93

As Table 4.4 indicates, both groups moved upward over time, which is consistent with the fact that both classes received regular English instruction over the same period. However, the upward movement was clearly larger in the Baamboozle group. When the two groups' mean gains are compared directly, the Baamboozle group improved by a substantially larger margin than the comparison group. Figure 4.1 shows this.

Figure 4.1. Motivation scores of the Baamboozle and comparison groups

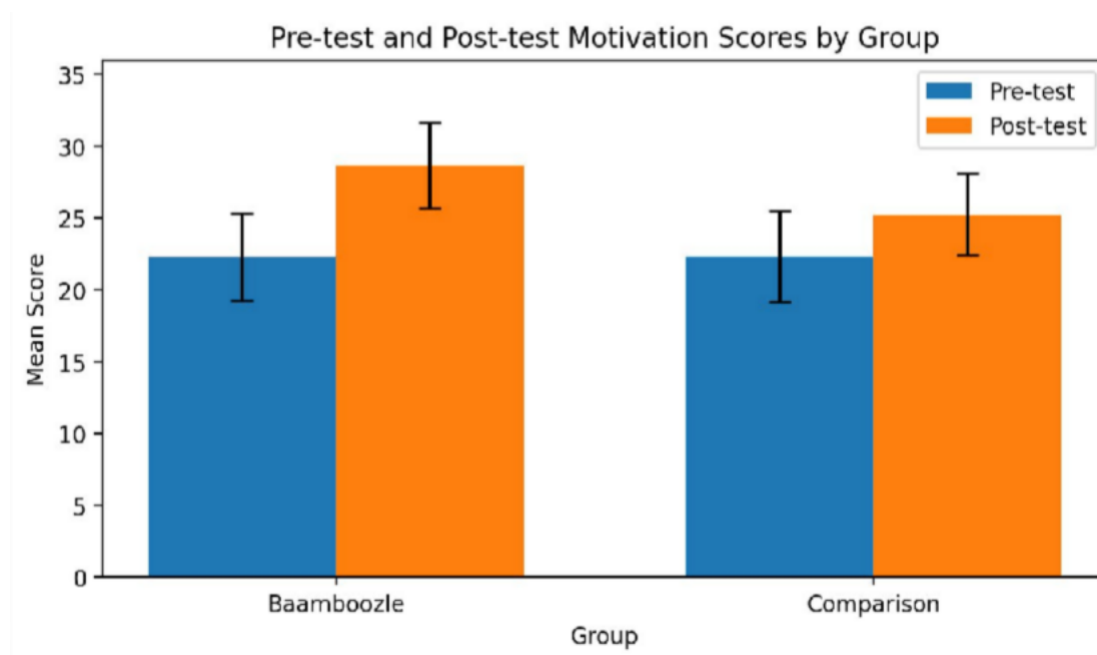


Figure 4.1 demonstrates that the two groups began the study at almost the same motivational level, with very similar pre-test mean scores (Baamboozle: $M = 22.30$, $SD = 3.04$; Comparison: $M = 22.33$, $SD = 3.15$). This indicates that the two classes were broadly comparable before the intervention. After the intervention period, both groups showed improvement, but the increase was clearly greater in the Baamboozle group. The Baamboozle group reached a post-test mean of 28.67 ($SD = 2.95$), whereas the comparison group reached 25.27 ($SD = 2.83$). In terms of mean gain, the Baamboozle group improved by 6.37 points, compared with 2.93 points in the comparison group. Visually, the figure therefore suggests that although ordinary teacher-centered review was associated with some motivational improvement, Baamboozle-integrated review was associated with a

substantially stronger increase in students' classroom motivation. The relatively similar standard deviations across pre-test and post-test scores also suggest that the difference between groups was not simply the result of unusually scattered responses in one class, but reflected a broader pattern of improvement in the Baamboozle group.

4.1.3 Baseline equivalence and post-test comparison

The independent-samples t-test in Table 4.5 for the pre-test confirmed the descriptive picture of baseline equivalence already suggested by the group means. As reported in Table 4.5, the pre-test mean difference between the two classes was only -0.03 points, with the comparison group scoring marginally higher than the Baamboozle group. This difference was statistically negligible, $t(58) = -0.04$, $p = .967$, and the effect size was virtually zero, $d = -0.01$. The 95% confidence interval, which ranged from -1.63 to 1.57, also crossed zero comfortably and remained fairly centered around it. Taken together, these results indicate that the two intact classes began the study from essentially the same position on the overall motivation measure. In other words, there was no statistical evidence at pre-test of a meaningful between-group difference that would suggest one class entered the intervention with a motivational advantage over the other.

Table 4.5 Baseline equivalence and post-test comparison

Comparison	Mean difference	t(df)	p	Cohen's d	95% CI
Pre-test: Baamboozle vs comparison	-0.03	-0.04 (58)	.967	-0.01	[-1.63, 1.57]
Post-test: Baamboozle vs comparison	3.40	4.56 (58)	< .001	1.18	[1.91, 4.89]

Table 4.5 shows equivalence before the intervention and clear separation after it. This baseline result is important for the logic of the quantitative strand

because the study did not use random assignment. Since the two classes were retained in their natural school grouping, the pre-test comparison serves as an important check on initial comparability. The near-zero mean difference, null significance value, trivial effect size, and confidence interval centered around zero all point in the same direction. Although baseline equivalence cannot eliminate all possible pre-existing differences between intact classes, it does strengthen the argument that the later post-test separation is unlikely to be explained simply by one group starting substantially higher than the other. Within the present dataset, the pre-test analysis therefore supports the view that the two groups were sufficiently comparable for the subsequent contrast to be interpreted meaningfully.

The post-test comparison produced a clearly different pattern. By the end of the intervention period, the Baamboozle group had a higher mean than the comparison group by 3.40 points. This difference was statistically significant, $t(58) = 4.56$, $p < .001$, and the corresponding effect size was large, $d = 1.18$. The 95% confidence interval ranged from 1.91 to 4.89, which is important for two reasons. First, the interval remained entirely above zero, meaning that the direction of difference consistently favored the Baamboozle group. Second, even the lower bound of the interval still represents a meaningful positive difference in favor of the Baamboozle class. The post-test result therefore indicates not merely a slight numerical separation, but a clear and substantial between-group contrast at the end of the six-week intervention period.

The contrast between the pre-test and post-test results is especially important when the two comparisons are read together rather than in isolation. At pre-test, the classes were statistically indistinguishable on the overall motivation measure; at post-test, they were clearly separated. This shift from near-equivalence to statistically significant difference forms one of the central quantitative patterns in the chapter. It does not by itself explain how that separation emerged, and it does not mean that only the Baamboozle class improved. However, it does show that the Baamboozle class finished the intervention period from a clearly stronger

position on the study's motivation measure than the comparison class. In this sense, the post-test comparison complements the descriptive and within-group analyses by showing that the motivational advantage of the Baamboozle group was visible not only in amount of change over time, but also in final outcome level.

Table 4.5 therefore captures two distinct phases in the quantitative pattern. The first phase is one of baseline comparability, in which the two groups appear effectively equivalent at the outset. The second phase is one of end-point separation, in which the Baamboozle group records a clearly higher post-test score than the comparison group. This transition from equivalence to difference is a key part of the quantitative answer to RQ1, because it shows that the intervention period was associated with a stronger motivational outcome in the Baamboozle class than in the class that received traditional review.

4.1.4 Within-group change

The paired-samples analyses in Table 4.6 showed that both groups improved significantly between pre-test and post-test, but the size of that improvement was clearly different across the two classes. In the Baamboozle group, the mean score increased by 6.37 points from pre-test to post-test. This change was statistically significant, $t(29) = 8.18$, $p < .001$, and the corresponding within-group effect size was very large, $d_z = 1.49$. The 95% confidence interval for the mean gain ranged from 4.77 to 7.96, which indicates that the improvement in the Baamboozle class was not only statistically reliable but also substantial in magnitude. Because the entire confidence interval remained well above zero, the direction of change was consistently positive across the group-level estimate. In descriptive terms, the Baamboozle class did not simply fluctuate upward slightly; it showed a strong rise in motivation scores over the intervention period.

Table 4.6 Within-group change from pre-test to post-test

Group	Mean gain	t(df)	p	Cohen's dz	95% CI of gain
Baamboozle	6.37	8.18 (29)	< .001	1.49	[4.77, 7.96]

Comparison	2.93	4.12 (29)	< .001	0.75	[1.48, 4.39]
------------	------	-----------	--------	------	--------------

The comparison group also showed significant improvement over time. Its mean gain was 2.93 points, with $t(29) = 4.12$, $p < .001$, and a medium-to-large within-group effect size of $d_z = 0.75$. The 95% confidence interval for this gain ranged from 1.48 to 4.39, which again remained fully above zero and therefore indicates a reliable positive change from pre-test to post-test. This result is important because it shows that motivation did not remain static in the comparison class. The comparison condition also produced upward movement during the study period, which suggests that traditional teacher-centered review remained an active instructional condition rather than a neutral or inactive baseline. The quantitative pattern therefore does not indicate improvement in only one class; instead, it shows improvement in both classes, but to different degrees.

When the two within-group results are read side by side, the difference in magnitude becomes clearer. The Baamboozle group's mean gain of 6.37 was more than twice the comparison group's mean gain of 2.93. The within-group effect size in the Baamboozle class was likewise much larger than that of the comparison class, with $d_z = 1.49$ compared with $d_z = 0.75$. The confidence intervals also reflect this difference in scale. Although both intervals indicate positive change, the interval for the Baamboozle group is positioned notably higher than that for the comparison group, which is consistent with a stronger upward shift in the Baamboozle class. These contrasts suggest that the motivational growth observed in the Baamboozle group was not simply a slightly stronger version of the comparison-class pattern, but a markedly larger change across the same time period.

This subsection is also important because it adds something different from the post-test comparison. The post-test analysis shows that the Baamboozle class finished from a higher position than the comparison class. The within-group analysis, by contrast, shows how much each class changed across time within its own pre-test to post-test trajectory. Read together, these results strengthen the

quantitative pattern in two ways. First, they show that the Baamboozle group finished higher at the end of the intervention period. Second, they show that this end-point advantage was accompanied by a stronger upward movement across time, not merely by a final score difference viewed in isolation. In that sense, the within-group results complement the post-test comparison by showing that the intervention class did not only end higher, but also changed more strongly between the beginning and end of the study period.

Table 4.6 therefore indicates that both forms of review were associated with increased motivation over time, but that the increase was substantially greater in the Baamboozle class. The quantitative pattern at this stage is therefore one of shared improvement but unequal magnitude. Both groups moved upward, yet the Baamboozle group showed the stronger within-group change on the study's motivation measure. This finding becomes especially important when read alongside the earlier evidence of baseline equivalence, because it suggests that the later difference between groups was associated not only with end-point separation but also with a distinctly larger rise in scores within the Baamboozle class itself..

4.1.5 Supplementary gain-score comparison

A supplementary independent-samples t-test was conducted on gain scores calculated from the pre-test and post-test results for each student. Unlike the earlier post-test comparison, which examined only the final difference between groups at the end of the intervention period, this analysis compared the amount of score increase directly across the two classes. It therefore provided an additional check on whether the Baamboozle advantage was visible not only in end-point means, but also in the size of motivational growth across time. As shown in Table 4.7, the mean gain difference between groups was 3.43 points in favor of the Baamboozle class. This difference was statistically significant, $t(58) = 3.25$, $p = .002$, with a Cohen's d of 0.84 and a 95% confidence interval ranging from 1.32 to 5.55.

Table 4.7 Gain-score comparison between groups

Comparison	Mean gain	t(df)	p	Cohen's d	95% CI
------------	-----------	-------	---	-----------	--------

	difference				
Baamboozle gain vs comparison gain	3.43	3.25 (58)	.002	0.84	[1.32, 5.55]

Several aspects of this result are important. First, the p value of .002 indicates that the difference in gain scores between the two groups is unlikely to be attributable to chance variation alone. Second, the effect size of 0.84 falls in the large range, which suggests that the difference in motivational growth between the two classes was not trivial in practical terms. Third, the confidence interval remained entirely above zero, which means that the direction of the gain-score difference consistently favored the Baamboozle group. Even at its lower bound, the interval still indicates a positive advantage for the Baamboozle class. Read together, these statistics show that the Baamboozle group did not merely finish with a higher post-test mean; it also showed a clearly stronger increase in motivation across the intervention period.

This supplementary analysis is useful because it strengthens the quantitative pattern from a different angle. The earlier subsections already showed baseline equivalence, significant within-group improvement in both classes, and a clear post-test difference favoring the Baamboozle group. The gain-score comparison adds to this pattern by testing the difference in improvement directly. In other words, it asks not only whether the Baamboozle class ended higher, but whether it changed more across time than the comparison class. The answer from Table 4.5 is yes. The Baamboozle class displayed a larger upward shift than the comparison class, and this pattern remained visible even when the analysis focused specifically on the size of score change rather than on pre-test or post-test means alone.

This subsection is also important because it reduces the risk that the quantitative interpretation rests too heavily on any single statistical comparison. In the present chapter, the evidence is cumulative rather than isolated. Baseline equivalence showed that the two classes started from comparable positions. The within-group tests showed that both groups improved, but that the Baamboozle

class improved more strongly. The post-test comparison showed that the Baamboozle class finished from a higher position. The gain-score comparison now confirms the same pattern by showing that the amount of growth itself was significantly larger in the Baamboozle group. The consistency of these results across several analytic steps strengthens the internal coherence of the quantitative strand.

Table 4.7 therefore reinforces the same quantitative narrative established in the earlier subsections: during the intervention period, the Baamboozle class showed a stronger upward movement on the study's motivation measure than the comparison class. The value of this analysis lies not in introducing a different conclusion, but in confirming the same conclusion through a direct comparison of score growth. In that sense, the gain-score test serves as a supplementary but important piece of evidence that the motivational advantage of the Baamboozle group was visible across time as well as at the final testing point.

4.1.6 SDT subdimension comparisons

Post-test scores were also compared across the three SDT-based subdimensions of autonomy, competence, and relatedness in order to show whether the quantitative advantage of the Baamboozle group appeared across the full motivation measure or only in one part of it. As shown in Table 4.8, the Baamboozle group recorded higher post-test means than the comparison group on all three subdimensions.

Table 4.8 Post-test SDT subdimension comparisons

Subdimension	Baamboozle M (SD)	Comparison M (SD)	t(58)	p	Cohen's d
Autonomy	9.57 (1.38)	8.33 (0.99)	3.97	< .001	1.02
Competence	9.50 (0.97)	8.50 (1.07)	3.78	< .001	0.98
Relatedness	9.77 (1.01)	8.57 (1.48)	3.68	< .001	0.95

For autonomy, the post-test mean was 9.57 (SD = 1.38) in the Baamboozle group and 8.33 (SD = 0.99) in the comparison group, $t(58) = 3.97$, $p < .001$, $d = 1.02$. For competence, the corresponding means were 9.50 (SD = 0.97) and 8.50 (SD = 1.07), $t(58) = 3.78$, $p < .001$, $d = 0.98$. For relatedness, the means were 9.77

(SD = 1.01) in the Baamboozle group and 8.57 (SD = 1.48) in the comparison group, $t(58) = 3.68$, $p < .001$, $d = 0.95$. All three comparisons were statistically significant, and all three effect sizes were large. The quantitative advantage of the Baamboozle group was therefore visible across autonomy, competence, and relatedness rather than being restricted to only one subdimension.

The three Baamboozle-class means were also relatively close to one another. Relatedness was the highest post-test subdimension mean, followed by autonomy and competence, but the gaps among them were small. This suggests a fairly even motivational profile rather than a sharply uneven one. In descriptive terms, the Baamboozle group did not appear to improve in only one isolated aspect of motivation. Instead, the post-test pattern indicates stronger scores across the three dimensions used to organize the study's motivation measure Table 4.8 is especially important for Objective 3 because it shows that post-test differences were visible across autonomy, competence, and relatedness rather than on only one part of the scale. Although the questionnaire was analyzed primarily at total-score and subdimension level, the wording of the items clarifies what these differences meant in classroom terms. Higher autonomy scores suggest that students in the Baamboozle class were more likely to want to join the English activity, value having their turn, feel involved, and want to answer when the teacher addressed the class. Higher competence scores suggest greater willingness to try, greater satisfaction in correct answers, and stronger perceived ability to do English class work. Higher relatedness scores suggest stronger feelings of peer help, shared success, and classroom belonging

4.1.7 Individual gain-pattern summary from raw data

Table 4.9 provides a distributional summary of how students' motivation scores changed between pre-test and post-test in the two classes. In the Baamboozle group, 26 of 30 students (86.7%) recorded a positive gain, 3 students (10.0%) showed no change, and only 1 student (3.3%) declined. In the comparison group, 21 of 30 students (70.0%) improved, 4 students (13.3%) remained

unchanged, and 5 students (16.7%) declined. The median gain was also notably different across the two groups, with a median increase of 6 points in the Baamboozle class compared with 2 points in the comparison class. These figures indicate that positive change was more common and more substantial in the Baamboozle class than in the class that received traditional review..

Table 4.9 Individual gain-pattern summary from raw data

Group	Improved	Unchanged	Declined	Median gain	Minimum gain	Maximum gain
Baamboozle	26 (86.7%)	3 (10.0%)	1 (3.3%)	6	-3	+16
Comparison	21 (70.0%)	4 (13.3%)	5 (16.7%)	2	-5	+10

The distribution of larger gains shows the same pattern even more clearly. In the Baamboozle group, 20 students gained at least 5 points, 14 gained at least 8 points, and 6 gained at least 10 points. In the comparison group, the corresponding counts were 12, 4, and 2. This means that the Baamboozle class did not only contain more students who improved; it also contained a denser concentration of medium and large positive gains. In other words, the difference between groups was visible not only at the level of the overall mean, but also in the way score improvements were distributed across the class. The Baamboozle group showed a broader spread of strong positive movement, whereas the comparison group displayed a more mixed pattern of modest improvement, stability, and decline.

The minimum and maximum gains also help clarify the shape of the two distributions. In the Baamboozle group, gains ranged from -3 to +16, while in the comparison group they ranged from -5 to +10. These ranges show that neither class changed in a perfectly uniform way. Both groups contained students whose motivation scores rose sharply and students whose scores changed little or moved downward. However, the two ranges also suggest that the Baamboozle group contained a stronger upper-end improvement pattern, while the comparison group contained a heavier cluster of weaker or negative outcomes. This distributional evidence therefore adds an important layer to the earlier group-level statistics by showing that the Baamboozle advantage was not created only by a small number

of extreme cases. Rather, it was reflected in a wider concentration of positive change across the intervention class.

This subsection is important because it complements the mean-based analyses reported in the earlier tables. The pre-test and post-test comparisons, the within-group tests, and the gain-score analysis all showed that the Baamboozle group performed more strongly overall. Table 4.9 adds a more detailed view by showing how that pattern was distributed among individual students. The results indicate that the Baamboozle class not only achieved a higher average gain, but also showed a broader pattern of improvement across the class as a whole. At the same time, the table preserves an important qualification: the quantitative pattern was clear, but not perfectly uniform. A small number of students in the Baamboozle group did not improve, and a number of comparison-class students still made meaningful gains. The student-level score patterns therefore support the main quantitative conclusion while also showing that the intervention effect was stronger at the group level than as a mechanically identical outcome for every individual learner.

4.1.8 Non-uniform and contradictory quantitative cases

Although the overall quantitative pattern favored the Baamboozle group, the student-level score patterns were not perfectly uniform. As shown in Table 4.7, most students in the Baamboozle group improved, but a small number remained unchanged and one student declined. The comparison group showed a more mixed pattern, with a larger cluster of unchanged and declining cases, even though several students in that group also made substantial gains. This indicates that the positive quantitative trend was widespread but not identical across all learners.

The individual score patterns also suggest that higher starting scores did not automatically lead to stronger improvement. Some students who began with relatively high pre-test scores showed little subsequent movement, while some students with lower starting scores recorded marked gains over the intervention period. This pattern is important descriptively because it shows that motivational

change was distributed unevenly within both classes rather than following a simple ceiling-free progression.

These non-uniform cases do not overturn the main quantitative result. Rather, they qualify it by showing that the Baamboozle advantage was a clear group-level pattern with some individual variation, while the comparison group displayed a more mixed profile of improvement, stability, and decline. The quantitative findings should therefore be read as showing a stronger overall motivational pattern in the Baamboozle group, not as implying that every student responded in the same way.

4.1.9 Quantitative answer to RQ1

Taken together, the quantitative findings yielded a clear and internally consistent answer to RQ1. First, the two groups were effectively equivalent at baseline. The pre-test means were almost identical, and the independent-samples comparison was statistically null, which indicates that the two intact classes entered the study from essentially the same starting point on the motivation measure. Second, both groups improved from pre-test to post-test. This shows that motivation rose over time under both instructional conditions and that the comparison group also experienced positive movement during the intervention period rather than remaining static. Third, the Baamboozle group displayed the stronger quantitative pattern on every major group-level comparison. It recorded the larger mean gain, the larger within-group effect size, the higher post-test mean, and the clearer end-point advantage over the comparison group. The post-test difference was not only statistically significant but also substantial in magnitude, while the supplementary gain-score comparison reinforced the same pattern by showing that the Baamboozle group improved more directly across time rather than merely finishing with a higher endpoint score.

The quantitative findings were also coherent at a more detailed level. The advantage of the Baamboozle group was not confined to one narrow part of the construct, but was visible across autonomy, competence, and relatedness in the

SDT-based subdimension comparisons. This means that the stronger pattern in the Baamboozle class appeared across the study's full motivational framework rather than emerging from only one isolated subscale. In addition, the raw score tables and gain-pattern summaries showed that the positive result was broadly distributed across the Baamboozle group. Twenty-six of thirty students in the Baamboozle class improved, compared with twenty-one of thirty in the comparison class, and medium as well as large gains were more densely clustered in the Baamboozle group. At the same time, the raw data were not perfectly uniform. A small number of students in the Baamboozle group showed no change or decline, and several students in the comparison group still improved strongly. These cases are important because they show that the quantitative pattern was clear without being mechanically identical for every learner.

The quantitative answer to RQ1 can therefore be stated descriptively as follows: during the six-week intervention period, the Baamboozle group showed stronger short-term motivational growth than the traditional review group on the study's motivation measure. That advantage was visible in baseline-to-post-test change, in end-point comparison, in gain-score analysis, in the three SDT subdimensions, and in the wider distribution of individual improvements across the Baamboozle class. The quantitative strand, therefore, supports the conclusion that the Baamboozle-integrated English review format was associated with a stronger positive motivational pattern than the traditional review format within the study period and sample..

4.2 Qualitative Findings

4.2.1 Overview of the qualitative corpus

The qualitative corpus comprised 24 observed review lessons, 14 student interview reports, and one teacher interview report. The 24 observed lessons were evenly distributed across the two classes, with 12 observed review lessons in the Baamboozle group and 12 in the comparison group. The opening and closing observation points mentioned in Chapter 3 were embedded in the first and final

observed lessons rather than counted as additional lessons outside this corpus. Of the 14 student interviewees, 8 came from the Baamboozle group and 6 from the comparison group. The teacher who taught both classes was interviewed after the intervention and was coded T1.

Four themes were identified in the corpus: (1) joyful engagement and sustained attention, (2) perceived competence and confidence through immediate feedback, (3) relatedness through team participation and peer support, and (4) motivation with conditions, including noise, pacing, fairness, and technological constraints. The themes emerged from repeated patterns across observation notes, student reports, and the teacher interview, with sample coding examples shown in Appendix H.

4.2.2 Observation summary

The classroom observation summaries showed consistently stronger visible participation in the Baamboozle group during the review phase. The average number of volunteer or hand-raise episodes per lesson was 11.8 in the Baamboozle group, compared with 5.1 in the comparison group. Average peer-help utterances were 9.4 in the Baamboozle group and 3.0 in the comparison group. Average visible positive-affect episodes were 14.1 and 6.2, respectively. The estimated on-task participation rate was 87.6% in the Baamboozle group, compared with 74.2% in the comparison group. The Baamboozle group also recorded fewer average off-task incidents per lesson (2.9) than the comparison group (5.4). These lesson-level averages were calculated from the checklist records for 12 observed review lessons in each class.

In relative terms, the Baamboozle lessons produced more than twice as many volunteer episodes and positive-affect episodes as the comparison lessons, and more than three times as many peer-help utterances. At the same time, the average number of off-task incidents in the Baamboozle group was almost half that of the comparison group. The observation summaries therefore did not show

only louder lessons; they showed a review phase that appeared more collectively participatory and more behaviorally focused.

These observational indicators are important because they document classroom behavior across repeated lessons rather than relying only on interview recall. They show that the differences later described by students and the teacher were also visible in the lesson process itself. Refusal or hesitation episodes were tracked during observation, but they were used qualitatively rather than reported as a separate aggregate line in the summary because short hesitations and ordinary wait-time were not always distinguishable with stable reliability in this age group.

Table 4.10 Observation summary across the 24 lessons

Indicator	Baamboozle group	Comparison group
Average volunteer / hand-raise episodes per lesson	11.8	5.1
Average peer-help utterances per lesson	9.4	3.0
Average visible positive-affect episodes per lesson	14.1	6.2
Estimated on-task participation rate	87.6%	74.2%
Average off-task incidents per lesson	2.9	5.4

Across all listed indicators, Table 4.10 shows stronger visible engagement in the Baamboozle condition, although the later qualitative theme on constraints also makes clear that this stronger engagement came with management demands.

4.2.3 Theme 1: Joyful engagement and sustained attention

The first theme concerned the way Baamboozle-integrated English review held students' attention and produced a more visibly energetic classroom atmosphere. Baamboozle-class participants repeatedly described the review stage as enjoyable and difficult to ignore. E1 stated, "I look at the screen all the time because I want to pick a number." E3 said, "When our team waits for the next tile, I don't want to talk to friends about other things." E5 commented, "In normal

review, I answer and then I forget, but in this game I keep waiting for what comes next." E7 said, "I want English class to go faster to the game part."

The teacher reported a similar pattern from a classroom-management perspective: "In the Baamboozle lessons, I had to manage excitement, but I almost never had the problem of students switching off completely during the review stage." This account is consistent with the observation notes, which repeatedly recorded board-watching, forward orientation, and turn-tracking during the Baamboozle lessons.

Comparison-class students described the review phase in more neutral terms. C1 said, "The lesson is okay, but sometimes I wait a long time." C3 noted, "I answer when teacher calls me, but I don't always want to raise my hand." C5 stated, "Worksheet time is quiet, but not exciting." The qualitative corpus therefore portrayed the experimental review phase as more attention-holding, while the control review phase was more often described as acceptable but comparatively static. This pattern was not limited to a few especially lively moments, but appeared repeatedly across the observed Baamboozle lessons.

Observation notes frequently recorded students looking toward the projected screen before their own turn arrived, following the selection of tiles by other teams, and reacting audibly to whether peers answered correctly or not. In several lessons, the review phase appeared to hold the attention of a wide portion of the class at the same time, rather than only of those immediately called on to speak. This pattern is also consistent with the observation summary reported earlier, in which the Baamboozle group showed a higher estimated on-task participation rate (87.6%) and fewer average off-task incidents per lesson (2.9) than the comparison group (74.2% and 5.4, respectively). Descriptively, the review phase in the Baamboozle class therefore appeared not only more enjoyable, but also more collectively attention-holding across repeated lessons.

A further difference was visible in the temporal rhythm of the two review formats. In the Baamboozle lessons, students often appeared to remain mentally

involved between turns, because each tile selection created a new moment of anticipation. By contrast, in the comparison lessons, observation notes more often recorded waiting, quieter passivity, or attention that rose mainly when the teacher directly nominated a student. Some comparison-class students still described the lessons as acceptable and manageable, but the qualitative corpus suggests that the ordinary review phase was less likely to sustain whole-class alertness over time. The contrast in this theme, therefore, was not between engagement and complete disengagement, but between a review stage that repeatedly renewed attention and one that more often depended on teacher prompting to reactivate it.

4.2.4 Theme 2: Perceived competence and confidence

A second theme concerned the way students described confidence, answer readiness, and the experience of making mistakes. Baamboozle-class students frequently referred to feeling less afraid to answer and more willing to continue after uncertainty. E2 stated, "When my answer is right, I feel I can do English." E4 said, "If we are wrong, teacher tells us and the next one comes, so I still want to try again." E6 commented, "I'm not so scared because my team helps me before I say it." E8 said, "Before, I didn't like answering in front. In the game I can say it because it feels like we all answer."

The teacher interview supported the same tendency. T1 reported that some students who were usually reluctant became more willing to attempt an answer during the Baamboozle lessons. Observation notes also recorded repeated episodes in which learners prepared an answer together, responded, received immediate confirmation or correction, and remained engaged with the next turn.

Comparison-class accounts referred more often to confidence as something linked to already knowing the answer or receiving teacher approval. C2 said, "I like it when teacher says 'good'." C4 stated, "I am happy when I know the answer first." These comments show that positive feelings existed in both conditions, but the way students described them differed.

The observation data added further detail to this theme by showing that confidence was often visible not only in successful answers, but also in students' willingness to remain involved after uncertainty. Across the Baamboozle lessons, observation notes repeatedly recorded short peer prompts before answering, brief teacher confirmation or correction, and rapid movement to the next turn without prolonged interruption. In that process, hesitation did not always lead to withdrawal. Some students paused, looked to teammates, or began tentatively, but still completed an answer with support. This is important descriptively because the competence pattern in the Baamboozle condition appeared to include recoverability after difficulty, not only confidence when the answer was already known.

The comparison-class data showed a narrower pattern. Confidence was certainly present in that class as well, but it was more often described in relation to individual certainty, teacher approval, or already knowing the correct response. Observation notes from the comparison lessons more often showed that uncertainty led to silence, waiting for teacher direction, or a shift to another student. In this sense, the two groups did not differ simply in whether students could feel successful, but in the conditions under which success appeared accessible. The qualitative corpus therefore suggests that the Baamboozle review format made participation feel more manageable for some students who might otherwise have remained hesitant.

4.2.5 Theme 3: Relatedness through team participation and peer support

The third theme was especially strong in the Baamboozle-class interviews. Students repeatedly described the review phase as something done with others rather than only in front of others. E1 said, "My friends help me remember the word." E3 said, "We want our team to win, so we listen together." E5 noted, "I like when my friend says, 'you can answer it.'" E7 explained, "If I forget, my team gives me the first sound." E8 said, "I am less shy when my team is with me."

The teacher summarized the same pattern in comparative terms: "In normal review, some children answer and others watch. In the game, even when only one child speaks, several are involved." This matches the numerical contrast in the observation summaries, where the average number of peer-help utterances per lesson was 9.4 in the Baamboozle group and 3.0 in the comparison group.

Comparison-class learners also mentioned classmates, but their comments were less strongly tied to the formal review task. C6 said, "I like sitting with my friends." C1 commented, "Sometimes my friend tells me the answer quietly." These remarks show that peer presence existed in both classes, but the Baamboozle condition generated more explicit descriptions of encouragement, cueing, and collective attentiveness during review.

Observation data strongly supported the interview-based picture of peer-supported participation. As reported in the observation summary, the Baamboozle group produced an average of 9.4 peer-help utterances per lesson, compared with 3.0 in the comparison group. These peer-help behaviors included reminding classmates of words, giving the first sound, whispering possible answers, encouraging a teammate to speak, and reacting collectively to whether an answer was correct. Such episodes appeared regularly across the Baamboozle lessons rather than as isolated incidents. Descriptively, this suggests that team participation changed the social organization of the review phase: the activity was often experienced as something shared within a small group, even when only one student was speaking aloud at a given moment.

By contrast, the comparison lessons showed a more limited form of peer involvement. Classmates were not absent from the process, and some support was still visible, but it appeared less structured and less frequent. The review stage more often remained organized around teacher question and individual student response, with classmates positioned mainly as listeners rather than as active co-participants in the same turn. The Baamboozle condition therefore appeared to widen the social basis of participation. Instead of answering only in front of others,

many students described and displayed participation with others, which was one of the most consistent patterns in the qualitative corpus.

4.2.6 Theme 4: Motivation with conditions: noise, pacing, fairness, and technological constraints

The qualitative corpus also contained substantial qualifying evidence. T1 reported, "Excitement was good for participation, but sometimes it rose too fast. I had to stop and reset the class." The teacher also noted that question preparation required additional time and that internet delay reduced lesson momentum when it occurred. These comments show that the intervention did not function as a friction-free classroom tool.

Students raised conditions on participation as well. E4 said, "I like the game, but sometimes I want more turns." E6 noted, "If one friend is very fast, teacher must stop them." These statements indicate that increased engagement did not eliminate turn-balance issues. Students still noticed whether participation opportunities were distributed fairly.

On the comparison side, several learners described benefits of calmer classroom work. C2 said, "I like quiet work sometimes," and C5 stated, "When teacher explains slowly, I understand more." The corpus therefore documented a stronger motivational profile in the Baamboozle group together with clear conditions, trade-offs, and continuing strengths in quieter teacher-led review.

At the same time, the qualitative corpus did not portray the Baamboozle review format as unproblematically positive. Observation notes repeatedly showed that the stronger energy of the Baamboozle lessons also required more active teacher management. In some lessons, excitement rose quickly after point changes, surprising answers, or disagreement over whose turn it was. The teacher interview reflected the same pattern, noting that participation was stronger but also harder to regulate smoothly. This means that the higher engagement described in Themes 1 to 3 was accompanied by practical classroom costs. The review phase could become noisier and faster-moving, and the benefits of the format appeared to

depend partly on how effectively the teacher controlled transitions, turn-taking, and response tempo.

A second qualifying pattern concerned equity and consistency across students. Although the Baamboozle group generally appeared more participatory, observation notes and student comments also suggested that not every child experienced the review stage in exactly the same way. Faster or more outspoken students could sometimes dominate response opportunities if turns were not carefully distributed, while quieter students still needed support to enter participation. In addition, a few comments in the qualitative corpus indicated that calm explanation and quieter work still had value for some learners, especially when they wanted more time to think. Technological conditions also remained relevant. Because the activity depended on projection and stable classroom technology, interruption or delay could weaken the flow that otherwise supported attention and anticipation. The theme of “motivation with conditions” therefore shows that the motivational advantage of the Baamboozle lessons was real in the corpus, but not automatic or free of constraint

4.2.7 Interview report summaries

The summary reports in Appendix I showed a clear internal pattern across the eight Baamboozle-class interviews. E1 emphasized tile selection and peer support. E2 focused on feeling more capable when answers were confirmed quickly. E3 highlighted anticipation and not wanting to miss the next turn. E4 stressed that mistakes felt less frightening in team format. E5 valued game rhythm and peer encouragement. E6 reported more confidence when classmates whispered support. E7 strongly preferred game-based review to worksheet review. E8 described reduced fear when answering within a team.

The six comparison-class summaries showed a different profile. C1 considered ordinary review acceptable but less exciting. C2 preferred calm explanation at times and valued teacher praise. C3 participated when called on but

rarely volunteered spontaneously. C4 linked confidence to already knowing the answer. C5 found worksheets manageable but not interesting. C6 appreciated classmates but described less active group involvement during review.

The teacher summary integrated both class experiences. T1 reported increased volunteerism, stronger visible engagement, and more peer support in the Baamboozle group, together with greater noise and higher preparation demands. Because the same teacher observed both classes across the intervention period, this comparative teacher account added an important cross-group perspective to the student reports.

4.2.8 Qualitative answer to RQ2

In direct answer to RQ2, students and the teacher experienced Baamboozle-integrated review as more engaging, confidence-supportive, and peer-supported than teacher-centered review, but not as unconditionally superior. The qualitative evidence showed that the tool widened participation and reduced the emotional risk of answering, while its effectiveness depended on teacher management of noise, pace, fairness, turn balance, and technology.

The observation data supported these perceptions with repeated behavioral evidence. In the Baamboozle lessons, the review phase was more often characterized by visible positive affect, higher on-task participation, more frequent peer-help utterances, and stronger whole-class attentiveness during turn-taking. Students did not appear to participate only when directly nominated by the teacher; rather, many seemed to remain involved while waiting, anticipating upcoming turns, reacting to other teams' answers, and supporting classmates during response opportunities. The interview data added further detail by showing that this stronger participation was closely tied to students' experience of answering in teams, receiving immediate feedback, and feeling that mistakes were easier to recover from in the gamified review format. By contrast, the comparison-class data described a calmer and more orderly review phase, but one that was generally less interactive and less likely to draw broad participation across the class.

4.3 Descriptive Mixed-Methods Joint Display

The mixed-methods design required more than separate reporting of quantitative and qualitative findings. A descriptive joint display was therefore used to place the two strands side by side. The purpose of the display was not to force identical meanings onto all data points, but to show where the strands converged, where one strand elaborated the other, and where areas of qualification remained visible.

Table 4.11 Descriptive mixed-methods joint display

Quantitative pattern	Qualitative pattern	Descriptive convergence
Higher post-test motivation in Baamboozle class	Baamboozle-class students and the teacher described stronger enjoyment, anticipation, and sustained attention	Both strands favored the Baamboozle condition
Larger mean gain in the Baamboozle class (6.37 vs. 2.93)	Baamboozle-class interviews repeatedly emphasized wanting turns, waiting for the next tile, and staying with the activity	The larger gain aligned with stronger reported lesson involvement
Higher autonomy mean	Students referred to choosing numbers, wanting turns, and remaining visually engaged with the board	Reports of involvement aligned with the autonomy pattern
Higher competence mean	Students said they felt less scared, more ready to try, and more able to continue after mistakes	Reports of answer confidence aligned with the competence pattern
Higher relatedness mean	Students and the teacher described peer prompting, team support, and shared participation	Reports of collaboration aligned with the relatedness pattern
Higher volunteerism, peer-help, positive affect, and on-task participation in observations	Baamboozle-integrated English review was described as lively and collectively interactive	Behavioral and interview evidence moved in the same direction
Both classes improved	Comparison-class students still valued teacher praise,	Improvement was not restricted to one

	calm explanation, and manageable worksheet work	class
Raw score tables showed unchanged and declining cases; qualitative data showed noise, fairness, pacing, and technology constraints	The intervention was described positively but not unconditionally	Neither strand showed a perfectly uniform positive effect

Table 4.11 shows that the strongest points of convergence occurred around the overall motivational advantage in the Baamboozle class, stronger gains over time, and the three SDT-related classroom processes of involvement, confidence, and peer-supported participation. The quantitative strand established the direction and relative strength of the difference, while the qualitative strand clarified how that difference was experienced during the review stage. In this sense, the strands converged not only at the level of outcome, but also at the level of participation conditions.

The table also shows complementarity more clearly when read against the Chapter 1 problem. The quantitative findings indicate that the Baamboozle class reported stronger classroom motivation, but the qualitative findings explain why that pattern matters pedagogically. Students did not simply report liking the activity more. Rather, they described a review stage that felt easier to enter, less intimidating, and more socially shared (Krashen, 1982). The mixed-methods integration therefore suggests that Baamboozle-integrated English review was associated not only with higher motivation scores, but also with broader and more psychologically safe readiness to participate during review.

At the same time, the joint display preserves qualification. Both classes improved, which means that teacher-centered review remained educationally meaningful within the case. Moreover, the qualitative strand retained recurring concerns about noise, pacing, fairness, turn balance, and technological reliability. These tensions prevent the integrated interpretation from collapsing into generic

praise of gamification. Instead, they indicate that the motivational advantage of Baamboozle was conditional on how the review stage was managed.

The joint display therefore prepares the transition to Chapter 5 in a more explicit way. Taken together, the two strands suggest that the central issue is not whether a digital tool is enjoyable in a general sense, but whether it reorganizes the review stage in ways that widen entry into participation, reduce the emotional cost of answering, and support peer-connected involvement within a bounded Vietnamese public primary-school case.

4.4 Chapter Summary

This chapter has presented the findings in a manner aligned with the two research questions and the analytical procedures described in Chapter 3. Quantitatively, the two classes were equivalent at pre-test, both classes improved over time, and the Baamboozle class showed the larger gain, the stronger post-test result, and higher scores on autonomy, competence, and relatedness. The raw score tables also revealed unchanged and declining cases in both classes, although such cases were more frequent in the comparison class.

Qualitatively, Baamboozle-integrated English review was described as more attention-holding, confidence-supportive, and socially interactive, while the qualitative corpus also retained constraints related to noise, pacing, fairness, preparation demands, and internet reliability. The joint display then showed substantial convergence between the two strands without erasing qualification. Chapter 5 now turns from reporting to interpretation in order to explain what these findings mean for the study's research questions, theoretical framework, and pedagogical contribution.

CHAPTER 5: DISCUSSION AND CONCLUSION

5.1 Discussion of RQ1: Was Baamboozle Associated with Stronger Classroom Motivation Within the Case?

The quantitative answer to RQ1 is affirmative, but the significance of that answer lies in more than the fact that one class scored higher than the other at the end of the study. The first point that matters is the near-perfect baseline equivalence between the two intact classes. The pre-test comparison was effectively null, $t(58) = -0.04$, $p = .967$, $d = -0.01$. In a school-based case study with an embedded class comparison, this does not create randomization after the fact, but it does narrow one major rival explanation. The later difference is therefore more plausibly associated with what happened during the intervention period than with an inherited motivational advantage in the Baamboozle class.

Against that starting point, the post-test difference of 3.40 points in favor of the Baamboozle group, together with a large effect size of $d = 1.18$, is analytically substantial. The result suggests not merely that Baamboozle made students report slightly more favorable feelings, but that the review phase came to be experienced under meaningfully different motivational conditions within this case. Because the study measured motivation as classroom readiness and willingness to participate, a post-test gap of this size indicates an educationally important divergence in the likelihood that learners would enter, remain in, and respond during review.

The internal structure of the quantitative findings strengthens that conclusion. Both groups improved, which is exactly what one would expect in two functioning classes taught over six weeks by the same teacher. However, improvement was not equivalent. The Baamboozle group gained 6.37 points, while the comparison group gained 2.93, and the within-group effect size in the Baamboozle class ($dz = 1.49$) was approximately double that of the comparison class ($dz = 0.75$). The embedded comparison therefore showed added motivational

value over an active instructional alternative that itself remained educationally meaningful.

The gain-score pattern adds a further layer because it prevents the interpretation from resting entirely on average scores. Improvement in the Baamboozle group was broader as well as larger. Twenty-six of thirty students improved, compared with twenty-one of thirty in the comparison class. More importantly, medium and large gains clustered far more densely in the intervention group. The pattern therefore appears as a class-wide shift in participation readiness rather than as an artifact generated by a small number of exceptional performers.

At the same time, the raw data prevent a simplistic or triumphalist reading. One student in the Baamboozle group declined, three showed no change, and several comparison-class students improved strongly despite the absence of the gamified review format. These cases matter because they show that Baamboozle did not operate as a deterministic motivational trigger. Its role in this study was probabilistic and conditional. It increased the likelihood, breadth, and magnitude of gain within the case, but it did not erase individual differences, classroom history, or the continuing value of conventional teaching.

The comparison-class improvement deserves careful treatment. Teacher-centered review remained an active and legitimate instructional format. Repetition, praise, increasing familiarity with routines, and continued exposure to the same teacher could all contribute to higher post-test motivation. Rather than weakening the intervention finding, this actually strengthens it. Because the comparison condition was educationally meaningful, the Baamboozle advantage cannot easily be dismissed as a weak-control artifact.

The SDT subdimension profile deepens the interpretation further. The post-test advantage appeared across autonomy, competence, and relatedness, with large effect sizes ranging from 0.95 to 1.02. This distribution is theoretically important because it suggests that the intervention did not produce only a short-lived rise in

fun or excitement. Instead, within this case, it was associated with simultaneous differences in involvement, answer confidence, and social support.

This structural reading is especially plausible in a Grade 2 EFL context. Very young learners are unlikely to sustain effort because of distant future self-images or long-term instrumental calculations. Their participation is more immediately shaped by whether the task feels accessible, whether a turn seems worth taking, whether mistakes are survivable, and whether peers make entry easier or harder. In this case, Baamboozle appears to have strengthened motivation not by changing what students learned, but by changing how the review phase positioned them as participants.

In direct answer to RQ1, then, the thesis supports the conclusion that Baamboozle-integrated English review was associated with stronger classroom motivation than teacher-centered review within this bounded case. That conclusion is supported by baseline equivalence, the larger post-test mean, the stronger gain pattern, and the parallel advantage across autonomy, competence, and relatedness. At the same time, the claim remains properly bounded: the study does not show that Baamboozle universally causes higher motivation in all settings, but that within one Vietnamese public primary-school case it was associated with a review structure that appears to have supported wider, stronger, and more sustained readiness to participate. Nevertheless, the convergence of baseline equivalence, post-test separation, within-group change, gain-score contrast, and subdimension advantage makes the overall quantitative conclusion robust within those bounds. The quantitative answer to RQ1 can therefore be stated descriptively as follows: during the six-week intervention period, the Baamboozle class showed stronger short-term motivational growth than the comparison class on the study's motivation measure. That advantage was visible in baseline-to-post-test change, in end-point comparison, in gain-score analysis, in the three SDT subdimensions, and in the wider distribution of individual improvements across the Baamboozle class. The quantitative strand therefore supports the conclusion that Baamboozle-

integrated English review was associated with a stronger positive motivational pattern than teacher-centered review within the study period and sample.

5.2 Discussion of RQ2: How Was Baamboozle Experienced?

The qualitative findings show that Baamboozle was experienced not simply as enjoyable, but as a reorganization of classroom review that changed how participation felt. If the tool had merely made lessons more entertaining, the educational meaning of the finding would be relatively shallow. What participants described, however, was something more specific: review became more eventful, more shared, and less intimidating. Observation data documented corresponding increases in volunteer responses, peer-help behavior, positive affect, and on-task participation, while interview data showed how students understood those changes from inside the classroom process.

One of the clearest experiential changes was anticipatory attention. Baamboozle-class students did not only say that they liked the activity after the fact; they described looking at the screen, waiting for the next tile, and staying oriented to the lesson because the next turn might involve them. This is analytically significant because it shows that the review phase ceased to be a passive interval between instructional moments and became itself a site of sustained expectancy. For young learners, such expectancy matters because attention in repetitive classroom work is often fragile and easily lost.

The observation results support that reading in concrete terms. The Baamboozle class averaged 11.8 volunteer or hand-raise episodes per lesson compared with 5.1 in the comparison group, and visible positive-affect episodes were likewise more than twice as frequent. These behavioral traces do not by themselves reveal students' internal states, but in combination with the interview reports they strongly suggest that the review phase in the intervention condition became something students actively wanted to enter rather than merely endure.

A second major experiential shift concerned emotional risk. Students repeatedly framed the game format as making it easier to answer even when they

were unsure. Error was not described as disappearing, but as becoming more recoverable. Teacher feedback came quickly, the activity moved on, and team membership reduced the sense of solitary exposure. In lower-primary classrooms, this is a highly consequential change because willingness to answer is often limited less by total ignorance than by fear of visible failure.

The teacher's account gives this pattern additional credibility. T1 did not claim that all learners became equally outspoken or that anxiety vanished altogether. Instead, the teacher observed that typically reluctant students entered the activity more often. That is a more restrained and therefore more defensible pedagogical claim. It suggests that Baamboozle lowered the threshold of entry into participation, which is precisely where motivational change becomes educationally meaningful for young learners.

A third experiential pattern was the centrality of peer support. Students described teammates as helping them remember words, supplying cues, encouraging answers, and making them feel less shy. In the Baamboozle class, peer involvement was not peripheral to the task; it was part of how the task operated. The observation count of 9.4 peer-help utterances per lesson, compared with 3.0 in the comparison group, indicates that this was not merely a retrospective impression or isolated anecdote. It was a repeated classroom practice.

The qualitative strand also clarifies why the findings should not be reduced to an uncritical celebration of gamification. Students and the teacher identified real conditions on effectiveness: excitement could escalate into noise, faster pupils could dominate turns, preparation demands were higher, and technology interruptions could break momentum. These constraints matter because they show that the platform functioned as a managed pedagogical structure, not as an automatically self-motivating tool.

Comparison-class interviews provide an additional qualification that strengthens rather than weakens the analysis. Several students did not reject ordinary review; instead, they valued its calmness, slower explanation, or

manageable pace. The qualitative evidence therefore suggests something more precise than a simple contrast between innovation and tradition: Baamboozle appears especially strong for energizing participation and peer-supported retrieval, whereas quieter teacher-led review may retain value for slower processing and direct explanation.

The qualitative findings provided a clear but carefully qualified answer to RQ2. Across the qualitative corpus, the use of Baamboozle in the Baamboozle class was consistently described as making the review stage more exciting, more participation-oriented, more confidence-supportive, and more socially interactive than teacher-centered review. This pattern was visible not only in student interview comments, but also in repeated classroom observations and in the teacher's cross-class comparison. Students in the Baamboozle class commonly described the review stage as enjoyable, active, and easier to join. They frequently referred to wanting turns, watching the screen attentively, listening to teammates, and feeling happier or less afraid when answering. The teacher likewise reported broader volunteerism, stronger visible enthusiasm, and more active peer involvement during Baamboozle-integrated English review than during the more conventional review format used in the comparison class.

At the same time, the qualitative answer to RQ2 remained conditional rather than absolute. The corpus did not portray Baamboozle as producing only positive outcomes or as benefiting all students in exactly the same way. Observation notes and interview comments also documented increased noise, occasional imbalance in turn-taking, the need for careful pacing, and the teacher's additional management work in keeping the activity fair and orderly. The teacher further indicated that the success of the review format depended on preparation, turn control, and stable classroom technology. In addition, some students in the comparison class described the value of quiet explanation, slower pacing, and

conventional teacher support, suggesting that teacher-centered review still provided a comfortable learning space for certain learners.

Taken together, the qualitative strand showed that Baamboozle-integrated English review was generally experienced as a stronger motivational format within this case, especially in relation to active participation, answer confidence, and peer-supported involvement. However, the same qualitative evidence also showed that these advantages were shaped by classroom conditions rather than guaranteed by the tool itself. The qualitative answer to RQ2, therefore, is that students and their teacher largely experienced Baamboozle as a more engaging and socially supportive review format than teacher-centered review, while also recognizing that its benefits depended on effective classroom management, balanced participation, and workable technological conditions.

5.3 Interpretation Through Self-Determination Theory

5.3.1 *Autonomy*

The autonomy findings require careful handling because autonomy in a Grade 2 classroom cannot be understood as unrestricted freedom or curricular self-direction. Yet Self-Determination Theory does not require absolute independence in order for autonomy support to be meaningful. What matters is whether learners experience participation as something they can enter with some degree of volition rather than something simply imposed on them.

In this study, autonomy took a developmentally bounded form. Students chose tiles, anticipated turns, tracked the unfolding activity, and experienced themselves as visible participants rather than passive recipients. Quantitatively, this pattern is reflected in the higher post-test autonomy mean for the Baamboozle group (9.57 versus 8.33, $d = 1.02$). Qualitatively, it appears in students' language of wanting turns, watching the board, and wanting the lesson to reach the game phase.

This is especially important in the Vietnamese primary-school context described in the earlier chapters, where teacher authority and whole-class organization remain central. In such settings, autonomy support is unlikely to resemble open-ended individual choice. It is more likely to appear in smaller forms such as turn ownership, bounded selection, and meaningful contribution within teacher-managed structure. The findings suggest that Baamboozle created exactly this kind of bounded autonomy.

At the same time, autonomy support was not automatic. E4's wish for more turns and E6's concern about faster classmates reveal that a sense of involvement could still become unevenly distributed. Autonomy here should therefore be understood not as a built-in property of the platform, but as a classroom condition produced by the interaction between tool design and teacher orchestration.

5.3.2 Competence

Competence may be the most pedagogically decisive SDT dimension in lower-primary language learning because children who repeatedly experience failure or embarrassment may withdraw quickly. The competence findings were strong in both numerical and qualitative terms. The post-test competence mean was 9.50 in the Baamboozle group and 8.50 in the comparison group, with $d = 0.98$. Students repeatedly described the game format as making them more ready to answer and less afraid of getting things wrong.

From an SDT perspective, this pattern makes sense because competence is supported not only by success itself, but by feedback structures that allow learners to act effectively and recover after error. E2's statement, "When my answer is right, I feel I can do English," reflects visible success. E4's comment about teacher feedback followed by the next question reflects recoverability. Together, these accounts show that competence in the intervention condition was not only about being correct; it was also about being able to remain in the activity after uncertainty.

A particularly important feature of the present data is that competence was often socially scaffolded. E6 and E8 both described feeling more able to answer because classmates helped before the response was spoken aloud. For very young learners, feeling capable can depend on the classroom structure that surrounds the answer attempt. The team format lowered exposure, distributed risk, and made success seem more reachable.

The comparison-class data sharpen this contrast. C2 valued teacher praise, and C4 linked confidence to already knowing the answer. These are legitimate sources of competence support, but they imply a more correctness-centered experience. In the Baamboozle condition, competence was described more often as a process involving quick confirmation, retry opportunities, and peer-supported entry.

The competence results should nonetheless be interpreted with restraint. The study measured motivation rather than objective language achievement. It cannot show that students became more proficient simply because they reported higher confidence and willingness to answer. What it does show is that the intervention increased perceived capability and readiness to attempt responses.

5.3.3 Relatedness

Relatedness was the highest post-test subdimension in the Baamboozle group, with a mean of 9.77 compared with 8.57 in the comparison group and an effect size of $d = 0.95$. This is a major finding because relatedness is sometimes underplayed in discussions of classroom technology, yet it may be especially central for very young learners. Children often participate through others before they participate confidently on their own.

The qualitative evidence was highly consistent on this point. Baamboozle-class students repeatedly described peer support, team listening, encouragement, cueing, and reduced shyness through group membership. The teacher reinforced the same pattern by observing that in the game several students remained involved even when only one child was speaking. The observation count of peer-help

utterances gives these perceptions further weight by showing that collaboration was regularly audible during the lessons.

These results complicate a common critique of gamification, namely that competition inevitably individualizes performance and produces exclusion. In the present study, competition did not disappear, but it was nested within teamwork. Students wanted their team to win, yet the desire to win often activated listening, prompting, and encouragement. Competition therefore appeared to function productively only because it was socially distributed rather than individually isolating.

However, the relatedness findings also contain a warning. Team membership by itself does not guarantee meaningful inclusion. Quieter children can still become symbolic members if stronger classmates dominate the interaction. The fairness concerns in the interviews therefore matter here too. Relatedness in this study was not simply the presence of groups; it was the degree to which group membership genuinely supported participation.

5.3.4 An Integrated SDT Interpretation

Taken together, the findings suggest that Baamboozle worked not as a generic entertainment device but as a participation architecture that touched all three SDT needs at once. Bounded choice and visible turn entry supported autonomy at a developmentally appropriate level. Immediate checking, retry logic, and reduced public exposure supported competence. Team-based response and peer prompting supported relatedness. These features operated interdependently rather than as separate classroom effects.

This interdependence is important because it helps explain the Chapter 1 problem of uneven participation in ordinary review. When review is predominantly teacher-fronted and rapid, some learners may experience weak autonomy because they enter only when selected, weak competence because error feels public, and weak relatedness because classmates function mainly as an audience rather than

as support. Under those conditions, participation readiness is likely to narrow around the already confident few. In the present case, Baamboozle appears to have altered that pattern by supporting autonomy, competence, and relatedness together during the same review stage.

The integrated SDT reading also clarifies why the qualitative findings are more important than simple statements that students liked the game. A child who experiences bounded agency is more willing to enter the task. A child who experiences peer support is more willing to risk an answer. A child who can answer, receive immediate feedback, and continue after error is more likely to feel capable rather than exposed. These processes reinforce one another. Motivation was therefore strengthened not by adding excitement from outside the lesson, but by reorganizing the internal conditions under which participation became possible and sustainable.

This synthesis also speaks directly to the question of safer and more equitable participation. In this case, autonomy made participation more invitational, competence made it more survivable, and relatedness made it more socially supported. Together, these need-supportive conditions help explain why the Baamboozle class showed stronger motivation and why students described the review stage as easier to join and less intimidating. The study therefore suggests that, for very young learners, uneven participation in ordinary review may be understood as a need-support problem as much as a classroom-management problem.

At the same time, the integrated interpretation remains bounded. The study does not show that every gamified review task will necessarily support all three needs, nor that Baamboozle did so perfectly in every lesson. Noise, turn imbalance, pacing pressure, and technological interruption all show that need support can be weakened in practice. Even so, the most coherent explanatory model emerging from the study is that Baamboozle-integrated English review reorganized the review stage in ways that more often supported autonomy, competence, and

relatedness simultaneously, and that this integrated need support was associated with stronger short-term motivation within this Vietnamese lower-primary case.

5.4 Comparison With Previous Studies

The findings broadly support the literature reviewed in Chapter 2, but their principal value lies in sharpening rather than merely repeating prior claims. Yip and Kwan (2006), for example, argued that game-based activity can make repetitive language practice more engaging. The present study supports that proposition while specifying the mechanism more precisely. Repetition did not become less repetitive in content terms; it became more motivationally sustainable because participation opportunities were redistributed and socially intensified.

The study also resonates with Reinders and Wattana's (2015) work on willingness to communicate in digital environments, although the comparison must be made carefully because the present participants were much younger and the tasks were teacher-mediated whole-class review activities rather than broader communicative game environments. Even so, a similar logic seems to operate: participation increased because the local emotional cost of entering classroom interaction was lowered.

The findings are likewise consistent with synthesis work, which reported generally positive effects of quiz-like platforms on engagement and classroom atmosphere. However, the present thesis goes beyond atmosphere language by offering a theory-based account of what improved. Motivation did not rise only as visible excitement; it rose through a classroom arrangement that made learners feel more involved, more able to try, and more socially held within the task.

Games are most effective when challenge and skill are meaningfully calibrated. The present study supports that position indirectly. Students were willing to retry, and the teacher reported broader participation, which implies that challenge remained within an acceptable range for most learners. At the same time, recurring concerns about pacing and turn balance show that this calibration was

not a property of the platform alone. It depended on teacher mediation and management discipline.

Butler's (2019) emphasis on age-appropriate game design is particularly well supported here. The motivational value of Baamboozle did not depend on sophisticated individual interactivity or complex mechanics. It depended on visual simplicity, collective attention to a shared screen, predictable turn structure, and opportunities for peer-supported oral response.

The present study also extends smaller-scale work more directly focused on Baamboozle. Nguyen (2021) reported positive classroom value for Baamboozle in vocabulary-oriented teaching, but the present thesis adds several layers not always visible in such classroom-use accounts. It places the platform within a mixed-methods comparison-class design, interprets the effects through SDT rather than generic enjoyment language, and documents not only benefits but also contradictions and implementation constraints.

At the same time, the thesis qualifies more celebratory strands of the gamification literature. Wang's (2015) warning about wear-out effects remains salient. A six-week intervention is long enough to move beyond a one-lesson novelty burst, but it is still not long enough to demonstrate semester-long stability. The study therefore supports a strong short-term conclusion while withholding any claim of durable motivational permanence.

Overall, then, the study both confirms and disciplines previous literature. It confirms that gamified review can support motivation, but it disciplines that claim by locating the effect in a specific age group, a specific classroom ecology, and a specific explanatory framework centered on autonomy, competence, and relatedness.

5.5 Triangulation and Strength of Evidence

One of the strongest features of the study is the convergence between its quantitative and qualitative strands. The questionnaire data showed higher post-test motivation in the Baamboozle group, larger gains over time, and advantages

across all three SDT subdimensions. The observation data showed more volunteering, more peer-help behavior, more visible positive affect, higher estimated on-task participation, and fewer off-task incidents. The interviews then explained these patterns in participant terms: students wanted turns, felt less afraid, and experienced participation as socially shared.

This convergence matters because each data source compensates for the limitations of the others. Questionnaire data alone might be questioned because the participants were very young and the instrument was simplified for developmental appropriateness. Observation data alone might show energy without clarifying what that energy meant to learners themselves. Interview data alone might produce memorable quotations without demonstrating whether they reflected a broader class-level tendency. Together, however, the three strands create a more persuasive evidential structure than any one source could provide independently.

The study also benefits from triangulation at the level of mechanism, not only at the level of direction. Higher autonomy scores aligned with students' references to selecting numbers, wanting turns, and staying visually engaged with the board. Higher competence scores aligned with comments about being less scared and more willing to try again after mistakes. Higher relatedness scores aligned with reports of peer prompting and team support. The observation summary then provided behavioral traces of those same processes.

At the same time, the triangulation was not artificially unanimous. The mixed-methods evidence also preserved tension and qualification. Quantitatively, the comparison group improved even without the intervention. Qualitatively, some comparison-class students valued calm explanation and quiet work. The Baamboozle group showed one decline, several no-change cases, and recurring concerns about noise, fairness, pacing, and technological interruption.

The strength of evidence can therefore be described as strong for a small-scale, school-based, mixed-methods case study with embedded quantitative

comparison. The design does not support the kind of universal causal claim associated with large randomized trials. Yet it does support a credible, internally coherent, and well-substantiated conclusion that Baamboozle-integrated review was associated with stronger short-term classroom motivation in this case and that this association is meaningfully explained by the classroom processes documented in the qualitative strand.

5.6 Pedagogical Implications

The first pedagogical implication is that classroom motivation in lower-primary EFL settings should be treated as something that can be designed during ordinary lesson phases, not simply as a stable trait that children either bring or do not bring to class. In this thesis, the strongest motivational differences emerged not from changing the syllabus or introducing individualized technology, but from reorganizing the review stage. This suggests that teachers who wish to widen participation do not necessarily need a new curriculum. They need review structures that make more learners willing and able to enter classroom interaction.

The second implication is that Baamboozle appears most useful when used strategically for the review stage of Grade 2 English lessons rather than as a substitute for all lesson stages. The findings do not support any broad claim that gamification should dominate primary English teaching. Explanation, modeling, slower clarification, and controlled practice still remain important, and some learners in the comparison class clearly valued calmer teacher-led support. What the present study supports is a more specific pedagogical claim: Baamboozle-integrated English review can function effectively as a consolidation format that re-energizes textbook-aligned review and broadens participation readiness during that bounded lesson stage.

Third, the study has concrete implications for Vietnamese public primary-school classrooms, where projector-based whole-class teaching remains common and one-device-per-student assumptions are often unrealistic. The results suggest

that meaningful digital enhancement is possible even under modest infrastructural conditions when the tool fits local classroom ecology. A teacher can use one projected screen, intact classes, and textbook-linked review content while still creating a more motivating review stage. This is practically significant because it makes the pedagogical value of the study locally realistic rather than dependent on ideal technological conditions that many schools do not have.

Fourth, the findings show that teacher mediation remains central. Baamboozle did not work as an independent motivational agent. Its value depended on how the teacher distributed turns, set pacing, managed noise, controlled dominance, and protected weaker learners from embarrassment. In practical terms, this means that the teacher should pre-plan turn rotation, build in short think time, actively invite quieter pupils, and permit limited team whisper support so that collaboration scaffolds participation rather than replacing it. Without such mediation, the same platform could easily become faster, louder, and less equitable rather than more inclusive.

Fifth, fairness should be treated as a core design principle. Because one of the aims of the study is broader and more equitable participation, teachers should not assume that team play automatically produces inclusion. Teams need to be organized carefully, and turn-taking needs to be structured so that quieter or lower-confidence learners are not hidden behind stronger classmates. In classrooms like the one studied here, psychologically safe participation may require explicit norms: no mocking of wrong answers, supportive prompting, praise for trying, balanced chances to answer, and a pace that allows hesitant learners to stay in the activity.

Sixth, the findings imply that pacing matters as much as platform choice. A review stage can become demotivating not only when it is dull, but also when it becomes too fast, too noisy, or too competitive. The teacher therefore needs to maintain a tempo that preserves excitement without sacrificing comprehension or fairness. This is especially important in lower-primary classrooms, where attention

can be sustained by lively rhythm, but confidence can be undermined quickly when learners feel rushed.

Seventh, infrastructural constraints must be planned for rather than ignored. Because the study identified technological reliability as one of the conditions shaping implementation, teachers in similar Vietnamese public-school settings should prepare backup procedures for the review stage. If internet access fails or the platform becomes unstable, the teacher should still be able to continue with equivalent review prompts through printed cards, the board, or oral questioning. This practical point matters because the pedagogical value of Baamboozle depends partly on continuity and classroom flow.

Overall, the pedagogical implication of the thesis is not that gamification is inherently superior to conventional teaching. Rather, it is that Baamboozle-integrated English review appears to be a practical way of making the review stage more participatory, less intimidating, and more socially supportive within a bounded Vietnamese lower-primary context, provided that teachers use it with deliberate attention to pacing, fairness, teacher mediation, and infrastructural reality.

5.7 Contributions of the Study

5.7.1 Theoretical Contribution

The theoretical contribution of the study lies first in showing that Self-Determination Theory can be applied analytically and productively to the problem of young learner participation during English review, rather than used only as a general motivational backdrop. In this thesis, SDT does explanatory work. The findings are organized not around diffuse enjoyment, but around autonomy, competence, and relatedness as distinct yet interdependent conditions of classroom participation. This makes the account of motivation more precise and more consistent with the actual empirical pattern.

A second theoretical contribution is developmental. The thesis clarifies what SDT needs look like in a Grade 2 EFL classroom. Autonomy does not appear

as independence from the teacher, but as bounded turn ownership, visible entry into participation, and a sense of being involved in the review process. Competence appears not as mastery in a broad achievement sense, but as answer confidence, visible success, and recoverability after error. Relatedness appears as peer prompting, team support, and socially shared participation. This refinement matters because motivation theory developed with older learners does not always translate directly into the experiential world of seven- and eight-year-old children (Reeve, 2015).

A third theoretical contribution is that the study links SDT more explicitly to the Chapter 1 problem of uneven participation. The thesis suggests that broad and psychologically safe participation during review may be understood as a need-support issue. When review structures support autonomy, competence, and relatedness together, motivation becomes more widely distributed across the class. In that sense, the study contributes not only to motivation theory in general, but also to the more specific theorization of participation readiness in lower-primary TESOL..

5.7.2 Methodological Contribution

The methodological contribution of the study lies in the coherence of its alignment chain. The research problem, purpose, questions, hypotheses, objectives, theoretical framework, instruments, analyses, and discussion are all organized around the same construct logic. That coherence strengthens internal consistency and helps the thesis avoid a common weakness in small-scale classroom technology research, namely, a mismatch between broad claims and narrower evidence.

A second methodological contribution is the use of a mixed-methods case study with an embedded comparison between two intact classes. The study did not overstate itself as a randomized experiment, but it also did more than provide a purely descriptive classroom account. Baseline comparison, pre-test/post-test analysis, effect sizes, subdimension analysis, observation data, interviews, and a

mixed-methods joint display together produced a denser and more credible evidential structure than a single-source design would have allowed.

A third methodological contribution is the way the thesis integrates developmental appropriateness with construct discipline. The emoji-based questionnaire, repeated observations, and brief interviews were suitable for very young learners, yet the study did not treat age-appropriate data collection as an excuse for theoretical looseness. Instead, it used multiple forms of evidence to capture classroom motivation as participation readiness. This offers a useful model for future TESOL research with children in school-based settings..

5.7.3 Pedagogical Contribution

The pedagogical contribution of the study lies in clarifying not merely that Baamboozle can be motivating, but what pedagogical work it appears to do during the review stage of Grade 2 English lessons. Within this case, its main value was not content delivery, but participation restructuring. It made the review stage easier to enter, more attention-holding, less intimidating for many learners, and more socially supported through team-based interaction. This is a more useful contribution than generic claims that the platform is fun, because it specifies the classroom mechanism through which the tool appears to matter.

A second pedagogical contribution is that the thesis identifies the bounded lesson stage in which the tool appears most useful. The findings do not justify broad claims about Baamboozle across the entire English lesson. Rather, they show its promise as a review-stage tool for textbook-aligned consolidation in lower-primary classrooms. This distinction is important because it prevents pedagogical overgeneralization and helps teachers use the platform where it is most defensible and practical.

A third pedagogical contribution is that the study makes the value of the tool conditional and actionable. It shows that benefits depend on turn balance, pacing control, emotionally safe error handling, team structure, preparation, and

technological reliability. These conditions matter because they translate the findings into usable classroom principles. The thesis therefore contributes not just a positive verdict on a digital tool, but a more disciplined account of how that tool should be used if it is to support broad and equitable participation rather than merely produce noise or excitement.

5.7.4 Contextual Contribution

The contextual contribution of the study lies in bringing evidence from a Vietnamese public primary-school classroom into a literature that often gives greater visibility to older learners, highly resourced settings, or one-device-per-student models. By focusing on intact Grade 2 classes, one teacher, projector-based whole-class teaching, and a textbook-aligned review stage, the thesis documents a form of digital pedagogy that is ordinary, low-cost, and locally realistic.

A second contextual contribution is that the study shows how the value of a digital tool is inseparable from classroom ecology. In this Vietnamese lower-primary case, Baamboozle did not operate outside the realities of teacher authority, fixed class groupings, curriculum pacing, uneven infrastructure, and whole-class projection. The thesis therefore contributes evidence that is not only locally situated, but also locally interpretable. It shows how technology, pedagogy, and context intersect in actual public-school conditions.

A third contextual contribution is that the study offers a bounded but meaningful account of what digital support may look like in settings where infrastructural constraints are real. It suggests that pedagogically useful innovation does not have to depend on fully individualized technology. Within this case, a single projected screen and teacher-mediated team participation were sufficient to reorganize review in motivationally important ways. That is a contextually valuable contribution for Vietnamese primary education and for comparable public-school contexts beyond it..

5.8 Limitations of the Study

Despite its strengths, the study has several limitations that should shape the force of its claims.

First, the study was a mixed-methods case study with an embedded comparison between two intact classes rather than a randomized experiment. Although the pre-test results showed near-equivalence between the two classes, class-level differences cannot be ruled out as confidently as they could in a true experiment.

Second, the study was conducted in one school with one teacher. This strengthened procedural consistency across groups, but it also constrains transferability. Different teachers may vary in classroom management, digital confidence, and feedback style, while different schools may differ in technological reliability, student culture, and institutional expectations.

Third, the intervention lasted six instructional weeks. That duration is sufficient to demonstrate more than a one-lesson response, but it remains too short to resolve the question of sustainability. The possibility that part of the pattern was strengthened by novelty cannot be dismissed.

Fourth, the thesis measured motivation rather than language achievement. It therefore cannot establish whether higher motivation translated into stronger vocabulary retention, grammatical control, oral production, or longer-term proficiency growth.

Fifth, the questionnaire necessarily simplified a complex construct in order to make it accessible to very young learners. The emoji-based format was appropriate and internally consistent, but it reduces nuance and depends on children's immediate self-report.

Sixth, the qualitative corpus was rich enough for triangulation but still limited in scope. Fourteen student interviews and one teacher interview provided valuable explanatory depth, yet quieter or less verbally expressive students may still be underrepresented.

Finally, the qualitative analysis relied on structured lesson records and interview summaries rather than full verbatim transcription of every recording. This was appropriate to the age group and the case study purpose, but it means that some fine-grained interactional detail remains condensed in the written account.

5.9 Recommendations for Future Research

Future research should build directly on the limitations and unresolved findings of the present study rather than merely extending the topic in a general way.

First, because the intervention lasted only six instructional weeks, longitudinal research is needed to determine whether the motivational pattern observed here remains stable over a longer period. This recommendation follows directly from the current study's inability to separate fully between sustained need support and possible novelty effects. A semester-long or year-long design would help clarify whether Baamboozle-integrated English review continues to support participation once the format becomes routine.

Second, because the present thesis measured classroom motivation rather than language achievement, future studies should investigate whether stronger motivation during review is accompanied by measurable learning gains. This could include delayed vocabulary retention, sentence production, oral responsiveness, pronunciation confidence, or broader classroom willingness to communicate. Such work would be important because the current findings show that the review stage became more motivating, but they do not yet show how far that motivational advantage transfers into language development.

Third, because the study identified broader but still imperfect participation, future research should examine more closely which learners benefit most and which remain peripheral. This recommendation follows from the presence of no-change and declining cases, as well as the recurring qualitative concern with fairness and turn imbalance. Future studies could track quieter, hesitant, or lower-confidence learners more closely through participation mapping, repeated

interviews, stimulated recall, or video-based micro-analysis in order to determine whether gamified review genuinely broadens equitable participation or sometimes redistributes it only partially.

Fourth, because the present findings suggest that teacher mediation is central, implementation-focused research should compare different pedagogical versions of the same platform. For example, future studies could compare more competitive and more collaborative designs, different forms of team formation, different turn-allocation rules, varying amounts of think time, or different error-handling practices. This would address an unresolved issue in the current thesis: whether the observed advantage belongs mainly to Baamboozle as a platform or to the specific way the teacher structured the review stage.

Fifth, because the case was bounded to one school, one teacher, and one grade level, broader comparative sampling is needed across Vietnamese public-school settings. Studies involving multiple schools, different regions, different teachers, and other primary grade levels would help determine how far the present pattern is transferable and where it may vary. This is especially important in Vietnam, where infrastructural conditions and classroom ecologies can differ substantially across schools.

Sixth, because the qualitative corpus in the present study relied on structured summaries rather than fine-grained transcription of every lesson, future research could use more interactionally detailed methods. Video-supported classroom discourse analysis, for example, would make it possible to trace exactly how peer prompting, hesitation, teacher mediation, and turn uptake unfold moment by moment during the review stage. Such work would deepen the explanatory account of how autonomy, competence, and relatedness are actually enacted in lower-primary classrooms.

Overall, the most useful future research will be research that keeps the same disciplinary caution as the present thesis: it should remain sensitive to the bounded nature of classroom cases, but it should also continue investigating how review-

stage design shapes broad, equitable, and psychologically safe participation for young EFL learners..

5.10 Chapter Conclusion

This chapter has interpreted the findings in relation to the research questions, Self-Determination Theory, and the TESOL literature in order to explain not only whether Baamboozle mattered, but why it appeared to matter in this specific setting. The evidence indicates that Baamboozle-integrated English review was associated with stronger short-term classroom motivation than teacher-centered review, and that this pattern is best understood as a result of restructured participation conditions rather than entertainment alone.

In direct answer to RQ1, the study found that the Baamboozle class demonstrated stronger overall motivational growth than the comparison class, with advantages visible in the total score, the gain pattern, and all three SDT subdimensions. In direct answer to RQ2, students and their teacher generally experienced Baamboozle-integrated English review as a more participatory, confidence-supportive, and socially engaging review format, while also identifying conditions that qualified these advantages, especially in relation to noise, pacing, fairness, and technology.

Taken together, these answers support a coherent SDT-based interpretation. Within this bounded case, the review stage appears to have become more motivationally effective when it offered bounded choice, visible success opportunities, and peer-supported participation. In that configuration, autonomy made entry into participation more invitational, competence made answering more survivable, and relatedness made classroom involvement more socially secure. The thesis therefore suggests that the Chapter 1 problem of uneven participation in ordinary review can be understood, at least in part, as a problem of insufficiently supported psychological needs.

At the same time, the thesis has retained the limits of its own evidence. The study was a mixed-methods case study with an embedded intact-class comparison, not a randomized experiment. Its claims therefore remain bounded to one Vietnamese public primary-school context, one teacher, one grade level, and one six-week review-stage intervention. The findings are persuasive within that case, but they are not presented as universal proof or as a blanket endorsement of gamification.

Even so, the practical significance of the thesis is clear. In a Vietnamese lower-primary classroom where whole-class projection, teacher mediation, fixed class groupings, and infrastructural constraints are everyday realities, a modest digital tool was able to reorganize review in ways that appear to support broader, safer, and more equitable participation. For TESOL, that is the central value of the study. It shows that classroom motivation for very young learners may be strengthened not only through major curricular reform, but also through carefully designed changes to an ordinary and often overlooked lesson stage.

REFERENCES

- Alfiah, A., & Sholihah, M. (2025). Boosting elementary science engagement: A qualitative study of baamboozle-based game learning in Indonesian classrooms. *ASATIZA Jurnal Pendidikan*, 6(2), 132–146. <https://doi.org/10.46963/asatiza.v6i2.2726>
- Amanah, N. P., Safitri, C. A., & Indriani, L. (2025). The use of Baamboozle website to improve EYL vocabulary mastery. *Journal of English Language Teaching and Cultural Studies*, 8(1), 1–13. <https://doi.org/10.48181/jelts.v8i1.27959>
- Arini, A. R., & Suwarso, P. N. (2024). EFL students' perception towards the use of Baamboozle for vocabulary learning. *ELTIN Journal*. <https://e-journal.stkipsiliwangi.ac.id/index.php/eltin/article/view/4674>
- Aryansyah, A. D. (2024). The application of the Baamboozle game to improve students' vocabulary mastery. *Jurnal Pendidikan Multikultural Indonesia*. <https://doi.org/10.36663/jpmi.v4i1.671>
- Azzahra, F., & Kembaren, F. R. W. (2025). Gamified Learning in EFL Classrooms: Students' perceptions of Baamboozle games for Vocabulary Mastery. *IJECA (International Journal of Education and Curriculum Application)*, 8(2), 177. <https://doi.org/10.31764/ijeca.v8i2.32640>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Bui, T. L. D., Nguyen, T. T. K., & Nguyen, K. D. (2025). Teachers' Assessment of Primary School EFL Learners' Engagement: Insights from Rural and Urban Classrooms in Vietnam. *Language Teaching Research Quarterly*, 47, 138–155. <https://doi.org/10.32038/ltrq.2025.47.08>
- Butler, Y. G. (2019). Teaching vocabulary to young second- or foreign-language learners. *Language Teaching for Young Learners*, 1(1), 4–33. <https://doi.org/10.1075/ltyl.00003.but>

Cameron, L. (2001). Teaching languages to young learners. Cambridge University Press. <https://doi.org/10.1017/CBO9780511733109>

Carello, J., & Butler, L. D. (2015). Practicing what we teach: Trauma-Informed Educational Practice. *Journal of Teaching in Social Work*, 35(3), 262–278. <https://doi.org/10.1080/08841233.2015.1030059>

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. Routledge.

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches. SAGE.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. Plenum. <https://doi.org/10.1007/978-1-4899-2271-7>

Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. *MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. <https://doi.org/10.1145/2181037.2181040>

Dichev, C., & Dicheva, D. (2017). Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *International Journal of Educational Technology in Higher Education*, 14(1). <https://doi.org/10.1186/s41239-017-0042-5>

Dörnyei, Z. (2007). Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies. Oxford University Press.

Dörnyei, Z., & Ushioda, E. (2009). *Motivation, language identity and the L2 self*. Multilingual Matters.

Field, A. (2018). Discovering statistics using IBM SPSS statistics. SAGE.

Gardner, R. C., Lalonde, R. N., & Moorcroft, R. (1985). The role of attitudes and motivation in second language learning: Correlational and experimental considerations. *Language Learning*, 35(2), 207–227. <https://doi.org/10.1111/j.1467-1770.1985.tb01025.x>

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.

Lamb, M. (2017). The motivational dimension of language teaching. *Language Teaching*, 50(3), 301–346. <https://doi.org/10.1017/s0261444817000088>

Lou, N. M., & Noels, K. A. (2017). Measuring language mindsets and modeling their relations with goal orientations and emotional and behavioral responses in failure situations. *Modern Language Journal*, 101(1), 214–243. <https://doi.org/10.1111/modl.12380>

Mellor, D., & Moore, K. A. (2014). The use of likert scales with children. *Journal of Pediatric Psychology*, 39(3), 369–379. <https://doi.org/10.1093/jpepsy/jst079>

Nguyen, N. M. (2021, November 23). *Using stories in presenting English grammar to Vietnamese young learners*. <https://ijte.org/index.php/journal/article/view/120>

Nhung, L. T. T., & Dieu, T. D. V. (2020). Impacts of using cartoons as an instructional tool on English communicative skills of young learners at a Vietnamese primary school. *HO CHI MINH CITY OPEN UNIVERSITY JOURNAL OF SCIENCE - SOCIAL SCIENCES*, 10(2), 78–86. <https://doi.org/10.46223/hcmcoujs.soci.en.10.2.553.2020>

Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-

determination theory. *Language Learning*, 50(1), 57–85.
<https://doi.org/10.1111/0023-8333.00111>

Nurlia, R., Wiratno, & Mulhuda, M. A. (2025). Improving English Vocabulary Using Gamification Of Baamboozle Platform For Smpn 2 Nglegok Grade 8. *Journal of Development Research*, 9(2).
<https://doi.org/10.28926/jdr.v9i2.479>

Oga-Baldwin, W. L. Q., & Nakata, Y. (2015). Structure also supports autonomy: Measuring and defining autonomy-supportive teaching in Japanese elementary foreign language classes. *Japanese Psychological Research*, 57(3), 167–179. <https://doi.org/10.1111/jpr.12077>

Oga-Baldwin, W. Q., Nakata, Y., Parker, P., & Ryan, R. M. (2017). Motivating young language learners: A longitudinal model of self-determined motivation in elementary school foreign language classes. *Contemporary Educational Psychology*, 49, 140–150.
<https://doi.org/10.1016/j.cedpsych.2017.01.010>

Phan, N. T. P., Truong, N. T. N., & Nguyen, N. T. (2024). Exploring English pronunciation teaching practices at Vietnamese primary schools. *HO CHI MINH CITY OPEN UNIVERSITY JOURNAL OF SCIENCE - SOCIAL SCIENCES*, 15(4), 21–38. <https://doi.org/10.46223/hcmcoujs.soci.en.15.4.3220.2025>

Phuong, T. T. H. (2025). Integration of ICT into EFL classrooms: Have Vietnamese primary schools progressed post-Covid-19? In *Issues in Educational Research*. <https://www.iier.org.au/iier35/phuong.pdf>

Pinter, A. (2017). *Teaching Young Language Learners, second edition*. Oxford University Press.

Prensky, M. (2001). Digital natives, digital immigrants, part 1. On the Horizon, 9(5), 1–6. <https://doi.org/10.1108/10748120110424816>

Rahayu, I. R., & Rukmana, D. (2022). THE EFFECT OF GAME-BASED LEARNING MODEL ASSISTED BY Baamboozle on the multiplication skills of

elementary school students. *Primary Jurnal Pendidikan Guru Sekolah Dasar*, 11(3), 806. <https://doi.org/10.33578/jpfkip.v11i3.8964>

Rahmawati, D. F., & Adityarini, H. (2026). The implementation of Baamboozle website for teaching English vocabulary to ‘Aisiyah Singopuran Kindergarten. *jurnal.unai.edu*. <https://doi.org/10.35974/re9a1021>

Ratinho, E., & Martins, C. (2023). The role of gamified learning strategies in student’s motivation in high school and higher education: A systematic review. *Heliyon*, 9(8), e19033. <https://doi.org/10.1016/j.heliyon.2023.e19033>

Reeve, J. (2015). Autonomy-Supportive Teaching: What It Is, How to Do It. In *Building Autonomous Learners* (pp. 129–152). https://doi.org/10.1007/978-981-287-630-0_7

Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL*, 27(1), 38–57. <https://doi.org/10.1017/s0958344014000226>

Ryan, R. M., & Deci, E. L. (2018). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press.

Safrizal, S., Selpia, R., Arafah, K., & Sulaimon, J. T. (2026). Game-based learning intervention: the effect of baamboozle on science learning interest in primary education. *Primary Jurnal Pendidikan Guru Sekolah Dasar*, 15(1), 15–26. <https://doi.org/10.33578/jpfkip.v15i1.15-26>

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, 69, 371–380. <https://doi.org/10.1016/j.chb.2016.12.033>

Sakdiyah, I. H., Maolida, E. H., & Nurviyani, V. (2024). Utilizing Game-Based Learning Baamboozle in developing students’ English Grammar mastery. *JALL (Journal of Applied Linguistics and Literacy)*, 8(1), 132. <https://doi.org/10.25157/jall.v8i1.12755>

Saud, S., Aeni, N., & Azizah, L. (2022). Leveraging quizzes and Baamboozles to engage EFL students in online classes. *International Journal of Language Education*, 6(2), 169. <https://doi.org/10.26858/ijole.v6i2.24301>

Tham, D. M., & Thuy, P. T. T. (2023). The application of digital storytelling in teaching English listening skills to young learners: Teachers' perceptions of benefits and constraints. *VNU Journal of Foreign Studies*, 39(2), 164–174. <https://doi.org/10.63023/2525-2445/jfs.ulis.5060>

Thi, D. T., & Thinh, L. (2023). Extramural English and Vietnamese EFL primary school students' listening and reading proficiency in Vietnam. *HO CHI MINH CITY OPEN UNIVERSITY JOURNAL OF SCIENCE - SOCIAL SCIENCES*, 13(2), 65–81. <https://doi.org/10.46223/hcmcoujs.soci.en.13.2.2805.2023>

Vasya, I. D. a. M. S., Puspitaloka, N., & Kartini, D. (2025). Utilizing Baamboozle as Media in teaching Vocabulary Mastery. *jes.ejournal.unri.ac.id*. <https://doi.org/10.31258/jes.9.4.p.2483-2495>

Walidaina, F. a. J., Fadhilawati, D., & Sutanti, N. (2024). The effectiveness of Baamboozle to escalate Seventh-Grade students' vocabulary learning outcomes. *Jurnal Pendidikan Riset Dan Konseptual*, 8(4), 786. https://doi.org/10.28926/riset_konseptual.v8i4.1109

Wang, A. I. (2015). The wear out effect of a game-based student response system. *Computers & Education*, 82, 217–227. <https://doi.org/10.1016/j.compedu.2014.11.004>

Weinberg, S. L., & Abramowitz, S. K. (2008). *Statistics using SPSS: An Integrative Approach*. Cambridge University Press.

Winaningsih, E. T., Syarif, F., & Pahrurroji, P. (2022). Baamboozle's effectiveness in increasing learning motivation in Islamic Studies. *Tarbawiyah Jurnal Ilmiah Pendidikan*, 6(2). <https://doi.org/10.32332/tarbawiyah.v6i2.5385>

Yip, F. W. M., & Kwan, A. C. M. (2006). Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational Media International*, 43(3), 233–249. <https://doi.org/10.1080/09523980600641445>

Zulbani, A., & Syafryadin. (2025). Examining the efficacy of the Baamboozle application in improving students' vocabulary proficiency. *Journal of English for Specific Purposes in Indonesia*, 4(1), 59–66. <https://doi.org/10.33369/espindonesia.v4i1.41055>

APPENDICES

APPENDIX A. MOTIVATION QUESTIONNAIRE

Title: English Class Feelings Questionnaire for Grade 2 Students

Instructions for teacher/researcher: Read each sentence aloud slowly. Ask students to circle one face.

Response options

- 1 = 😞 Not true for me
2 = 😐 Sometimes true for me
3 = 😊 True for me

Autonomy Items

1. *I want to join the English activity.*
2. *I like having my turn in English class.*
3. *I feel involved in what we do in English.*
4. *I want to answer when teacher asks the class.*

Competence Items

5. *I feel I can do English class work.*
6. *I feel happy when I answer correctly.*
7. *I want to try again when I make a mistake.*
8. *I understand many things in English class.*

Relatedness Items

9. *I like playing or working with my classmates in English.*
10. *My friends help me in English class.*
11. *I feel good when my team or class does well.*
12. *I feel I belong in English class.*

Scoring: Add all 12 items. Minimum = 12. Maximum = 36.

APPENDIX B. CLASSROOM OBSERVATION CHECKLIST

Observation Categories

- Number of student volunteer responses
- Number of peer-help utterances
- Number of visible positive-affect episodes
- Number of off-task incidents
- Number of refusal/hesitation episodes
- Estimated on-task participation rate
- Field notes on pacing, noise, fairness, and teacher support

Observation Template

Date:

Class:

Lesson topic:

Condition: Baamboozle / Comparison

Observer:

Appendix Table B1. Observation recording template

Indicator	Count / Estimate	Notes
Volunteer responses		
Peer-help utterances		
Positive-affect episodes		
Off-task incidents		
Refusal/hesitation episodes		
On-task participation		
Notable classroom moments		

APPENDIX C. STUDENT AND TEACHER INTERVIEW PROTOCOLS

C1. Student Interview Protocol

- 1. What part of English class do you like most? Why?*
- 2. How do you feel when it is your turn to answer?*
- 3. Do you feel different in the game lesson and the normal review lesson?
How?*
- 4. What do your classmates do when your team is answering?*
- 5. What happens when you do not know the answer?*
- 6. What makes you want to join more in English class?*
- 7. Is there anything you do not like about the game lesson?*
- 8. Which class makes you feel more brave to answer? Why?*

C2. Teacher Interview Protocol

- 1. How would you describe student participation in the Baamboozle and comparison groups?*
- 2. What differences did you notice in volunteerism and attentiveness?*
- 3. Did Baamboozle seem to affect lower-confidence students?*
- 4. How did the intervention affect classroom management?*
- 5. What benefits did you observe for peer interaction?*
- 6. What challenges emerged during implementation?*
- 7. Would you continue using Baamboozle? Under what conditions?*
- 8. How do you interpret the role of such a platform in primary English teaching?*

APPENDIX D. PARENTAL CONSENT FORM

Title: Parental Consent Form for Participation in English Classroom Research

Dear Parent or Guardian,

Your child is invited to participate in a classroom-based study about English learning motivation. The study examines different ways of reviewing vocabulary and grammar in Grade 2 English lessons. If you agree, your child may complete a short emoji-based questionnaire and may be observed during ordinary class activities. A small number of students may also be invited to participate in a short interview about their classroom experience.

Participation is voluntary. Your child may stop participating at any time without any penalty. The study is intended for educational research only. Your child's name will not appear in any report. All information will be kept confidential.

There are no serious risks expected from participation. The lessons will remain age-appropriate and aligned with the school curriculum. The potential benefit is that the study may help teachers understand how to make English lessons more motivating for young learners.

Please sign below if you agree that your child may participate.

Child's name: _____

Parent/Guardian name: _____

Signature: _____

Date: _____

APPENDIX E. STUDENT ASSENT FORM

Title: Student Assent Script

Hello. We are doing a project about English class. We want to know how children feel in English lessons. You may answer some easy questions with faces and maybe talk for a few minutes about your class. This is not a test. There are no right or wrong answers. You can say yes or no. If you start and later want to stop, that is okay. Your teacher will not be angry, and your grade will not change.

Do you want to join?

Yes / No

Student name: _____

Date: _____

APPENDIX F. BAAMBOOZLE INTERVENTION LESSON PLANS

Lesson Plan 1

Unit Topic: Colors

Language Focus: red, blue, yellow, green; “What color is it?” / “It is ____.”

Objective: Students identify and say target colors confidently.

Warm-up: Flashcard naming race.

Presentation: Teacher introduces color cards and models sentence pattern.

Guided Practice: Choral repetition and pair pointing activity.

Baamboozle Review: Teams choose numbered tiles. Each tile shows a color object or prompt. Students answer orally using the target structure.

Teacher Role: Encourage full-sentence answers, ensure even turns, slow down dominant students.

Closure: Quick recap with four class questions.

Lesson Plan 2

Unit Topic: Animals

Language Focus: cat, dog, bird, fish; “What is this?” / “It is a ____.”

Objective: Students recognize and name common animals.

Warm-up: Animal sound guessing.

Presentation: Teacher shows pictures and models pronunciation.

Guided Practice: Matching pictures and words on the board.

Baamboozle Review: Animal picture tiles. Bonus tile asks one student to act like an animal while teammates answer.

Teacher Role: Keep excitement controlled, reinforce pronunciation supportively.

Closure: Exit chant with animal names.

Lesson Plan 3

Unit Topic: Numbers

Language Focus: 1–10; “How many?”

Objective: Students count classroom objects and respond using number words.

Warm-up: Count-and-clap activity.

Presentation: Number cards and board sequencing.

Guided Practice: Counting objects in pairs.

Baamboozle Review: Tiles show grouped objects; teams answer “There are ____.”

Teacher Role: Give brief think time before answers; invite quieter students first.

Closure: Whole-class count backward from 10.

Lesson Plan 4

Unit Topic: Classroom Objects

Language Focus: book, pen, ruler, bag; “Is this a ____?” / “Yes, it is.”

Objective: Students identify classroom objects and answer short questions.

Warm-up: Hidden-object bag game.

Presentation: Real object demonstration.

Guided Practice: Teacher asks individual recognition questions.

Baamboozle Review: Object image tiles and yes/no challenge tiles.

Teacher Role: Protect weaker learners from public embarrassment by allowing team whisper support.

Closure: Students point to an object and say its name.

Lesson Plan 5

Unit Topic: Family

Language Focus: mother, father, brother, sister; “This is my ____.”

Objective: Students recognize and say family vocabulary.

Warm-up: Family picture show-and-tell.

Presentation: Teacher models with picture cards.

Guided Practice: Students repeat and identify family roles.

Baamboozle Review: Family tree picture tiles with naming prompts.

Teacher Role: Connect correct answers to praise and encouragement, not only points.

Closure: Students say one family word before leaving.

Lesson Plan 6

Unit Topic: Food

Language Focus: rice, milk, bread, banana; “I like ____.”

Objective: Students express simple preferences using target food words.

Warm-up: Food flashcard reveal.

Presentation: Vocabulary modeling and preference sentence frame.

Guided Practice: Students repeat and choose favorite foods.

Baamboozle Review: Food image tiles, surprise double-point tile, team sentence completion.

Teacher Role: Ensure that all teams receive balanced opportunity.

Closure: Whole class says “I like ____” with one chosen food.

APPENDIX G. RAW QUANTITATIVE DATA

Appendix Table G1. Baamboozle Group

ID	Pre-test	Post-test	Gain
E01	25	27	2
E02	23	28	5
E03	25	25	0
E04	28	25	-3
E05	22	31	9
E06	22	33	11
E07	28	28	0
E08	26	32	6
E09	21	30	9
E10	25	27	2
E11	21	30	9
E12	21	33	12
E13	24	28	4
E14	17	33	16
E15	17	21	4
E16	21	31	10
E17	20	29	9
E18	24	28	4
E19	20	29	9
E20	18	23	5
E21	28	28	0

E22	22	30	8
E23	23	33	10
E24	18	27	9
E25	21	26	5
E26	23	27	4
E27	19	31	12
E28	24	30	6
E29	21	27	6
E30	22	30	8

Baamboozle Group Mean: Pre = 22.30, Post = 28.67

Appendix Table G2. Comparison Group

ID	Pre-test	Post-test	Gain
C01	21	26	5
C02	29	28	-1
C03	23	23	0
C04	19	24	5
C05	26	24	-2
C06	19	21	2
C07	24	26	2
C08	16	26	10
C09	18	25	7
C10	23	25	2
C11	25	21	-4
C12	23	24	1

C13	22	24	2
C14	22	23	1
C15	18	25	7
C16	20	27	7
C17	21	31	10
C18	26	26	0
C19	24	26	2
C20	17	25	8
C21	24	19	-5
C22	21	25	4
C23	20	26	6
C24	25	33	8
C25	26	25	-1
C26	26	26	0
C27	20	25	5
C28	22	22	0
C29	24	29	5
C30	26	28	2

Comparison Group Mean: Pre = 22.33, Post = 25.27

APPENDIX H. QUALITATIVE CODING MATRIX

Table H1. Sample Qualitative Coding Matrix

Excerpt	Initial Code	Theme
“I look at the screen all the time because I want to pick a number.” (E1)	anticipatory attention	Joyful engagement and sustained attention
“If we are wrong, teacher tells us and the next one comes, so I still want to try again.” (E4)	recovery after error	Perceived competence
“My friends help me remember the word.” (E1)	peer support	Relatedness
“Excitement was good for participation, but sometimes it rose too fast.” (T1)	over- arousal management	Motivation with conditions
“I like quiet work sometimes.” (C2)	value of calm pacing	Motivation with conditions
“In normal review, some children answer and others watch. In the game, even when only one child speaks, several are involved.” (T1)	distributed participation	Relatedness

APPENDIX I. INTERVIEW REPORT SUMMARIES

Baamboozle Group

E1: Highly engaged with tile selection; emphasized peer support and excitement.

E2: Reported feeling more capable when answers were confirmed quickly.

E3: Focused on anticipation and not wanting to miss the next turn.

E4: Noted that mistakes felt less frightening in team format.

E5: Valued game rhythm and team encouragement.

E6: Reported more confidence when classmates whispered help.

E7: Expressed strong preference for game-based review over worksheet review.

E8: Previously shy; reported feeling less afraid to answer in a team.

Comparison Group

C1: Considered ordinary review acceptable but less exciting.

C2: Preferred calm explanation at times; valued teacher praise.

C3: Participated when called on but rarely volunteered spontaneously.

C4: Linked confidence to already knowing the answer.

C5: Found worksheets manageable but not interesting.

C6: Appreciated classmates but described less active group involvement.

Teacher

T1: Observed increased volunteerism, stronger visible engagement, and more peer support in the Baamboozle group, but also greater noise and higher preparation demands.