UCLouvain
Faculté de philosophie, arts et lettres – FIAL

Faculté de philosophie, arts et lettres

Fostering Agentic Engagement Through Game-Based Language Learning: A Mixed-Methods Study

Case Study of *The Travel Kit*

Author : Clara Dauby
Supervisor : Prof. Ferran Suner Munoz
Co-supervisor: Bailing He
Academic year: 2025-2026
GERM2MD – Master en langues et lettres modernes, orientation germaniques, à finalité didactique

Declaration of honour

Dissertation/Individual Final Project

Faculty of Philosophy, Arts and Letters, UCLouvain

I hereby declare that this is an original and personal work, that all referenced sources have been fully cited, regardless of their origin, and that the use of artificial intelligence tools complies with the general instructions published by UCLouvain and available at the following address :

https://cdn.uclouvain.be/groups/cms-editors-p1/portail/reglement/Consignes-ChatGPT-version-etudiante.pdf?itok=SebXXm87 (French version)

I am aware that failure to cite a source, to cite it clearly and fully, to announce in a specific section at the beginning of the dissertation the use I have made of artificial intelligence tools, or to use them in a manner contrary to the guidelines set forth in the institutional note, constitutes serious irregularities in accordance with the ***Academic regulations and procedures*** (Chapter 4, section 7, articles 107 to 114) within the University. In particular, I am aware of the risks of academic and disciplinary sanctions.

In light of the above, I declare on my honour that I have not committed plagiarism or any other form of fraud and that I have not used artificial intelligence tools beyond what is permitted by UCLouvain.

Last name, First name : Dauby Clara

Date : 16/08/2026

Student's signature :

Acknowledgements

I am deeply grateful to Professor Suner for his continued support and guidance throughout my academic career at the Université Catholique de Louvain, from my Bachelor’s until the end of my Masters’ degree, and for his encouragement to pursue academic research with confidence and curiosity by joining a research team.

I would also like to extend my deepest and most sincere gratitude to Bailing, my co-supervisor and research team leader, who took me under her wing and mentored me through the research process with extraordinary patience, encouragement, and care.

Special thanks are due to Professor Auracher at the University of Singapore for kindly permitting our team to recruit his students for this experiment and for offering invaluable support from an unofficial capacity in the coordination of the main study in Singapore. My sincere thanks also go to the IT staff for their practical support throughout the experiment.

Likewise, I am grateful to Christina from the International School of Le Verseau, for permitting us to recruit her students, for her assistance in coordinating the game session and for accommodating us as we navigated the technical difficulties that can arise in an experiment such as this one. I am also thankful to the school secretary, Françoise, for her patience with our many comings and goings, and for facilitating room access.

I would like to warmly thank my friend Ines Lhoest for acting as my research assistant during a substantial part of this study, and for her dedicated help and support.

Lastly, my sincere appreciation goes to the examiners for the time and care they will devote to reading this thesis.

AI Use Statement

Generative artificial intelligence (AI) tools were used in a limited capacity during the preparation of this thesis. Specifically, ChatGPT was used to refine linguistic expressions such as grammar and style as well as vocabulary refinement. It was also used to refine the thesis structure, to provide guidance on the use of SPSS during the analytical procedures and on the qualitative analysis procedures of this thesis. However, all academic content, study design, data analysis, and interpretation of results were developed, verified, and written by the author. The author also reviewed and validated all AI-assisted suggestions and takes full responsibility for the content of this thesis.

Table of Contents

1. Introduction

1.1. Background and Context

In contemporary education, student engagement has become one of the main constructs used to explain why some learners seem to learn more, or struggle less, than others, both in terms of learning processes and outcomes (Buntins et al., 2021; Fredricks et al., 2004; Kahu, 2013; Li & Xue, 2023). Research across educational psychology and higher education increasingly shows that engaged students tend to learn more deeply and report more positive experiences in the learning process than those who stay passive (Fredricks et al., 2004; Kahu, 2013; Xu et al., 2023). Due to these findings, engagement is now frequently treated as a central construct for understanding how instructional practices and structures as well as learning environments shape students' participation and long-term academic trajectories (Astin, 1999, originally 1984; Wong & Liem, 2022).

Over time, the construct of engagement has evolved from a relatively narrow perspective that mainly associated it with attendance, compliance, and time-on-task, toward more multidimensional perspectives that view students as active participants in their learning process and environment (Astin, 1999; Fredricks et al., 2004; Reeve & Tseng, 2011). Early models studied behavioral involvement above all, like turning up to class, completing assignments, participating in activities, whereas more recent frameworks recognize emotional and cognitive factors such as enjoyment, interest, or self-regulation (Fredricks et al., 2004; Korhonen et al., 2019; Reeve & Tseng, 2011). These developments have culminated in the idea of engagement as a meta-construct that integrates behavioral, emotional, cognitive, and, more recently, agentic dimensions (Mameli & Passini, 2017; Reeve & Tseng, 2011; Wong & Liem, 2022).

The field of second language acquisition (SLA) provides a particularly rich context in which engagement matters. Communicative classrooms depend on learners' willingness to speak, take risks, to negotiate meaning, and to persist despite difficulties, which ties engagement directly to progress and success in the target language (Shernoff et al., 2003; Sikandar et al., 2025). Decades of research on motivation, self-efficacy, and cognitive load in L2 learning have shown that these constructs strongly influence whether learners will invest effort and take initiative, but engagement provides a complementary approach as to what learners actually do, feel, and think in classroom and task settings (Bandura, 1994, 2006; Martin, 2007; Sweller, 1994; van Merriënboer & Sweller, 2005).

At the same time, the rapid evolution of educational technology has reshaped the environments in which engagement is studied: classrooms. Technology-enhanced learning, ranging from mobile apps to game-based environments and virtual reality, can open new possibilities for collaboration, feedback, personalization, and authentic task design, which may positively or negatively influence engagement (Bond & Bedenlier, 2019; Nkomo et al., 2021). Digital game-based language learning (DGBLL) and immersive augmented reality (AR) or virtual reality (VR) are particularly interesting because they combine interaction, narrative, and simulation to create experiences that align with many of the conditions thought to support engagement and motivation, such as clear goals, meaningful choices, and social interaction (Gee, 2003, 2005; Reinhardt, 2019; Hung et al., 2018; Thi, 2025).

Within this context, the present thesis focuses on agentic engagement, defined as students' proactive, constructive contribution to the flow of instruction, in a VR-based, game-based L2 German environment, and more specifically, how it evolves across repeated VR game-based scenarios. Rather than treating learners as passive recipients of teaching, the concept of agentic engagement emphasizes their initiatives to ask questions, express preferences, suggest alternative approaches, and negotiate task procedures, thereby actively shaping their learning conditions (Patall, 2024; Reeve & Tseng, 2011; Reeve, 2013). Investigating how such agentic contributions develop in immersive VR GBLL environments offers a way to connect contemporary engagement theory with emerging technologies with practical concerns in L2 German speaking instruction.

1.2. Problem Statement

Although engagement is now widely recognized as a multidimensional and dynamic construct, the research field still shows significant conceptual and methodological inconsistencies. Reviews of engagement models have repeatedly pointed out the lack of conceptual consensus, noting that the construct is used in divergent ways, sometimes treated as a broad umbrella term that encompasses motivation, participation, and outcomes, and sometimes defined more narrowly in relation to specific behaviors (Buntins et al., 2021; Nkomo et al., 2021; Wong & Liem, 2022). Issues like over-generalization, jingle-jangle fallacies, object ambiguity, and under-theorization continue to complicate attempts to compare findings across studies and to build cumulative knowledge (Kahu, 2013; Wong & Liem, 2022).

In technology-enhanced and game-based contexts, these issues are even more complex. Studies reviewing student engagement and educational technology in higher education indicate that many studies either rely on narrow operationalizations focused on a single dimension, often behavioral or affective/emotional, or employ instruments that are not clearly linked to multidimensional frameworks (Bond et al., 2020; Buntins et al., 2021). In digital game-based environments, engagement is frequently deduced from usage metrics, enjoyment ratings, or motivational self-reports, but agentic forms of participation and the distinction between engagement and its motivational antecedents is still understudied (Hung et al., 2018; Ryan & Deci, 2020; Xu et al., 2018).

Virtual reality creates even further complexity. Meta-analyses and systematic reviews suggest that immersive VR can increase presence and various engagement dimensions, but they also show that heightened immersion can be accompanied by increased cognitive load caused by interface demands and technical issues, which can undermine agentic behaviors and learning (Han et al, 2021; Makransky et al., 2019; Makransky & Petersen, 2021; Poupard et al., 2025). Learners must divide their attention between navigating the environment, manipulating the controls, processing instructional content, and using their skills, which creates conditions where cognitive load may fluctuate in ways that are difficult to predict and that may affect how, and to what extent, engagement can be expressed (Sweller, 1994; Andersen & Makransky, 2021).

Within L2 education, immersive AR/VR and GBLL studies have documented positive effects on motivation, enjoyment, and sometimes engagement, but relatively few have examined agentic engagement as a distinct but related dimension. A lot of the existing research either does not include agentic engagement explicitly in its measurement model, or treats learners' proactive initiatives as unrelated observations rather than a systematically measured behavior (Buntins et al., 2021; Hung et al., 2018; Huang et al., 2021). Moreover, Humanities and language learning contexts are still underrepresented in VR engagement research compared to STEM disciplines, leaving open questions about how agentic engagement manifests in communicative, language-focused scenarios (Bond et al., 2020; Nkomo et al., 2021).

Finally, while constructs such as motivation, self-efficacy, and cognitive load have each been studied extensively, there is only limited empirical work that examines their interaction together with agentic engagement in VR GBLL contexts. Martin's distinction between motivation and engagement, Bandura's work on self-efficacy, SDT research on psychological needs, and cognitive load theory all point to rich theoretical connections, but empirical studies that bring them together in immersive L2 settings remain rare (Bandura, 2006; Deci & Ryan, 2012; Makransky & Petersen, 2021; Martin, 2007; van Merriënboer & Sweller, 2005).

1.3. Literature-Based Research Gap

Theoretical work on engagement now converges on a multidimensional view that recognizes behavioral, affective/emotional, and cognitive dimensions as interconnected, and introduces agentic engagement as a fourth, distinct dimension. Reeve and Tseng's (2011) foundational study validated a measure of agentic engagement, demonstrated its positive relations with classroom motivation and academic achievement, and showed that agentic engagement forms a separate factor that cannot be reduced to a part of the other dimensions (Reeve & Tseng, 2011). Subsequent research has refined the Agentic Engagement Scale and confirmed its structural validity in school settings, as well as its positive associations with autonomy support, basic need satisfaction, and motivational outcomes (Mameli & Passini, 2017, 2019; Reeve & Jang, 2022; Patall, 2024).

Self-Determination Theory provides a complementary motivational perspective. SDT emphasizes that learners' basic psychological needs for autonomy, competence, and relatedness must be supported in order to foster intrinsic motivation and high-quality engagement (Deci & Ryan, 1985, 2000, 2012; Ryan & Deci, 2017, 2020). In educational contexts, autonomy-supportive teaching and structured environments that provide clarity and choice have been shown to increase engagement, including agentic engagement, by inviting students to actively participate in shaping their learning (Jang et al., 2010; Reeve & Halusic, 2009; Reeve et al., 2004). In game-based settings, SDT-informed studies show how meaningful choices, optimally challenging tasks, and social interaction contribute to the motivational pull of games (Przybylski et al., 2010; Rigby & Ryan, 2011; Ryan et al., 2006).

GBLL research has demonstrated many benefits, including increased motivation, enjoyment, and willingness to communicate. However, it also points to limitations such as overreliance on extrinsic rewards and insufficient alignment between game mechanics and pedagogical goals (Hung et al., 2018; Smiderle et al., 2020; Thi, 2025; Xu et al., 2020). Reviews argue that language-learning games are most effective when they include speaking practice with meaningful, problem-based scenarios that require collaboration and strategic decision-making. These conditions are also considered to foster agentic engagement. Yet few studies explicitly measure or theorize students' proactive contributions to the flow of instruction within GBLL environments (Reinhardt, 2019; Reeve, 2013).

VR-based education research shows that immersive environments can promote presence and engagement, but also shows that these benefits are moderated by cognitive load and design choices (Chen et al., 2024; Han et al., 2021; Makransky et al., 2019; Makransky & Petersen, 2021). Some studies find that well-designed VR experiences support deeper cognitive engagement without overloading learners, while others report that poorly managed immersion leads to high cognitive load and weaker learning outcomes. In language-learning applications, AR/VR interventions show potential for enhancing communicative skills and learner motivation, but systematic attention to engagement dimensions, especially agentic engagement, remains limited (Huang et al., 2021; Wang et al., 2019).

Finally, measurement research suggests that mixed-methods, multimodal approaches are best to measure engagement in technology-enhanced environments. Four-dimensional scales including agentic engagement and instruments like the Classroom Engagement Inventory provide peer-reviewed and validated tools, although reviews emphasize that no single instrument can fully grasp the dynamic and context-sensitive nature of agentic engagement, especially in VR (Buntins et al., 2021; Fredricks et al., 2004, 2016; Kassab et al., 2023; Li & Lerner, 2013; Mameli & Passini, 2017; Wang et al., 2014). Recent work therefore recommends integrating self-report data with observational indicators and other tools to better understand how engagement unfolds during the learning process (Kassab et al., 2023; Wang et al., 2025).

Taken together, this literature points to a clear gap: while agentic engagement is theoretically and empirically well-established, its development over time and associations in repeated-scenarios of VR-based GBLL L2 contexts have not been systematically investigated, particularly using mixed-methods designs that incorporate intrinsic motivation, self-efficacy, cognitive load, and overall classroom engagement.

1.4. Purpose of the Study

The purpose of the present study is to examine how agentic engagement develops over repeated scenarios in a VR-based, game-based L2 German speaking intervention, and how it relates to other key cognitive and affective constructs. In a series of scenarios implemented in *The Travel Kit* VR game, pairs of learners collaborate in everyday communicative tasks designed for A2-level German, such as planning a trip, searching for a job, visiting a city, ordering food, and looking for lost belongings. These scenarios are structured to create opportunities for learners to express preferences, negotiate decisions, and seek feedback, thereby providing a natural context in which agentic engagement can emerge.

The study adopts a concurrent mixed-methods, within-subject longitudinal design. Quantitatively, it collects scenario-level data on agentic engagement, general classroom engagement (behavioral, emotional, cognitive), intrinsic motivation, self-efficacy, and cognitive load after each VR session, allowing changes and relations among constructs to be examined within learners over time (Andersen & Makransky, 2021; Leppink et al., 2013; Makransky & Petersen, 2021; Mameli & Passini, 2017; Reeve & Tseng, 2011). Qualitatively, it uses post-scenario interviews to understand how learners interpret their experiences and describe their initiatives. By integrating these strands at the discussion stage, the study aims to offer a nuanced account of agentic engagement in VR GBLL L2 German and to situate it within a network of motivational and cognitive factors.

1.5. Research Questions

Guided by multidimensional engagement theory, SDT, social cognitive theory, CDST, flow theory, and cognitive load theory, the study addresses the following research questions:

1. *How does agentic engagement develop over time in a VR game-based language learning intervention?* This question focuses on the trajectory of agentic engagement across multiple VR scenarios, treating it as a dynamic, context-sensitive construct rather than a fixed learner trait (Larsen-Freeman, 2019; Reeve & Tseng, 2011).
    a. *What patterns of agentic engagement emerge from the quantitative data ?*
    b. *What obstacles and facilitators can be identified to help explain the development of agentic engagement over time ?*
2. *To what extent does overall engagement (i.e., behavioural, emotional, and cognitive engagement), self-efficacy, intrinsic motivation, and cognitive load shape patterns of learners' agentic engagement in VR game-based language learning environments over time?* Here, agentic engagement is examined in relation to its theoretically relevant correlates and potential antecedents, drawing on SDT, social cognitive theory, and cognitive load theory to interpret these relationships (Bandura, 2006; Deci & Ryan, 2012; Makransky & Petersen, 2021; Mameli & Passini, 2017; Martin, 2007; van Merriënboer & Sweller, 2005).

1.6. Significance of the Study

The study contributes to the theoretical understanding of engagement by examining agentic engagement within an immersive VR GBLL environment for L2 German. By positioning agentic engagement among intrinsic motivation, self-efficacy, cognitive load, and general classroom engagement, and by grounding this analysis in SDT, social cognitive theory, CDST, and flow theory, the thesis contributes to ongoing efforts to refine engagement theory and to clarify its relation to motivational and cognitive constructs (Bandura, 2006; Csikszentmihalyi, 1990; Deci & Ryan, 1985, 2012; Fredricks et al., 2004; Larsen-Freeman, 2019; Thi, 2025).

Methodologically, the study adds to the literature on measuring engagement in technology-enhanced contexts by demonstrating how validated instruments for agentic engagement can be combined with motivational, self-efficacy, and cognitive load scales in a longitudinal, scenario-based design (Andersen & Makransky, 2021; Fredricks et al., 2004; Leppink et al., 2013; Mameli & Passini, 2017). The integration of quantitative results with qualitative accounts of learners' experiences exemplifies the kind of mixed-methods, multimodal approach that recent reviews recommend for a more comprehensive account of engagement's dynamic nature in VR contexts (Kassab et al., 2023; Wang et al., 2025).

Practically, the study aims to offer insights for teachers and researchers interested in VR-based, game-based language learning environments. By identifying when and how learners take agentic steps in VR scenarios, and how these are linked to motivational and cognitive conditions, the findings can inform the design of tasks, scaffolding strategies, and interface features that support autonomy, competence, and relatedness while managing cognitive load (Chen et al., 2024; Makransky & Petersen, 2021; Rigby & Ryan, 2011;Ryan & Deci, 2017, 2020). More generally, the thesis contributes to a still relatively small and new body of work that examines student engagement in Humanities disciplines and language-learning contexts within immersive, technology-enhanced environments (Bond & Bedenlier, 2019; Bond et al., 2020; Huang et al., 2021).

1.7. Overview of the Thesis Structure

The remainder of the thesis is structured as follows. Chapter two reviews the literature on student engagement, tracing its evolution from participation-based constructs to multidimensional frameworks, introducing agentic engagement, before examining measures of engagement and research on technology-enhanced, game-based, and VR-based engagement, and finally introducing the constructs that will be examined in relation to agentic engagement.

Chapter three presents the theoretical framework, drawing primarily on Self-Determination Theory and integrating CDST, social cognitive theory, and flow theory. Chapter four states the research questions. Chapter five describes the study design and the methodology, including the mixed-methods approach, research context and development phases, participant selection, materials and scenario description, experimental procedures, data collection instruments, and data analysis procedures. Chapter six reports the quantitative results. Chapter seven presents the qualitative findings, and chapter eight offers an integrated discussion that brings together quantitative and qualitative results in light of the theoretical framework and research questions. Finally, chapter nine concludes the thesis by summarizing the main findings, discussing limitations, and outlining directions for future research on agentic engagement in VR-based game-based language learning.

2. Literature Review

This chapter reviews the literature that informs the present study's investigation of agentic engagement in VR-based L2 German learning environments as well as the constructs that interact with it such as overall engagement, motivation, self-efficacy, and cognitive load. It begins by examining the evolution of student engagement from early participation-based views to multidimensional frameworks and the more recent addition of agentic engagement as a fourth dimension of engagement. It briefly defines and describes the constructs examined in relation to agentic engagement within the study, before presenting research on game-based language learning and virtual reality in educational contexts, with a particular focus on agentic engagement. Finally, it synthesizes the literature to identify the gap that the current study addresses.

2.1. Student Engagement in L2 Learning

2.1.1. Evolution from Participation to Multidimensional Construct

The conceptualization of student engagement has undergone significant transformation from early notions of participation and time-on-task towards more complex, multidimensional frameworks that view students as active co-constructors of their learning experience (Kahu, 2013). Early research tended to equate engagement with observable indicators such as attendance, on-task behavior, involvement in school activities, and completion of assignments, positioning it primarily as a behavioral response to institutional structures and teaching practices (Astin, 1999).

In higher education, Astin's theory of student involvement framed engagement as the quantity and quality of physical and psychological energy students invest in academic and social aspects of college life (1999). Although these perspectives do highlight the importance of students' behavioral involvement, they treat engagement as a primarily participation-based construct, without distinguishing emotional and cognitive dimensions (Fredricks et al., 2004; Kahu, 2013).

A major turning point came with Fredricks, Blumenfeld, and Paris' influential 2004 work, which moved beyond traditional conceptualizations of engagement as separate components and recognized their interconnectedness in a *meta-construct*. They proposed a three-dimensional framework that captured the complexity of student involvement in learning, stating that "engagement has considerable potential as a multidimensional construct that unites the three components" (Fredricks et al., 2004, p. 60): behavioral engagement (doing), emotional/affective engagement (feeling), and cognitive engagement (thinking).

In this framework, behavioral engagement refers to a student's participation in academic and social activities, including effort and persistence as well as adherence to classroom norms. Emotional engagement refers to affective reactions such as interest, enjoyment, boredom, or anxiety in relation to the school environment and learning. Lastly, cognitive engagement refers to the investment of psychological resources in learning, including for example self-regulation, strategic processing, or willingness to exert effort on a complex task (Fredricks et al., 2004).

The authors emphasized that engagement is dynamic and responsive to contextual features, arguing that studying these dimensions together offers a more nuanced understanding of how students behave, feel, and think in learning environments (Fredricks et al., 2004). Their work established the foundation for treating engagement as a multidimensional construct rather than a simple synonym for participation.

This framework is the first student engagement model to move away from two main categories of conceptualizations of engagement existing in research that were identified by Wong & Liem (2022), namely active involvement in learning and school involvement, and presents a hybrid conceptualization of engagement that "incorporate[s] both learning and school engagement elements" (p. 12), thus positioning engagement as a meta-construct (Fredricks et al., 2004; Wong & Liem, 2022). However, while Fredricks et al. highlight the number of varying definitions of engagement that exist in the literature and acknowledge that the existence of so many different conceptualizations with no clearly-defined limits can be confusing, they do not seek to refine these constructs into a clearly delimited conception of engagement (2004) beyond proposing three dimensions of it.

Later research has further refined this model, recognizing engagement as a dynamic and context-dependent construct. Wong and Liem (2022) propose the *Dual Component Framework of Student Engagement*, viewing "learning engagement and school engagement as two separate constructs under the umbrella of student engagement" (p.118). This framework emphasizes that engagement varies across individuals, subjects, and situations, and can be influenced through interventions. Korhonen et al. (2019) also highlight this multidimensional nature of engagement in higher education, demonstrating how different dimensions evolve during students' transition to university.

Moreover, Kahu (2013) proposes a framework that synthesizes behavioral, psychological, sociocultural, and holistic perspectives and conceptualizes engagement as an individual student's psychosocial state—an evolving behavioral, emotional, and cognitive connection to their learning—that is shaped by both personal characteristics and features of the educational environment. This framework is one of several that builds on Fredricks et al.'s (2004) perception of engagement as multidimensional to draw attention to the mechanisms through which the school and learning environments influence engagement, and it stresses the importance of distinguishing engagement itself from its antecedents (e.g. prior achievement, contextual affordances) and its consequences (e.g. persistence, academic success) (Kahu, 2013).

Building on Fredrick et al.'s (2004) framework, other models have also attempted to refine how engagement is conceptualised. Flow theory, for instance, which was originally developed as a theory of optimal experience rather than enjoyment per se, has been applied in educational research to conceptualize engagement as a dynamic state characterized by students' affective (e.g. enjoyment) and cognitive (e.g. absorption) experience in learning, highlighting its malleable and situational nature (Csikszentmihalyi, 1990; Shernoff et al., 2003; Wong & Liem, 2022, p.113). Martin's (2007) motivation and engagement wheel addresses the lack of clear delimitations between motivation and engagement by defining motivation as "the inclination, energy, and drive toward learning, working, and achieving" and engagement as the "behaviors that are aligned with such inclination, energy, and drive" (Wong & Liem, 2022, p.114; Martin, 2007). More recent work has also incorporated social engagement, expanding the meta-construct to include students' interpersonal involvement with peers and the wider school community (Fredricks et al., 2016; Wong & Liem, 2022).

Despite these advances in the conceptualization of engagement in research, there still appears to be definitional haziness and inconsistent operationalization of the construct (Buntins et al., 2021; Nkomo et al., 2021; Wong & Liem, 2022; Kahu, 2013). In their review of existing engagement models and analysis of their strengths and limitations, Wong and Liem (2022) highlight this lack of definitional consensus for the concept of student engagement and identify two apparent directions of interpretation in earlier student engagement models: "student engagement [that] is construed as active involvement in learning" and "students' interpersonal involvement in the school community" (p.117). In other words, they found that research on engagement was largely divided into engagement focused on active involvement in learning activities ("learning" or "classroom" engagement), and engagement focused on involvement in the school community ("school engagement").

Moreover, they identify four central issues that impede conceptual clarity in engagement research: first, overgeneralization, which occurs when diverse constructs and outcomes are grouped together under the broad label of "engagement", obscuring nuanced differences and leading to imprecise interpretations (Wong & Liem, 2022). For example, some studies treat classroom participation, motivation, and school attachment as interchangeable forms of engagement.

Second, jingle-jangle fallacies, where similar labels are used to refer to fundamentally different constructs (jingle), or different terms refer to the same construct (jangle). A concrete instance exemplified in their review is the use of "academic engagement" and "school engagement" in some research to mean both broad involvement in school and specific behaviors like attentiveness or persistence, despite underlying conceptual distinctions (Wong & Liem, 2022).

Third, object ambiguity, which arises when the target of engagement (the "object", such as classroom, activity, or subject) is either not specified or inconsistently defined within measurement tools and research frameworks. For example, Wong and Liem (2022) note how certain engagement measurement instruments fail to clarify whether they assess engagement in a subject, a pedagogical approach, or the broader school experience, leading to confusion about what is actually being measured.

Finally, under-theorization, which reflects a lack of explicit theoretical grounding that limits the development and application of engagement research. Wong and Liem (2022) note that models such as the "meta-construct" of engagement are sometimes applied without detailed connections to psychological theory or consideration of psychological processes, leading to superficial conceptualizations.

In light of this evolution, the present study adopts a multidimensional view of student engagement that recognizes the behavioral, emotional, and cognitive dimensions as interrelated components of a broader meta-construct, while also responding to calls for greater conceptual precision. Engagement is here understood as students' situated behavioral, emotional, and cognitive involvement in specific learning activities, in this case VR-based L2 German learning scenarios, rather than as a diffuse disposition towards schooling in general. This study also aims to analytically distinguish engagement from both motivation (as the force that energizes and directs action) and from later outcomes such as achievement.

This provides the conceptual basis for subsequently introducing agentic engagement in the following section as one of the more newly identified dimensions of engagement, and represents the primary dimension on which this study focuses in order to maintain a scope appropriate for a master's thesis.

2.1.2. The Emergence of Agentic Engagement Theory

Agentic engagement emerged as a fourth dimension of engagement in response to limitations of the traditional three-dimensional model that focused on behavioral, emotional, and cognitive aspects of students' involvement in learning (Fredricks et al., 2004). The most significant theoretical advancement in engagement research came with Reeve and Tseng (2011), who introduced agentic engagement, defining it as "students' constructive contribution into the flow of the instruction they receive" (p.258) via expressing preferences, making suggestions, communicating their needs, and personalizing learning, for instance. This conceptualization fundamentally shifted the understanding of learners from passive recipients of teaching to co-creators of their educational experiences and learning environments, emphasizing that students actively shape the very conditions under which engagement and learning occur (Reeve, 2013; Reeve & Tseng, 2011).

In educational psychology, Bandura (2006) had already underscored agency as individuals' capacity to intentionally influence their own functioning and the environment through forethought, self-regulation (or self-reactiveness), and self-reflection. From this perspective, students are not merely reacting to instructions, but are also capable of initiating actions that alter teaching and learning, providing a theoretical foundation for conceptualizing agentic engagement as a classroom manifestation of personal agency (Bandura, 2001; Reeve, 2013).

From the perspective of Complex Dynamic Systems Theory (CDST), agency is understood as an emergent phenomenon that develops through ongoing interactions between learners and their environment, rather than a fixed trait (Larsen-Freeman, 2019). CDST views agency as a dynamic and relational process that evolves over time, shaped by both contextual conditions and temporal factors as learners engage with, and respond to, the affordances and constraints of their learning environments (Hiver et al., 2023). This view aligns with Bandura's (2001) conceptualization of agency as learners' capacity to intentionally influence their own learning processes and outcomes, and it underscores that agentic engagement will vary across contexts, tasks, and moments as learners continuously adapt their actions.

Self-Determination Theory further complements this perspective by emphasizing the role of autonomy, competence, and relatedness in fostering intrinsic motivation and supporting proactive engagement in learning (Deci & Ryan, 1985; Reeve Ryan & Deci, 2000). When classroom environments support students' basic psychological needs, learners are more likely to take initiative or voice out preferences, thus seeking out optimal conditions for their learning, which are behaviors central to agentic engagement (Ryan & Deci, 2000; Reeve, 2013). Similarly, Reeve & Tseng (2011) conceptualize agentic engagement as a malleable, context-dependent variable that cannot be studied in isolation, but rather emerges when teachers provide supportive instruction.

In their seminal study, Reeve and Tseng (2011) pursued three main goals: validating a new measure of agentic engagement, testing whether agentic engagement functions as a distinct component of classroom engagement, and examining whether agentic engagement is an important educational construct. To address the first goal, they hypothesized that agentic engagement would correlate positively and significantly with the first three dimensions of engagement (H1), would be closely associated with classroom motivation (H2), and would predict academic achievement (H3) (p.258).

To investigate the second goal, whether agentic engagement constitutes a distinct component of engagement, they hypothesized that agentic engagement would emerge as a separate factor that is not reducible to the other three dimensions (H4) (p.259). Lastly, to address their third goal of testing its educational significance, they posited that agentic engagement would mediate the effect that student motivation will have on academic achievement, even after controlling for behavioral, emotional, and cognitive engagement (H5), thereby demonstrating its unique contribution to learning processes (Reeve & Tseng, 2011).

The results of this study supported all five hypotheses, underscoring the theoretical and practical significance of agentic engagement. For the first three hypotheses—correlation with other dimensions, association with classroom motivation, and prediction of academic achievement—they found that agentic engagement correlated positively and significantly with the other three dimensions of engagement as well as classroom motivation and “significantly and rather substantially predicted achievement” (Reeve & Tseng, 2011, p.261). For the fourth hypothesis—emergence of agentic engagement as a separate factor—the results also showed that “all five items of the Agentic Engagement Scale (…) did not cross-load onto any other engagement factor, and no item from the other three aspects cross-loaded onto the agentic engagement factor” (Reeve & Tseng, 2011, p.261), confirming that it functions as a separate factor. Lastly, path analyses revealed that “the path from agentic engagement to academic achievement remained individually significant” (Reeve & Tseng, 2011, p.263) even in models that factored in the other dimensions of engagement, demonstrating that AE explains unique variance in achievement beyond the traditional engagement components.

The authors thus bring forward two reasons why agentic engagement is particularly relevant to the learning process. First, they argue that agentic engagement contributes uniquely to achievement because “it is through intentional, proactive, and constructive acts that students find ways to improve their opportunity to learn by enriching the learning experience and by enhancing the conditions under which they learn” (Reeve & Tseng, 2011, p.263). Second, they argue that adding agency as a new dimension of engagement affords a “fuller portrayal of how students engage themselves in learning activities” (Reeve & Tseng, 2011, p.263).

Reeve (2013) further elaborated this conceptualization. He researched how students construct supportive learning environments through initiative-taking behaviors, such as asking questions or expressing preference about tasks. This work further positions agentic engagement as a student-initiated pathway through which learners increase the level of autonomy support, feedback quality, and overall instructional clarity they receive, thereby enhancing their own motivation and engagement (Reeve, 2013; Reeve & Jang, 2022).

More recent research has broadened the empirical base for agentic engagement, showing that it is a distinct factor in four-dimensional engagement scales and that it relates positively to students' perceived autonomy support, basic need satisfaction, and motivational outcomes in diverse educational contexts (Mameli & Passini, 2017; Mameli & Passini, 2019; Reeve & Tseng, 2011; Reeve & Jang, 2022). Patall (2024) synthesizes these developments in the field by arguing that agentic engagement enables students to move beyond passive motivation and compliance, allowing them to actively shape optimal conditions for their learning according to their psychological needs.

Moreover, Reeve and Jang (2022) synthesize evidence from correlational, longitudinal, and experimental studies, including two laboratory experiments on the benefits of greater agentic engagement, to show that when students are encouraged and supported to show initiative and make constructive contributions in class, they demonstrate higher levels of motivation, improved self-regulation, and enhanced academic outcomes (see p. 99). This demonstrates how agentic engagement, when assessed alongside behavioral, emotional/affective, and cognitive engagement, has the potential to predict indicators of academic progress and success, thereby strengthening the case for agentic engagement as a distinct and educationally significant fourth dimension.

As a whole, these findings underline agentic engagement as a theoretically grounded, empirically validated fourth dimension of engagement that captures learners' proactive role in shaping their instructional environments, and they provide a solid foundation for examining agentic engagement in technology-enhanced and game-based language learning contexts.

Building on this, the present study conceptualizes agentic engagement in VR-based L2 German learning as learners' proactive contributions to the design and enactment of their learning environment, such as suggesting alternative solutions, negotiating task procedures with peers, or requesting additional support and clarification from a guiding figure such as the teacher (or, in the case of this study, NPC). Treating agentic engagement as a dynamic, context-sensitive dimension alongside the other dimensions of engagement allows this study to examine how and to what degree students' agentic engagement manifests over multiple VR GBLL sessions.

2.1.3. Measurement of Engagement

Foundational Three-Dimensional Measures

The measurement of student engagement in educational research is a complex challenge, reflecting the construct's multidimensional and evolving nature (Buntins et al., 2021; Fredricks et al., 2004, 2016; Wong & Liem, 2022). Accurate assessment requires instruments and methodologies that can capture not only behavioral participation, but also emotional, cognitive, and, more recently, agentic dimensions (Fredricks et al., 2004; Reeve & Tseng, 2011; Mameli & Passini, 2017).

Fredricks et al. (2004) set an early benchmark, establishing engagement as a meta-construct comprised of behavioral, emotional, and cognitive facets. This tripartite model became the basis for widely adopted instruments such as the School Engagement Scale and the Classroom Engagement Inventory (Fredricks et al., 2004; Li & Lerner, 2013; Wang et al., 2014). These scales assess observable participation (behavioral engagement), positive feelings and belonging (emotional engagement), and investment in learning strategies and effort (cognitive engagement) (Fredricks et al., 2004; Mameli & Passini, 2017; Wang et al., 2014).

Typically, such instruments take the form of student self-report questionnaires in which learners rate the extent to which specific statements about their classroom behavior, feelings, and thinking are true for them, using Likert-type response scales (e.g. from "not at all" to "very true"), allowing separate subscores for behavioral, emotional, and cognitive engagement (Wang et al., 2014; Li & Lerner, 2013).

In the Classroom Engagement Inventory, for instance, affective engagement is measured through items covering interest, excitement, pride, happiness, and amusement in class, while behavioral engagement-compliance is assessed via items on listening carefully, paying attention, and completing assignments, and behavioral engagement-effortful participation is measured via items on getting involved in classroom activities, active contribution to discussions, and collaborative work with peers (Wang et al., 2014). Cognitive engagement is measured through items that ask students whether they review material they do not understand, reflect on mistakes, or use strategies or resources when working on tasks. Finally, disengagement is measured by items describing zoning out, for instance, or pretending to work (Wang et al., 2014).

Likewise, school-level engagement scales designed for younger students include subscales for emotional engagement (like enjoyment or feelings about school), behavioral engagement (like participation, attendance, or following the rules), and cognitive engagement (dedication to learning and strategy for learning) (Fredricks et al., 2004; Li & Lerner, 2013).

These various measuring instruments for the three-dimensional engagement framework have been peer-reviewed and validated, demonstrating satisfactory internal consistency and structural validity, and providing researchers with tools for quantifying these facets of engagement in school and classroom contexts (Wang et al., 2014; Mameli & Passini, 2017).

Taken together, these foundational three-dimensional measures provide a widely used starting point for assessing engagement. However, these approaches typically overlook students' proactive contributions to instruction and the learning process, resulting in the later inclusion of agentic engagement as a distinct dimension in measurement instruments.

Measuring Agentic Engagement as a Fourth Dimension

More recently, agentic engagement emerged as a fourth dimension, leveraging students' proactive contributions to their own learning environments (Reeve & Tseng, 2011; Reeve, 2013; Patall, 2024). Instruments like the Agentic Engagement Scale (AES) have been developed to capture classroom behaviors such as asking questions, expressing preferences, requesting clarification, suggesting changes to tasks, and communicating needs, thereby operationalizing students' constructive input into the flow of instruction (Reeve & Tseng, 2011; Reeve, 2013). Subsequent validation studies have combined the AES with three-dimensional student engagement scale to create a four-dimensional student engagement scale, enhancing measurement sensitivity to student agency while retaining the established three-dimensional model (Mameli & Passini, 2017; Reeve, 2013).

Mameli & Passini (2017) demonstrated that agentic engagement could be reliably measured as a distinct factor that does not collapse into behavioral, emotional, or cognitive engagement, and that is positively related to these other engagement types. Subsequent research has further validated that agentic engagement is positively associated with classroom motivation and perceived autonomy support, and holds unique predictive power for educational outcomes (Mameli & Passini, 2017; Reeve & Jang., 2022).

The measurement of agentic behavior thus moves beyond passive indicators, covering behaviors like question-asking, suggestion-making, and expressing preferences (Reeve, 2013;

Mameli & Passini, 2017). Experimental designs confirm agentic engagement's value: Reeve et al. (2022) found that opportunity and support for agentic behaviors predicted motivation, achievement, and improved self-regulation. These findings echo Patall's (2024) call for measuring agency as a path to understanding how students transcend passive participation in favor of active learning.

For the purpose of the present study, these four-dimensional instruments (in particular the AES) provide a solid framework for assessing agentic engagement alongside the other dimensions, and they support the decision to treat agentic engagement as a separate but related construct when examining students' proactive contributions in VR GBLL German learning scenarios.

Methodological Issues and Innovations

The field of student engagement and its measurement has faced persistent challenges. Fredricks et al. (2016) stress the definitional inconsistencies and methodological shortcomings in engagement measurement, urging for context-sensitive approaches and integration of qualitative methods to supplement surveys and self-reports. Similarly, Wong & Liem (2022) advocate for combining multiple methods and real-time behavioral indicators to capture the malleable, dynamic nature of engagement.

Recent reviews therefore recommend mixed-methods strategies: self-report surveys or interviews provide subjective insight, while observational data and real-time analytics are better suited to capture authentic behavioral engagement (Fredricks et al., 2016; Buntins et al., 2021; Kassab et al., 2023), as "the comprehensive nature of the engagement construct made it almost impossible to measure all its components using a single instrument" (Kassab et al., 2023, p.9-10). Technologies such as digital logfiles, participation tracking in virtual environments, and sophisticated physiological measures (e.g., ECG, EMG for emotional engagement in VR) have enriched the measurement toolkit (Wang et al., 2025).

Chen et al. (2024) also highlight the need for adapted measurement strategies in virtual reality (VR) and online settings. VR's immersive nature enhances emotional and cognitive engagement and can foster agentic behaviors, but also requires novel assessment approaches due to its dynamic, often non-linear interactions. Their meta-analysis shows large positive effects of VR on learning engagement, particularly for cognitive engagement and immersive experiences (Chen et al., 2024). At the same time, reviews of engagement measurement emphasize that classic self-report scales may miss nuanced behavioral and physiological signs of engagement unique to VR experiences (Kassab et al., 2023; Wang et al., 2025). Physiological measures (heart rate, muscle response), presence questionnaires, and the integration of machine learning to analyze complex engagement patterns have therefore become increasingly important (Kassab et al., 2023; Wang et al., 2025).

Towards Multimodal and Contextualized Measurement

In summary, measuring engagement requires a nuanced, multidimensional approach that blends foundational validated scales (Fredrick et al., 2004; Mameli & Passini, 2017; Wang et al., 2014), context-sensitive methods, and innovative digital and physiological analytics suited to contemporary technology-enhanced learning environments (Chen et al., 2024; Buntins et al., 2021; Wang et al., 2025). However, Buntins et al. (2021) observe that most of the tools and methods that are being used in student engagement research are still methodologically underdeveloped, especially when it is paired with educational technology.

In light of these findings, the present study will adopt a mixed-methods approach to capture a comprehensive picture of students' agentic engagement. This strategy will enable the triangulation of quantitative and qualitative data, integrating validated engagement scales with behavioral observation and contextual insight via interviews, and is especially suited to technology-enhanced learning environments such as game-based and VR contexts (Chen et al., 2024; Thi, 2025). The specific procedures and rationale for this mixed-methods design will be detailed in the forthcoming study design and methodology section.

2.1.4. Technology-Enhanced Student Engagement

Technology as an Engagement Affordance

Technology-enhanced learning environments have fundamentally changed the field of student engagement research by integrating digital tools, virtual simulations, game-based learning, and online platforms, each contributing to richer, more interactive educational experiences (Bond & Bedenlier, 2019; Fredricks et al., 2004; Nkomo, 2021; Przybylski et al., 2010). Engagement, as a multidimensional construct, is influenced by the affordances and constraints of its surrounding learning environment (whether face-to-face, online, or hybrid) and is strengthened by the strategic development of technology (Fredricks et al., 2016; Kearsley & Schneiderman, 1998; Korhonen et al., 2019).

A pivotal shift in research has been the recognition that technology is not merely an instructional supplement, but a determinant of engagement itself. Bond and Bedenlier (2019) propose a bioecological framework for student engagement based on Bronfenbrenner's (1979, 1986) bioecological theory of human development. Their conceptual model emphasizes technology's role as both direct and mediated influence: digital platforms can facilitate collaboration, personalize learning, and adapt instruction in ways unavailable to traditional settings (Bond & Bedenlier, 2019; Wang et al., 2014).

Bioecological Framework

According to Bond and Bedenlier (2019), there are a number of factors that can increase or decrease student engagement in digital/technology-enhanced contexts. The authors present a bioecological framework of engagement that takes into account all levels of Bronfenbrenner's (1979, 1986) bioecological theory of human development: the microsystem (immediate settings, like classroom, peers, teachers, family, etc.), the mesosystem (the connections between those settings, like parent-teacher communication), the exosystem (indirect external influences on the microsystem, like a parent's workplace or the neighborhood or school policies), the macrosystem (broad cultural, social, economic, and political contexts that affect the other systems, like societal values, laws, etc.), and finally the chronosystem (how changes over time influence the development of an individual, such as historical events, personal experiences, etc.).

Bond and Bedenlier's (2019) bioecological framework of engagement within technology-enhanced environments examines student engagement at all levels of Bronfenbrenner's theory. On the microsystem level, they observe that internet access and access to digital tools play a significant and direct role in student and parent engagement with their learning (Bond & Bedenlier, 2019). In the exosystem, they note that "institutions that develop a culture of student success (…) and that invest in support services and infrastructure, such as reliable internet connections and technology (…) are far more likely to promote positive student engagement" (Bond & Bedenlier, 2019, p.4), indicating that there is a strong conceptual and empirical link between school leadership and higher student engagement.

On the mesosystem level, the authors note that the relationships between microsystem elements (e.g. home and school) play a crucial role in student engagement and success. Socioeconomic background, situated in the exo- and macrosystems, shapes these relationships by influencing students' and parents' attitudes towards technology, like how not being able to afford the necessary digital tools to keep up with peers or with school may result in a negative attitude towards technology and, in turn, lower engagement (Bond & Bedenlier, 2019). Finally, the relationships between the individual and the microsystem elements such as teacher, peers, curriculum, and learning environment, are also instrumental to student engagement in technology-enhanced environments. A strong teacher-student relationship in which the teacher is perceived as "knowledgeable, supportive, invested and effective" (Bond & Bedenlier, 2019, p. 6) as well as confident in their mastery of digital tools, or "relevant and challenging" learning content (p. 6), for instance, have been shown to positively influence student engagement.

In summary, Bond and Bedenlier's (2019) framework points out that engagement in technology-enhanced learning is not only a function of individual motivation or classroom practices, but also of broader structural, institutional, and relational conditions that enable or constrain meaningful participation. This multilevel perspective has direct implications for how engagement is measured in digital contexts. Buntins et al. (2021) highlight the importance of convergence in measurement instruments, noting that technology-mediated environments require tools sensitive to new forms of engagement, such as real-time interactivity and online participation, which emerge from the complex interplay of micro-, meso-, exo-, and macrosystem influences. In other words, the very features of technology-enhanced environments that Bond and Bedenlier (2019) identify as engagement-shaping (access, infrastructure, relationships, content, and broader socioeconomic conditions) are also the features that measurement approaches must be able to capture if they are to reflect engagement's full multidimensional nature.

Empirical Evidence

Recent systematic reviews and empirical research confirm that digital technologies can strengthen student engagement by enabling deeper participation, self-paced exploration, and timely feedback (Bond & Bedenlier, 2019; Chen et al., 2024; Chiu et al., 2021; Nkomo et al., 2021**).** For instance, digital platforms and cloud classrooms replicate real-world professional environments, allowing for practical skill development and authentic assessment (Wang et al., 2024). Interactive features such as gamified quizzes, discussion boards, and adaptive tutorial systems make learning both fun and productive, inspiring students' intrinsic motivation and sustaining engagement through personalized pathways and immediate feedback (Huang et al., 2021).

The transformative impact of technology on engagement is strongly supported by meta-analytic studies, particularly in the realm of virtual reality (VR). Chen et al. (2024**)** found that many studies report increases in emotional, cognitive, and agentic engagement among students using

VR, with immersive experiences supporting active inquiry and persistence. Similarly, Thi (2025) documents that game-based learning leverages clear goals, adaptive challenges, and social interaction to behavioral, cognitive, and emotional engagement. Theoretical models such as the Self-Determination Theory (SDT) and Flow Theory help explain these outcomes by highlighting how technology-facilitated autonomy, competence, and relatedness promote sustained engagement and deeper learning (Csikszentmihalyi, 1990; Deci & Ryan, 1985; Fredricks et al., 2016; Ryan & Deci, 2000; Wang et al., 2014).

Notably, the adoption of technology has introduced new challenged and areas for further refinement. Overgeneralization and "jingle-jangle" fallacies, as noted by Wong and Liem (2022), persist in studies that conflate different types or objects of engagement (e.g. participation vs. agentic engagement), or use inconsistent terminology when measuring effects in technology-rich environments. Object ambiguity is evident when research fails to specify whether engagement refers to the online platform, the subject matter, or the social context of learning (Wong & Liem, 2022; Fredricks et al., 2016). Under-theorization is addressed in contemporary frameworks like Bond and Bedenlier's, which connect engagement outcomes more robustly to learning environments, digital tools, and classroom practices (2019).

Despite these challenges, growing evidence stresses the potential of technology-enhanced learning environments to foster engagement across diverse learner profiles and settings. Studies show that students' technology use, when designed to support students' basic psychological needs by matching it to meaningful learning experiences, mediates satisfaction, participation, and achievement (Chiu et al., 2021; Wang et al., 2025). Game-based and VR-enabled learning have especially notable effects for at-risk and disengaged students, providing adaptive, motivational experiences that support both academic and personal development (Chen et al., 2024; Maples-Keller et al., 2017; Thi, 2025; Buntins et al., 2021).

2.2. Motivation in L2 Learning

Within second language acquisition, motivation is increasingly regarded as a central driving force that directs and sustains learners' language learning (Sikandar et al., 2025; Wu, 2022). It has been consistently linked to achievement across skills and contexts, with studies reporting substantial associations between motivational strength and outcomes of L2 proficiency (Dörnyei, 2019; Yousefi & Mahmoodi, 2022).

Motivation must be differentiated from engagement, which refers to what learners do, feel, think, and proactively contribute as they are learning. Motivation concerns the underlying reasons and curiosity that make learners invest effort into their learning process and persist over time, making it a key antecedent to engagement and not a synonym (Fredricks et al., 2004; Reeve & Tseng, 2011; Sikandar et al., 2025).

Building on self-determination theory (Deci & Ryan, 1985, 2012; Ryan & Deci, 2000), motivation can be understood as a continuum ranging from externally controlled forms (external motivation) to autonomous forms (intrinsic motivation), with more intrinsic motivation generally associated with deeper involvement and greater willingness to take initiative in learning (i.e. agentic engagement). In their "Motivation in Second Language Acquisition: A State of the Art Article", Zareian and Jodaei (2015) define intrinsic motivation as "a type of motivation that derives from internal satisfaction and enjoyment" (p.302).

Recent bibliometric and systematic reviews in SLA support that research has increasingly focused on how intrinsic motivation supports learning in L2 contexts, including digital and gamified environments (Sikandar et al., 2025, Wu, 2022; Zareian & Jodaei, 2015). These studies suggest that when learners perceive language learning as personally meaningful or aligned with their goals, they are more likely to invest effort and engage in agentic behaviors such as negotiating tasks, and reach better learning outcomes (Sikandar et al., 2025; Zareian & Jodaei, 2015).

At the same time, the engagement literature has repeatedly noted conceptual blurring between motivation and engagement, with some studies treating the two as interchangeable, contributing to jingle-jangle fallacies (Wong & Liem, 2022). Martin's (2007) student motivation scale and motivation and engagement wheel explicitly address this problem by conceptualizing motivation as the inclination/drive towards learning/achieving, and engagement as the behaviors resulting from this inclination, thus differentiation the "why" (motivation) from the "what/how" (engagement) of students' involvement in classroom activities (Martin, 2007; Wong & Liem, 2022). This distinction is particularly useful for the present study, which conceptualizes intrinsic motivation as a separate construct from engagement and aims to examine the relationship that exists between them in VR-based L2 German learning.

In technology-enhanced and game-based environments, including VR, motivation can increase through the availability of feedback, which "enable[s] autonomous learning and let[s] achieve intrinsic and extrinsic motivations for learning" (Sikandar et al., 2025, p. 325), as well as culture-promoting contexts that also increase intrinsic motivation regardless of initial motivation (Sikandar et al., 2025). Reviews of motivation in digitally mediated L2 contexts emphasize that features such as interactivity or feedback can also increase enjoyment and perceived relevance, which in turn foster both drive to participate in learning activities (motivation) and willingness to take initiative in the learning activities (agentic engagement) (Ryan & Deci, 2020; Makransky & Petersen, 2021; Sikandar et al., 2025).

From a SDT perspective, the immersive and interactive nature of VR has the potential to support autonomy by allowing meaningful choices, competence by providing feedback and optimal challenging tasks, and relatedness through collaborative or social elements, thereby promoting autonomous forms of motivation (intrinsic) that favor agentic engagement (Ryan & Deci, 2000; Makransky & Petersen, 2021; Sikandar et al., 2025).

Consequently, the present study treats motivation, and more specifically intrinsic motivation, as a theoretically grounded antecedent to agentic engagement, and examines intrinsic motivation in relation to it within immersive VR L2 German learning scenarios.

2.3. Self-Efficacy in L2 Learning

Self-efficacy is a construct rooted in social cognitive theory and is defined as "people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives" (Bandura, 1994, p.71), or, in layman's terms, people's beliefs in their ability to successfully perform a task or reach a desired goal (Bandura, 1994; Makransky & Petersen, 2021; Raoofi et al., 2012). More specifically, in second language learning contexts, self-efficacy is typically conceptualized as a task/skill-specific belief about the ability to perform language activities, and it has been associated with positive learning outcomes such as higher achievement or lower language anxiety (Raoofi et al., 2012; Yang & Gan, 2024).

In technology and VR-enhanced learning environments, self-efficacy holds particular importance because learners are required to process linguistic content but also to be able to navigate digital interfaces and coordinate multimodal interaction (Chen et al., 2024; De Witte et al., 2026). Studies suggest that learners with higher self-efficacy are more likely to display behaviors that align closely with agentic engagement—such as explore spaces, experiment with solutions and strategies, and express preferences about task implementation—while low self-efficacy may lead to passivity or generally less proactive contribution, thereby limiting engagement in immersive learning tasks (Raoofi et al., 2012). Overall, technology-enhanced learning environments are considered to have a positive impact of self-efficacy (Makransky & Petersen, 2021).

Taken together, these strands of research support conceptualizing self-efficacy in the present VR-based German L2 study as an important variable that can facilitate or constrain agentic engagement in game-based scenarios, as it is one of the main factors contributing to the efforts that learners will allocate for a language learning task, and one of the strongest predicators of achievement in formal education (Raoofi et al., 2012, Makransky & Petersen, 2021). Learners who believe they can successfully complete communicative tasks and manage navigating the VR environment might be more likely to take initiative in their learning process and within the game, whereas those with lower self-efficacy may be more hesitant to act agentically (Bandura,1994; Makransky & Petersen, 2021; Raoofi et al., 2012; Yang & Gang, 2024).

2.4. Cognitive Load in L2 Learning

Cognitive load theory "describes how cognitive overload occurs when the information to be processed during learning exceeds the limited capacity of working memory" (Makransky & Petersen, 2021, p.946). Within CLT, cognitive load refers to the amount of mental effort required to process information and perform tasks, and it is typically divided into intrinsic load, which is determined by the inherent complexity of the material, extraneous load, which is imposed by the way information is presented, and germane load, which is not considered to be a load in its own right but rather assumes the role of redistributing aspects of the extraneous load to the intrinsic load (Buchner et al., 2021; Chen et al., 2024; Makransky & Petersen, 2021; Sweller, 1994; van Merriënboer & Sweller, 2005).

As van Merriënboer and Sweller (2005) note, "modern instructional theories often incorporate real-life learning tasks as a driving force, cognitive load considerations are becoming increasingly important" (p.171). In the context of technology-enhanced learning, recent systemic reviews report diverging results regarding the effects of immersive VR on cognitive load and learning performance. While some studies have suggested that VR learning environments might potentially lower cognitive load and enhance performance, for instance by providing immediate feedback or by distributing the cognitive demand among several learners in collaborative settings, the majority of studies has found that VR may increase extraneous cognitive load, especially when learners must manage navigation of the environment, interaction, and content comprehension simultaneously (Buchner et al., 2021; Han et al., 2021; Makransky & Petersen, 2021; Poupard et al., 2025).

It should be noted that the negative effects of VR on learning and cognitive load seem to be mitigated or even reversed in certain circumstances. Studies have reported that a larger knowledge base of the learning content in learners may lead to reduced cognitive load and improved performance, while others also suggest that task designers should reduce the numbers and complexity of tasks for similar outcomes (Han et al., 2021; Makransky & Petersen, 2021; Poupard et al., 2025).

For the present study, cognitive load is an essential construct to consider in relation to student engagement. Learners must simultaneously process linguistic information, navigate an immersive environment, and coordinate interaction with a peer (De Witte et al., 2026). Excessive extraneous load—for example, due to confusing interface elements or challenging task goals—may reduce learners' available cognitive resources for agentic engagement, while an appropriate balance of intrinsic and extraneous cognitive load may encourage higher agentic engagement and deeper investment in the tasks (Chen et al., 2024; Han et al., 2021; Makransky & Petersen, 2021; Poupard et al., 2025).

In line with these insights, the present study treats cognitive load as a contextual and task-related factor tat interacts with agentic engagement in VR-enhanced Germn L2 scenarios. Scenario design decisions, such as providing an initial introduction to the VR environment and providing real-life familiar scenarios, are intended to minimize unnecessary extraneous load. The study examines how learners' cognitive load, engagement, motivation, and self-efficacy, relate to each other across multiple VR learning sessions, contributing to a more nuanced understanding of how these constructs shape agentic engagement in game-based L2 learning.

2.5. Game-Based Language Learning (GBLL)

Previous studies have highlighted that supportive learning environments that encourage student engagement contribute to more effective learning processes, while also fostering higher-order cognitive skills such as critical thinking, and improving knowledge retention (Bond &

Bedenlier, 2019; Kahu, 2013; Nkomo et al., 2021; Thi, 2025). Within this broader line of research, game-based and gamified language learning environments, which introduce game mechanics to learning, have gained attention as a way to create interactive and motivating contexts that sustain learners' involvement in tasks and improve skills (Deterding et al., 2011; Gee, 2003; Hung et al., 2018; Thi, 2025).

The use of gamification in language learning has become increasingly widespread as a means of enhancing learner engagement, typically through the integration of game-like elements such as points, badges leaderboards, and levels into instructional activities (Deterding et al., 2011; Thi, 2025; Patall, 2024). From the perspective of Self-Determination Theory (Deci & Ryan, 1985), digital game-based contexts are particularly well suited to language learning goals, as they can support learners' basic psychological needs for autonomy (meaningful choices and control), competence (optimal challenge and feedback), and relatedness (social interaction and collaboration) (Thi, 2025). When these needs are effectively addressed, learners are more likely to experience intrinsic motivation, which in turn contributes to more sustained and meaningful engagement with the learning process (Patall, 2024; Reeve, 2013; Reeve & Jang, 2022; Ryan & Deci, 2020).

In terms of design, work on games and language learning has highlighted that the motivational potential of gamification should be considered with caution. An overreliance on external incentives, such as points, badges, or leaderboards, may inadvertently shift learners' attention away from the learning activity itself toward the rewards associated with it, thereby undermining intrinsic motivation, which may in turn limit deeper forms of engagement and reduce opportunities for learners to exercise agency in their learning (Deci, 1971; Deci & Ryan, 2012; Smiderle et al., 2020). Therefore, effective game-based language environments should insert language practice in rich, goal-oriented scenarios rather than relying only on isolated drills or reward structures (Gee, 2003; Thi, 2025; Wang et al., 2019). Design principles such as meaningful choices, problem solving, narrative framing, and opportunities for social interaction are generally considered to foster deeper cognitive and social enagement with language use (Hung et al., 2018, Thi, 2025; Wang et al., 2019). In multitplayer or virtual world-contexts, these principles aslo support agentic behaviors, as learners negotiate tasks, explore environments, and co-construct interaction patterns (Patall, 2024; Reeve, 2013; Wang et al., 2019).

From an engagement perspective, game-based language learning has been shown to influence multiple engagement dimensions (Thi, 2025). Reviews of digital game-based language learning indicate that affective or psychological outcomes—such as interest, enjoyment, reduced anxiety, and positive attitudes toward language learning—are among the most frequently reported benefits, closely followed by language acquisition outcomes (Hung et al., 2018; Thi, 2025; Xu et al., 2023). These environments often promote behavioral engagement by increasing time on task and participation in activities, increase affective/emotional engagement by fostering enjoyment and interest (intrinsic motivation), and support cognitive engagement by encouraging strategy use and problem solving (Fredricks et al., 2016; Hung et al., 2018; Thi, 2025; Wang et al., 2019; Xu et al., 2020, Xu et al., 2023). In addition, multiplayer and collaborative game contexts foster social engagement by providing opportunities for interaction, cooperation, and peer support, which aligns closely with relatedness needs in Self-Determination Theory, and can contribute to more agentic forms of participation (Przybylski, 2010; Thi, 2025; Thorne et al., 2009).

Empirical evidence for the effectiveness of game-based language learning is generally positive, though heterogeneous. Hung et al.'s (2018) scoping review of 50 studies on digital game-based language learning between 2007 to 2016 shows that the majority of studies report beneficial effects, particularly in terms of affective and psychological states (e.g. increased motivation and enjoyment) and language outcomes. Xu et al. (2020)'s review similarly finds that games have been used successfully to support vocabulary, grammar, listening, speaking, and reading, with many studies reporting improved performance and positive learner perceptions, although results may vary according to game genre, task design, and learner profile. Syntheses of game-based L2 teaching practices further confirm that, when grounded in solid learning theories and aligned with learners' needs, game-based approaches can enhance engagement and support language skill acquisition (Gee, 2003; Thi, 2025). These works emphasize that the balance between intrinsic and extrinsic motivation, the quality of game tasks, and the extent to which learners can act agentically within the game are crucial components to determine whether a game-based task leads to deep and sustained engagement and learning, or if it remains superficial (Hung et al., 2018; Ryan & Deci, 2020; Reeve, 2013; Smiderle et al., 2020).

2.6. Virtual Reality in Educational Contexts

Virtual Reality (VR) in education is commonly understood as a technology that provides learners with an interactive three-dimensional environment (Yu, 2021). It is a part of a wider set of immersive technologies that also include augmented reality (AR), both of which have been used to enhance and support digital game-based learning. Systematic reviews and meta-analyses across disciplines suggest that VR can enhance learning outcomes, motivation, and various dimensions of engagement, while also introducing new design challenges and cognitive demands (Chen et al., 2024; Jensen & Konradsen, 2017; Wang et al., 2019; Yu, 2021). In language education specifically, VR applications have been used to immerse learners in virtual worlds and interactive tasks that support various communicative competences (Huang et al., 2021; Wang et al., 2019).

2.6.1. General Effectiveness of VR and AR in Education

Across domains, meta-analytic evidence suggests that immersive VR using head-mounted displays (HMDs) tends to yield higher learning performance than non-immersive or semi-immersive approaches, although the effect sizes are generally moderate (Chen et al. 2024; Hamilton et al., 2021; Wu et al., 2020). For example, Wu et al. (2020) report that HMD-based immersive VR is more effective than less immersive desktop VR or traditional instruction for a range of learning tasks, echoing Yu's (2021) meta-analysis that finds a positive overall effect of VR technology on educational outcomes, suggesting that VR can support both knowledge acquisition and skill development across a range of subjects.

However, many studies and reviews do note increased extraneous cognitive load reports, mitigating the positive effects of VR and AR on learning outcomes (Garzón et al., 2019; Jensen& Konradsen, 2017; Poupard et al., 2025; Yu, 2021). Unfamiliarity with VR environments and with new interaction patterns, as well as technical issues, can add unnecessary mental effort and create confusion, especially when the learners are not prepared for the immersive experience, underlining the need for appropriate guidance and familiarization with both the VR environment and the task instructions prior to starting the learning activity (Garzón et al., 2019; Jensen & Konradsen, 2017; Mallek et al., 2024; Poupard et al., 2025; Yu, 2021).

These mixed findings suggest that, while there seems to be a tendency to agree that immersive VR can be useful and effective in certain contexts, the educational impact depends strongly on factors such as individual learners' characteristics, the content domain, or the design quality of the task. The main takeaway seems to be that clear pedagogical goals and manageable cognitive demands are essential to the design of an effective VR learning activity in terms of learning outcomes.

2.6.2. Immersive VR, Presence, and Cognitive Load

A key feature of immersive VR is the sense of presence it affords (Jensen & Konradsen, 2017; Makransky et al., 2019; Makransky & Petersen, 2021). Jensen and Konradsen (2017) found that VR HMDs can create a strong sense of "presence" in the virtual environment, and that this may support experiential learning, but can also draw attention away from core instructional content if not designed properly. Makransky et al. (2019) similarly found that adding immersive VR to a science lab simulation increased learners' sense of presence, but decreased the overall learning performance compared to a low-immersion version of the same simulation, suggesting that heightened presence can come at the cost of increased cognitive load.

2.6.3. VR, Language Learning, and Engagement

While AR and VR are being increasingly studies across many fields of education, Huang et al. (2021) note that there still remains a lack of studies on VR and game-based learning in language learning, although existing studies indicate positive outcomes on linguistic and affective skills (Hamilton et al., 2021; Huang et al., 2021; Li et al. 2017; Wang et al., 2019). The authors provide a systematic review of AR- and VR-enhanced learning environments and identify several benefits of immersing learners in virtual worlds, including increased motivation and improved learning outcomes, which echoes other studies (Huang et al., 2021, Wang et al., 2019; Yu, 2021).

With respect to student engagement, Chen et al. (2024) synthesize studies that measured behavioral, affective/emotional, cognitive, and social engagement in VR-based learning. Their meta-analysis indicates that VR has a significant positive impact on student engagement across these dimensions, with the strongest impact on cognitive engagement and the weakest on affective/emotional engagement (Chen et al., 2024). At the same time, they acknowledge the fact that the results of VR on engagement as well as learning outcomes varies greatly across studies, and note that engagement can decrease when the cognitive load is too high (Chen et al., 2024; Makransky & Petersen, 2021). They conclude that it is difficult to accurately determine the effect of VR on engagement, and call for additional research. Moreover, the majority of studies conducted on VR and engagement have been conducted in STEM fields and focus on behavioral, affective, cognitive, and social engagement, but the Humanities and agentic engagement remain understudied in comparison (Chen et al., 2024; Li et al., 2017; Patall, 2024).

For language learning, these findings suggest that VR and game-based environments can be effective in promoting engagement and supporting communicative development, but that much remains to be understood about how different engagement dimensions—especially agentic engagement—are fostered in immersive contexts. The mixed evidence on engagement and cognitive load, combined with the disciplinary skew towards STEM fields, underscores the need for more research on VR-based language learning that examines students' proactive contributions to shaping their learning experiences.

Studies have suggested methods to help reduce cognitive load and offset its negative impacts on the learning performance, such as introducing a virtual instructor to the virtual environment to provide learners with feedback in order to satisfy needs for competence and social relatedness, developing multiple user tasks to increase authenticity, or reducing the complexity of the learning activities (Makransky & Petersen, 2021).

2.7. Summary and Research Gap

Research has increasingly conceptualized student engagement as a multidimensional meta-construct covering behavioral, emotional, and cognitive dimensions (Fredricks et al., 2004; Wong & Liem, 2022). Building on this work, agentic engagement has emerged as a theoretically grounded fourth dimensions that captures learners' proactive contributions to their learning environments, with studies showing that it is a significant predictive factor for academic achievement (Patall, 2024; Reeve & Tseng, 2011; Reeve, 2013).

2.7.1. Technology-Enhanced, Game-Based, and VR Learning Contexts

In the field of technology-enhanced learning, research has shown that digital tools and platforms can support multiple dimensions of student engagement, particularly behavioral and affective engagement, but this work can often lack theoretical grounding and precise operationalization of engagement constructs (Bond & Bedenlier, 2019; Bond et al., 2020; Buntins et al., 2021; Nkomo et al., 2021; Wong & Liem, 2022). Bond and Bedenlier's (2019) bioecological framework explicitly positions the immediate learning environment as a microsystemic influence on engagement, highlighting how relationships, learning activities, and digital environments shape behavioral, cognitive, emotional, and social indicators of students' involvement. Systematic reviews of technology-mediated engagement further reveal that many empirical studies do not define engagement clearly, tend to focus on single dimensions, and rarely connect their indicators to established multidimensional models or to newer constructs such as agentic engagement (Bond et al., 2020; Bond & Bedenlier, 2019; Chen et al., 2024; Patall, 2024; Reeve & Jang, 2022).

In game-based and VR environments, findings generally point to enhanced motivation and engagement, but the evidence base remains uneven and methodologically heterogeneous (Chen et al., 2024; Buntins et al., 2021). Reviews of digital game-based technology in language learning report positive effects on leaners' interest and willingness to communicate, while also noting that most studies rely on self-report measures and do not capture learners' proactive contributions to the instructional design or task structure (Fredricks et al., 2016; Patall, 2024). Recent meta-analytic work on VR indicates that immersive experiences can increase learning engagement and motivation, but also stresses that engagement in VR is dynamic, context-sensitive, and influenced by factors such as cognitive load (Chen et al., 2024; Makransky & Petersen, 2021). Moreover, few studies explicitly examine agentic engagement in VR contexts. Reviews of engagement measurement in technology-rich contexts therefore calls for multi-method approaches that combine validated survey instruments with qualitative data to better capture the multidimensional and dynamic nature of agentic engagement in digital environments (Buntins et al., 2021; Chen et al., 2024; Fredricks et al., 2016; Wong & Liem, 2022).

2.7.2. Four Interacting Constructs

In this study, four additional constructs are examined in relation to agentic engagement: intrinsic motivation, self-efficacy, cognitive load, and overall classroom engagement. Within second language acquisition, motivation is widely regarded as a central driving force that directs and sustains learners' investment in language learning, and has been consistently linked to achievement (Sikandar et al., 2025; Yousefi & Mahmoodi, 2022). Within self-determination theory, intrinsic motivation (i.e. the internal drive to take part in learning activities rather than the external factors like rewards) is conceptualized as an antecedent of engagement that powers learners' behaviors, emotions, and cognitive factors in ways that make agentic behaviors more likely (Deci & Ryan, 2009; Ryan & Deci, 2000, 2020; Zareian & Jodaei, 2015). In the context of VR game-based language learning, intrinsic motivation is therefore likely to support students' willingness to explore the environment and contribute to the flow of instruction (Patall, 2024; Reeve, 2013; Ryan & Deci, 2000, 2020; Sikandar et al., 2015).

Self-efficacy, which refers to learners' beliefs about their capability to successfully perform specific language tasks, has also been shown to predict willingness to engage in communicative activities (Bandura, 2006; Raoofi et al., 2012). From a socio-cognitive perspective, higher self-efficacy should enable students to take initiative and express preferences in complex VR scenarios, thus functioning as a psychological resource that supports agentic engagement (Bandura, 2006; Patall, 2024; Reeve, 2013).

Cognitive load, in turn, refers to the mental effort required by a task and is typically differentiated into intrinsic, extraneous, and germane load (Sweller, 1994; van Merriënboer & Sweller, 2005). In immersive VR environments, poorly designed tasks or interfaces can create high extraneous load and restrict learners' ability to take initiative, whereas well-designed tasks can lessen extraneous load and thus free up resources for agentic behaviors (Chen et al., 2024; Han et al., 2021; Poupard et al., 2025; Reeve, 2013).

Finally, classroom engagement (in this study also referred to as simply engagement)—as a broader meta-construct covering behavioral, affective/emotional, and cognitive dimensions—provides contextual information about students' general involvement in the VR GBLL course (Fredricks et al., 2004; Fredricks et al., 2016; Mameli & Passini, 2017). Together, these four constructs are treated as theoretically grounded correlates that interact with agentic engagement and help explain how and why students' proactive contributions emerge and develop over time in VR-based L2 German learning (Patall, 2024; Reeve & Jang, 2022; Reeve & Tseng, 2011).

2.7.3. Research Gap

Taken together, existing research has increasingly established student engagement as a multidimensional meta-construct and has validated agentic engagement as a distinct, educationally significant fourth dimension that reflects learners' proactive contributions to their learning environment (Fredricks et al., 2004; Mameli & Passini, 2017; Reeve, 2013; Reeve & Tseng, 2011; Wong & Liem, 2022). At the same time, technology-enhanced, game-based, and VR learning contexts are increasingly recognized as promising environments for fostering engagement and agency (Bond & Bedenlier, 2019; Bond et al., 2020; Buntins et al., 2021; Chen et al., 2024). And yet, empirical work in these settings remains uneven, often relying on self-report scales alone, and rarely examines agentic engagement longitudinally and in relation to multiple affective and cognitive factors.

3. Theoretical Framework

3.1. Self-Determination Theory as Foundation

The Self-Determination Theory (SDT) gradually emerged in the late 1970s and early 1980s from the groundbreaking work of Edward L. Deci and Richard M. Ryan (Deci & Ryan, 1985, 2012; Ryan & Deci, 2000, 2020). The theory's origins can be traced to Deci's initial experimental investigations into the effects of external rewards on intrinsic motivation, which challenged prevailing behaviorist assumptions about human motivation, and was later formalized in Deci and Ryan's writings (Deci, 1971; Deci & Ryan, 1985, 2012; Ryan & Deci, 2000). These early studies demonstrated that external rewards could paradoxically undermine intrinsic motivation under certain conditions, leading to what became known as the "undermining effect" (Deci et al., 1999; Deci & Ryan, 2012; Ryan & Deci, 2000).

The theoretical development of the SDT represented a significant departure from dominant motivation theories of the time, which typically viewed motivation as a unitary concept varying only in quantity (Ryan & Deci, 2000). Instead, the SDT proposed that "the human organism is evolved to be inherently active, intrinsically motivated, and oriented toward developing naturally through integrative process" (Deci & Ryan, 2012, p.417). From this perspective, motivation varies not only in level but also in type, distinguishing between intrinsic motivation, various forms of extrinsic motivation, and amotivation (Deci & Ryan, 2012; Ryan & Deci, 2000; 2017, 2020). This multidimensional conceptualization has proven to be one of the SDT's most enduring and influential contributions to motivation science (Ryan & Deci, 2020; Vansteenkiste et al., 2020).

Core Theoretical Principles of SDT

The SDT is grounded in the fundamental assumption that humans are inherently active, growth-oriented organisms with natural tendencies towards psychological development and well-being (Ryan & Deci, 2000, 2017). The theory posits that optimal motivation, performance, and well-being are fostered when three basic psychological needs are satisfied: autonomy, competence, and relatedness (Deci & Ryan, 2000; Ryan & Deci, 2000, 2017).

Autonomy refers to the need to experience behavior as volitional and self-endorsed rather than controlled by external forces (Deci & Ryan, 2000; Ryan & Deci, 2000). This need involves feeling that one's actions emanate from the self and are consistent with one's values and interests (Deci & Ryan, 2000; Ryan & Deci, 2000, 2017; Vansteenkiste et al., 2020).

Competence encompasses the need to feel effective in one's interactions with the environment and to experience opportunities to exercise and express one's capacities (Ryan & Deci, 2000, 2017). Relatedness involves the need to feel connected to others, to love and care for others, and to sense that others love and care for oneself (Baumeister & Leary, 1995, Ryan & Deci, 2000).

Central to the SDT is the proposition that these three needs are universal, innate, and essential for psychological growth, integrity, and well-being across cultures and developmental stages (Deci & Ryan, 2000; Ryan & Deci, 2017). When these needs are satisfied, individuals experience enhanced intrinsic motivation, better performance, and greater psychological well-being (Ryan & Deci, 2000; Vansteenkiste et al., 2020). Conversely, need frustration leads to diminished motivation, poor performance, and psychological maladjustment (Vansteenkiste & Ryan, 2013).

SDT's Impact on Engagement Research

The influence of the SDT on student engagement research has been deep and multifaceted. Traditional models of engagement, exemplified by Fredricks et al.'s (2004) three-dimensional framework comprising behavioral, emotional, and cognitive engagement, have been significantly enriched by the SDT's theoretical insights. The theory provides a compelling explanation for why certain educational environments foster higher levels of engagement while others fail to motivate students effectively (Reeve, 2013; Reeve et al., 2004; Skinner & Pitzer, 2012).

The SDT's contribution to engagement research extends beyond merely explaining engagement outcomes, as it provides actionable principles for creating engaging learning environments. Research consistently demonstrates that when educational contexts support students' autonomy, competence, and relatedness needs, students exhibit higher levels of behavioral, emotional, cognitive, and agentic engagement (Jang et al., 2010; Reeve & Halusic, 2009; Reeve & Jang, 2022). More significantly, the SDT informed the theoretical foundation for Reeve and Tseng's (2011) conceptualization of agentic engagement as a fourth dimension of student engagement, representing students' proactive contributions to the flow of instruction they receive.

The integration of the SDT with engagement research has also led to more sophisticated understanding of engagement as a dynamic, contextually influenced phenomenon rather than a static student characteristic (Skinner & Pitzer, 2012). This perspective recognizes that engagement fluctuates based on the extent to which learning environments support or hinder students' basic psychological needs (Jang et al., 2010; Reeve et al., 2004; Skinner & Pitzer, 2012).

SDT and Game-Based Learning Research

The application of SDT to game-based learning has yielded particularly compelling insights into why games can be such powerful educational tools. Ryan et al. (2006) were among the first to systematically apply the SDT to understand the "motivational pull" of video games, demonstrating that well-designed games naturally satisfy players' needs for autonomy, competence, and relatedness. This seminal work established the theoretical foundation for subsequent research on game-based learning and educational game design (Przybylski et al., 2010; Rigby & Ryan, 2011).

Games support autonomy through meaningful choices, customization options, and player agency in determining how to approach challenges and objectives (Ryan et al., 2006; Rigby & Ryan, 2011). The interactive nature of games allow players to exercise volition and feel ownership over their experiences, contrasting sharply with many traditional educational approaches that emphasize compliance and conformity (Gee, 2003, 2005).

Competence needs are addressed through carefully calibrated challenge systems, clear feedback mechanisms, and progressive skill development opportunities that games typically provide (Csikszentmihalyi, 1990; Ryan et al., 2006). Well-designed games maintain an optimal level of challenge that stretches players' abilities without overwhelming them, creating what Csikszentmihalyi (1990) termed "flow" experiences that are inherently motivating and engaging.

Relatedness is fostered through multiplayer interactions, social features, and shared gaming experiences that create communities of practice around game-based learning (Korhonen et al., 2019; Lave & Wenger, 1991). In language learning contexts, games provide authentic social contexts for communication and collaboration that support both linguistic development and social connection (Gee, 2003, 2005; Thorne et al., 2009; Yang et al., 2010).

Integration with the Present Study

The SDT serves as the foundational theoretical framework for the present investigation of agentic engagement in game-based language learning environments for several interconnected reasons. First, the SDT provides the motivational foundation for understanding why students engage agentically with their learning environments. Agentic engagement, as conceptualized by Reeve (2013), is closely related to autonomy need satisfaction, as students who feel autonomous are more likely to express preferences, ask questions, and contribute to instructional decision-making.

Second, the SDT offers design principles for creating game-based learning environments that support all dimensions of engagement, including agentic engagement. By ensuring that learning games support autonomy through meaningful choices, competence through appropriate challenges and feedback, and relatedness through social interaction opportunities, educational designers can create contexts that naturally foster agentic behaviors (Chen et al., 2024; Rigby & Ryan, 2011; Ryan et al., 2006; Thi, 2025).

Third, the SDT provides a theoretical lens for interpreting research findings and understanding how game-based learning environments influence student engagement. The theory's emphasis on the quality rather than merely the quantity of motivation aligns with contemporary views of engagement as a multidimensional construct that varies not only in intensity, but also in type and direction (Deci & Ryan, 2000; Wong & Liem, 2022).

The present study's focus on VR game-based language learning is particularly well-suited to SDT application because virtual reality environments offer unprecedented opportunities for supporting basic psychological needs in distinctive ways. VR's immersive nature and interactive capabilities can provide highly personalized, competence-supporting experiences while maintaining social connections through collaborative features (Bond & Bedenlier, 2019; Makransky & Petersen, 2021; Wang et al., 2025). The agentic potential of VR environments—where students can make meaningful choices about their learning experiences and see immediate consequences of their actions—aligns directly with the SDT's emphasis on autonomy and competence need satisfaction (Ryan & Deci, 2017; Yang et al., 2010).

Furthermore, the study's mixed-methods approach, incorporating both quantitative measures of engagement and qualitative exploration of student experiences, reflects the SDT's focus on understanding motivation from both behavioral and phenomenological perspectives (Ryan & Deci, 2017). This methodological alignment ensures that the research can capture not only whether students are engaged, but also the quality and subjective experience of that engagement, providing insights into the psychological processes underlying agentic engagement in game-based learning environments.

3.2. Social Cognitive Theory

Social Cognitive Theory (SCT) provides a broad framework for understanding how people learn and act through the interaction of personal, behavioral, and environmental factors. Bandura (2001) describes SCT as founded on an agentic perspective in which individuals are self-organizing, proactive, self-reflective, and self-regulating rather than just reactive to external forces (Bandura, 2001). From this perspective, human functioning is explained by triadic reciprocal causation, a phenomenon in which "internal personal factors in the form of cognitive, affective, and biological events, behavioral patterns, and environmental influences all operate as interacting determinants that influence one another bidirectionally" (Bandura, 2001, pp. 14-15). Learners are therefore seen as contributors to their life circumstances, not just products of them. This conception is central to the way this thesis conceptualizes agentic engagement in classroom and VR contexts.

SCT's agentic perspective is further elaborated in Bandura's work on human agency. He identifies four core features of personal agency: intentionality, forethought, self-reactiveness, and self-reflectiveness (Bandura, 2006b). Together, these features enable individuals to plan, execute, monitor, and evaluate their actions. Intentionality refers to committing to courses of action, forethought to anticipating outcomes and setting goals, self-reactiveness to self-regulation during action, and self-reflectiveness to evaluating one's functioning and making adjustments (Bandura, 2006b). SCT thus portrays learners as agents who can intentionally shape their learning processes and environments, which aligns with the idea of agentic engagement as students' proactive contributions to the flow of instruction (Bandura, 2001, 2006b; Reeve & Tseng, 2011).

Within SCT, self-efficacy is a central construct. It links agency to motivation, persistence, and performance. Bandura (1994) defines self-efficacy as the belief someone holds in their own capabilities to organize and execute the course of action needed to produce results. In this thesis, self-efficacy is important because learners' beliefs about their ability to speak German and to manage the VR environment may shape whether they take initiative, ask questions, suggest changes, or persist when tasks become challenging. In this thesis, SCT is used to explain why learners who feel capable of managing both the German-speaking task and the VR environment may be more likely to act agentically, while those with lower self-efficacy may stay more passive.

3.3. Complex Dynamic Systems Theory

Complex Dynamic Systems Theory (CDST) provides a conceptual perspective that emphasizes nonlinearity, emergence, and context-sensitivity in human behavior and learning (Bülow et al., 2017; de Bot, 2017). In applied linguistics, CDST has been proposed as an umbrella term that brings together complexity theory (CT) and dynamic systems theory (DST), both of which view language and development as systems composed of multiple interacting components that evolve over time (Bülow et al., 2017; de Bot, 2017; Larsen-Freeman, 2019). De Bot (2017) argues that although CT and DST originate from different intellectual approaches to the same field of Systems Theory and Cybernetics—DST offers a mathematical approach, CT includes social complexity—, they are used to refer to essentially the same phenomena in language development, and also argues that adopting the unified label CDST offers greater conceptual clarity for describing such systems.

Bülow et al. (2017) apply CDST to linguistical research on language variation and change, showing how individual usage patterns, social interaction, and structural properties of varieties interact over time to produce non-linear trajectories of change. The authors treat idiolects and abstract varieties as dynamic, complex adaptive systems in which variation is not random "free variation" but a manifestation of system behavior under changing conditions, thereby illustrating how CDST can capture both micro-level behavior and macro-level outcomes in language (Bülow et al., 2017). This application to variation and change underscores CDST's relevance for the present study, which examines how learners' agentic engagement patterns evolve in response to contextual and interactional variables over time.

Language learning and classroom interaction, when combined with CDST research, are viewed as complex, adaptive systems in which cognitive, affective, social, and environmental subsystems constantly interact and co-determine each other's trajectories (Hiver et al., 2023; Larsen-Freeman, 2019). Rather than treating constructs such as motivation or engagement as stable traits, CDST conceptualizes them as processes that emerge, stabilize, and sometimes change as learners respond to the affordances and constraints of their environment (Bülow et al., 2017; Hiver et al., 2023; Larsen-Freeman, 2019). This view directly supports the process-oriented focus of the present study, in which agentic engagement is examined across multiple VR scenarios as an evolving pattern rather than a fixed characteristic.

Within this CDST framework, agency is viewed as a complex and dynamic construct that is “relational, emergent, spatially, and temporally situated” (Larsen-Freeman, 2019, p.73). Learners’ capacity to influence their own learning is not reduced to an internal disposition, but is understood as arising from ongoing interactions between learners, tasks, tools, and interlocutors (Hiver et al., 2023; Larsen-Freeman, 2019). In the present study, agentic engagement is therefore treated as a dynamic expression of agency that can strengthen, weaken, or change form as learners gain experience with the VR environment and the collaborative tasks.

A key implication of CDST is that time and trajectory become central analytical concerns. Because complex systems are inherently dynamic, CDST encourages researchers to examine how patterns of behavior and experience emerge and transform rather than focusing solely on static snapshots or pre-post differences (Hiver et al., 2023). In the specific case of student engagement, this means tracing how behavioral, emotional, cognitive, and agentic dimensions unfold over successive learning episodes, and how particular configurations of contextual factors give rise to temporary states called “attractor states”, where agentic engagement is high, but may later be destabilized (Hiver et al., 2023; Larsen-Freeman, 2019). This thesis adopts a within-subject longitudinal design, which is directly informed by Hiver et al. (2023)’s suggestion that this type of design is “most favored in order to understand the process of individual learners’ development” (p. 1).

CDST also provides a way to integrate the multiple theoretical perspectives used in this study. SDT explains how satisfaction of autonomy, competence, and relatedness needs support high-quality motivation and engagement, while social cognitive theory clarifies how self-efficacy beliefs influence whether learners take initiative or remain passive (Bandura, 2006; Deci & Ryan, 1985, 2012; Ryan & Deci, 2017, 2020). CDST frames these constructs as interacting components of a larger system: basic psychological needs, self-efficacy, cognitive load, and contextual affordances jointly shape the evolving pattern of agentic engagement in VR game-based tasks (Hiver et al., 2023; Larsen-Freeman, 2019; Makransky & Petersen, 2021). For instance, a scenario that supports autonomy and competence, moderates extraneous cognitive load, and offers opportunities for agentic behaviors and collaboration may function as an attractor state characterized by strong agentic engagement, whereas a scenario that frustrates needs or imposes excessive cognitive load may push the system toward lower engagement (Chen et al., 2024; Deci & Ryan, 2012; Sweller, 1994; van Merriënboer & Sweller, 2005).

In the context of VR-based game-based language learning, CDST highlights the complexity of the learning environment created by immersive technologies. VR scenarios in The Travel Kit offer multimodal features such as spatial immersion, embodied interaction, and collaboration, that can simultaneously support and challenge learners, depending on how they are designed and experienced (Chen et al., 2024; Makransky & Petersen, 2021). From a CDST standpoint, learners' agentic behaviors in these scenarios (e.g. requesting clarification or feedback, negotiating task procedure, proposing alternative strategies) are not simply outcomes of the system but actions that alter the system itself, reshaping social dynamics, task demands, and perceived affordances, and thereby also influencing subsequent engagement trajectories (Hiver et al., 2023; Reeve & Tseng, 2011; Reeve, 2013). Treating agentic engagement as a context-sensitive, emergent dimension within this complex system justifies the study's focus on within-learner change across scenarios and on the integration of quantitative and qualitative data to capture how agentic engagement develops over time in VR-based L2 German game-based learning.

3.4. Flow Theory

Flow theory describes an optimal experience state in which individuals become deeply absorbed in an activity, concentrate immensely, and experience the task as intrinsically rewarding (Csikszentmihalyi, 1990). It is often associated with a strong sense of involvement, time distortion, and focused attention during an activity (Csikszentmihalyi, 1990). In educational settings, flow is useful because it helps explain why some tasks hold learners' attention and support sustained concentration more effectively than others (Shernoff et al., 2003).

A central idea in flow theory is the balance between challenge and skill (Csikszentmihalyi, 1990). When a task is too easy, learners may become bored, but when it is too difficult, they may experience anxiety or frustration rather than deep involvement (Csikszentmihalyi, 1990). In contrast, when challenge and skill are balanced, learners are more likely to experience flow and remain fully engaged with the activity (Csikszentmihalyi, 1990). This is particularly relevant in game-based activities and immersive environments, where task design can be adjusted to support this balance and encourage sustained participation (Acland, 2020; Thi, 2025; Shernoff et al., 2003).

In educational research, Shernoff et al. (2003) use flow theory to conceptualize student engagement as a combination of concentration, interest, and enjoyment. Their work is especially relevant to the present study because it shows that engagement is not only about participation, but also about the quality of the learning experience itself (Shernoff et al., 2003). From this perspective, a learner can be actively involved in a task without necessarily experiencing flow, but flow can help explain moments when involvement becomes especially deep and rewarding (Shernoff et al., 2003).

Acland (2020) extends this logic into an online game context by showing that flow theory is useful for understanding motivation during the game. His study suggests that players are more motivated when they face challenges that are demanding but manageable, which aligns with the challenge-skill balance proposed by flow theory (Acland, 2020; Csikszentmihalyi, 1990). This makes flow especially relevant for VR-based game-based learning, where learners must manage both the content of the task and the demands of the initial environment (Acland, 2020).

For the present thesis, flow theory functions as a complementary perspective that helps interpret how learners experience the VR-based German tasks in *The Travel Kit*. A flow-like state may emerge when learners find the scenarios challenging but manageable, when the interface is familiar enough to not distract them from the task, and when they can remain focused on the communicative goal (Acland, 2020; Csikszentmihalyi, 1990; Shernoff et al., 2003). Flow is therefore not treated as identical to engagement, but rather as one important experiential quality that may accompany and support it (Shernoff et al., 2003).

Within the broader theoretical framework of this thesis, flow theory is used as a complement to SDT, SCT, and CDST. SDT explains why autonomy, competence, and relatedness support motivation, SCT explains why self-efficacy shapes initiative, and CDST explains how these factors change over time, while flow theory explains the subjective experience of being absorbed in the learning task (Acland, 2020; Csikszentmihalyi, 1990; Shernoff et al., 2003). Together, these perspectives allow the thesis to account not only for whether learners participate, but also for how they experience and sustain their involvement in VR-based L2 German learning.

4. Research Questions

The present study investigates how learners' agentic engagement envovles across repeated immersive VR game-based German speaking scenarios and how this proactive dimension of enagement relates to other cognitive and affective variables. Building on previous work that conceptualizes agentic engagement as a distinct, student-initiated contribution to shaping the learning environment, the study adopts a longitudinal perspective to examine both trajectories of agentic engagement over time and their association with the traditional dimensions of engagement, motivation, self-efficacy, intrinsic motivation, and cognitive load in immersive VR contexts.

Accordingly, the study addresses the following research questions:

1. **RQ1: How does agentic engagement develop over time in a VR game-based language learning intervention?**
    a. **RQ1a: What patterns of agentic engagement emerge from the quantitative data ?**
    b. **RQ1b: What obstacles and facilitators can be identified to help explain the development of agentic engagement over time ?**
2. **RQ2: To what extent do overall engagement (i.e., behavioural, emotional, and cognitive engagement), self-efficacy, intrinsic motivation, and cognitive load inform/shape patterns of learners' agentic engagement in VR game-based language learning environments over time?**

5. Study Description and Methodology

5.1. Research design

5.1.1. Overall research approach

The present study adopts a concurrent mixed-methods design with a within-subject approach to examine the relationship between agentic engagement and other cognitive and affective variables involved in speaking interaction. The quantitative strain of this study focuses on scenario-level data (motivation, cognitive load, self-efficacy, overall engagement, agentic engagement), allowing for statistical examination of within-learner change and relationships among constructs across VR scenarios. The qualitative strand uses post-session interviews to explore how learners enact agentic engagement in the VR environment.

This concurrent design serves complementary and triangulation purposes: quantitative data provide patterns of change and association among constructs, while qualitative data offer explanatory accounts and contextualized examples of agentic behavior, enabling integrated interpretations. The two strains are integrated at the discussion stage through joint displays that juxtapose quantitative trajectories (e.g. changes in agentic engagement scores across scenarios) with qualitative explanations from learner narratives.

5.1.2. Research context and development phases

The research process included preliminary testing phases prior to the main study, which were used to refine the instruments and procedures. The study was carried out by an interdisciplinary research team, with each member focusing on a specific dimension of speaking ability, such as speaking fluency, speaking anxiety, or agentic engagement. As a result, multiple types of data were collected, reflecting the different perspectives and methodological approaches brought into the project.

As the broader study involved a comprehensive series of statistical analyses, only the data and analytical procedures relevant to the research questions discussed in the present paper are addressed in the following sections. Analyses conducted and data collected for other purposes within the wider research project are beyond the scope of this thesis and therefore not discussed.

The pilot study aimed to assess the feasibility and acceptability of the VR intervention (recruitment, technical reliability), as well as the clarity and reliability of the measurement instruments. Feedback from participants and initial data analyses indicated, among other points, that unfamiliarity with the technology added extraneous cognitive load and that certain questionnaire items required wording adjustments, leading to the introduction of a preliminary "Scenario 0" and revisions to the instruments before the main trial.

A within-subject longitudinal design was selected, by which participants completed the same questionnaire and interview following each of the seven scenarios. This repeated-measures approach allowed changes in the measured constructs to be examined within individuals over time while controlling for individual differences. To minimize order effects, VR sessions were spaced over several weeks and the scenario sequence was kept constant across participants.

The main study took place concurrently in Louvain la Neuve (Belgium) and Singapore. While the study design remains largely the same as in the pilot study—within-subject longitudinal design with participants serving as the control group—some elements of the study have been modified according to feedback received during the pilot study as well as recommendations from literature and research of similar nature. In the main study, the number of scenarios was enhanced from four to seven, that the participants played in pairs, with a a questionnaire and an interview after each session.

5.2. Participants and Sampling

5.2.1. Pilot study

Participants (P) for the pilot study were recruited from students of L2 German at the Catholic University of Louvain, all with A2 level proficiency in the Common European Framework of Reference for Languages (CEFR). The final sample for this initial study consisted of seven participants, two male students and five female students, aged between 18 and 22 years at the time of the study. Although they had diverse durations of German language learning, most had studied German for approximately one year. The majority of participants were enrolled in language studies degree programs, with a particular focus on German.

The most commonly reported native language was French, with a minority of students also reporting Dutch. Participants indicated competency in many foreign languages, including English, Spanish, and Dutch. With regard to German-speaking environments, the majority of participants reported minimal to no experience living in a German-speaking area, with only one person noting a short visit.

Participants were recruited through their German language instructor.

5.2.2. Main study

The main study took place concurrently with two separate groups of participants: one in Wavre (Belgium) and one in Singapore. The participants in the latter location were 10 (6 female, 4 male) university students aged between 18-25 and enrolled at the National University of Singapore (NUS) in different majors, all taking German as an elective class. They were selected from two German classes, all with approximate A2 level proficiency in the CEFR. The majority reported English, Mandarin Chinese or Malay as first language.

The Belgium participants were 15 high school students aged between 14-18 and enrolled at the International School Le Verseau in Wavre, Brabant Wallon. They were selected from three separate class levels ranging from A2- to B1+/B2 in the CEFR. Due to the school's status as an international institution, the reported first languages varied significantly across these students, including English, French, and Dutch, among others.

For both groups, participant recruitment was conducted through their German language instructors. Participants were initially invited to form pairs of their own choosing, and those who were unable to identify a partner were subsequently assigned to a pair based on their availability and scheduling compatibility.

5.3. Materials and Experimental Environment

Author's note: the following section is largely adapted from a co-authored manuscript in preparation by He Bailing, Suner Munoz, and the author. Where applicable, content has been revised to align with the scope and terminology of this thesis.

The Travel Kit VR game environment

The primary instructional tool used in this study is a social VR-based game titled *The Travel Kit*, which was developed specifically for research purposes and has no commercial intent. The game provides an immersive environment designed to support collaborative language learning through action-oriented gameplay. It features an introduction to VR and six everyday life scenarios: in the classroom, during a job search, before a trip departure, on a campus visit, at a restaurant, and in a park. These contexts were selected to reflect common communicative situations that learners may encounter outside the classroom.

Within the game, two characters were created to facilitate interaction. For clarity and neutrality, they were given the names "Alex" and "Min," as these do not carry strong gender associations. The scenarios are designed to support learners at the A2 level of the Common European Framework of Reference (CEFR), with a particular focus on developing interactive speaking skills and related linguistic features relevant to everyday communication.

The Travel Kit was played via Metaquest 2 and Metaquest 3 headsets with hand-held controllers in a seated gaming configuration. Participants sat in separate rooms under the surveillance of a research assistant, and interacted with each other inside the VR environment through voice and movement (i.e. the in-game avatar's movements). The NPC was seated in a different room. The combination of immersive visuals and auditory input with embodied interaction was intended to support action-oriented language use and to create authentic opportunities to initiate and adapt the flow of communication within each scenario.

Scenario Design and Learning Objectives

In what follows, we describe the tasks and learning objectives for each scenario.

Scenario 0: Introduction to immersive VR

Participants took part in an individual introductory session through an in-game tutorial. They were guided through the main commands and followed simple tasks to get used to the movements and features of immersive VR.

The goal of this introductory session was to reduce extraneous cognitive load during the experiment by familiarizing the participants with the interface.

Scenario 1: In the classroom

In this scenario, learners collaboratively plan a trip by discussing their preferences regarding destinations, modes of transportation, and accommodation. The scenario focuses on developing learners' ability to express travel preferences while expanding their travel-related vocabulary and reinforcing the use of German locative and directional prepositions (*Wo? Wohin? Woher?*). At a broader level, the activity aims to foster collaborative decision-making, whereas its specific learning objectives target the acquisition of travel-related lexical items and the accurate use of the corresponding grammatical structures.

Scenario 2: searching for a job

Learners search for a job and complete it to earn money for their trip to Berlin. They have a choice between working in a restaurant or a tourism agency. In this scenario, they are asked to compare the two job offers together, then visit the location of both jobs in-game to ask the employer more questions about their tasks and benefits, then make a final decision. The students have the opportunity to interact with the work environment.

Scenario 3: Before departure

The day of the departure has come, but Alex is already at the airport and forgot their passport. The students help Alex find their missing passport by describing possible locations to Min, who is still at the apartment. This information gap is used to encourage the students to simulate a phone conversation. The second part of this scenario involves giving and following driving directions after a navigation issue.

Scenario 4: University campus visit

Learners have arrived at their destination and their first visit is a university campus. They follow a local university student (NPC) and visit the premises while exchanging information about their own school campus. They are invited to explore and interact with their environment, for example by visiting classrooms or playing soccer in the university gym. The task promotes school- and extracurricular activities-related vocabulary.

Scenario 5: In a restaurant

Learners order food and express preferences while interacting with an NPC restaurant employee. The task promotes the use of polite language and vocabulary related to food and dining.

Scenario 6: At the park

After losing personal items, learners visit a nearby park to find them.

5.4. Experimental Procedure

Participants were informed about the purpose of the study and that participation was voluntary. Prior to data collection, all participants provided written informed consent and were informed that their responses would be anonymized and treated confidentially. Ethical approval for the study was obtained from the Université Catholique de Louvain (reference CE-ILC/2025/09).

The pilot study, which took place prior to the main study that is discussed in this paper but informed its design, took place over a span of 7 weeks. Because the research team decided on a within-subject design, the participants served as a control group for themselves. They played the VR-based game concurrently with taking regular classes. They played a total of four scenarios at a rate of two 30-minute scenarios per week, in pairs, and received a questionnaire to fill out after each scenario, as well as interviews. Data from the pilot study were not included in the analyses reported in this thesis.

The main study, examined here, took place in a similar fashion with minor adjustments. Scenarios were increased from four to seven and participants took part in two game scenarios per week over the course of four weeks, unless there were schedule constraints. Each scenario was followed by a retelling task (not relevant to this study), then the questionnaire, and then an interview.

5.5. Data Collection Instruments

To collect quantitative data on learners' agentic engagement, participants completed a Likert-type questionnaire (in English) following the completion of each VR scenario, with answers labelled from "strongly disagree" (1) to "strongly agree" (5). The questionnaire was designed to measure multiple constructs simultaneously and therefore incorporated several validated measurement scales. Although the questionnaire also included additional constructs as part of the broader research project, only the data from the constructs relevant to the present study–

namely overall classroom engagement, agentic engagement, intrinsic motivation, cognitive load, and self-efficacy—will be retained for the data analysis.

For the main construct examined in this study, agentic engagement, we selected the updated Agentic Engagement Scale (AES) refined by Reeve (2013) based on the AES developed by Reeve and Tseng (2011). It is comprised of five items that capture the extent to which learners introduce their own ideas and preferences, ask for additional information or support, and attempt to influence the flow of instruction.

General engagement (the original three dimensions of classroom engagement, i.e. behavioral, emotional/affective, and cognitive) was assessed using the Classroom Engagement Inventory (CEI) developed by Wang et al. (2014). To limit participant burden in a multi-scale, repeated measures questionnaire, a reduced set of items from the CEI was selected: two items for affective (emotional), one for behavioral, and one for cognitive engagement. These items are used as engagement indicators rather than as the full validated CEI scale.

Intrinsic motivation was measured using items adapted from Makransky and Petersen's (2021) work on motivation. In their work, intrinsic motivation is conceptualized in line with self-determination theory as the drive to participate in an activity for the inherent enjoyment, interest, and challenge it provides, rather than for external rewards (Makransky & Petersen, 2021). Four items were retained that reflect task enjoyment and interest.

Cognitive load was measured by integrating multiple scales. One item each were adapted for intrinsic load (the inherent complexity of the learning material) and germane load (the mental effort invested in schema construction and meaningful learning) from Leppink's (2013) Cognitive Load Questionnaire. The remaining eight items related to extraneous cognitive load (the complexity of the task or of the way that the information is presented) were taken from Andersen & Makransky's (2021) Multidimensional Cognitive Load Scale.

Self-efficacy was measured similarly, by combining items from multiple scales. Items were adapted from Bandura's (2006) general self-efficacy scale and adapted to language learning contexts. An additional four items were adapted for VR from Meyer et al. (2019) and Pintrich et al.'s (1991) Motivated Strategies for Learning Questionnaire (MSLQ).

To ensure contextual relevance, the terminology used in the original scale items was modified when necessary to fit the particular immersive VR context of the study, for instance by replacing ‘teacher’ with ‘NPC’ to reflect the role of the virtual agent within the experimental setting.

The qualitative data collection was conducted through structured interviews after each scenario to gain a deeper and more nuanced understanding of the participants' replies to the questionnaires. This will be further elaborated in the findings section. The integration of both quantitative and qualitative data allows for a better understanding of the relationship between agentic engagement and other cognitive and affective factors in DGBLL.

5.6. Data Analysis Procedure

5.6.1. Quantitative Data

The quantitative data collected through the post-scenario questionnaires were analyzed using IBM SPSS Statistics following a series of sequential steps.

Data Preparation

The dataset was first anonymized, and questionnaire items were systematically coded according to their respective constructs (e.g., items measuring agentic engagement were labeled AE1, AE2, etc.). The dataset was then screened for coding errors, missing values, and incorrectly labelled variables to ensure accuracy. Because Linear Mixed Models use maximum likelihood estimation, participants with incomplete data were retained in the analysis, and no cases were excluded.

Reliability Analysis

Internal consistency was assessed for each construct using Cronbach’s alpha. Questionnaire items belonging to the same construct were grouped, and reliability analyses were conducted across all seven scenarios to evaluate the overall reliability of each scale.

Composite Score Computation

Following confirmation of acceptable internal reliability, composite scores were calculated for each construct. Specifically, the mean of the corresponding questionnaire items was computed for each participant at each scenario, resulting in a single construct score per participant for every measurement occasion (scenario).

Descriptive Statistics and Data Visualization

Descriptive statistics were subsequently calculated to examine the distribution of each construct across the seven scenarios. Mean scores, standard deviations, and valid sample sizes were obtained for each scenario.

Inferential Analyses

To address the first research question (agentic engagement evolution), changes in learners' agentic engagement across the seven scenarios were examined using a Linear Mixed Model (LMM). This method was selected because it appropriately handles the hierarchical structure of repeated measurements embedded within participants, accomodates for missing data without requiring listwise deletion, and enables the simultaneous modeling of both fixed (the effects studied) and random (natural differences between participants) effects. The model included "Scenario" as a fixed effect (to account for changes associated with learners' progression through the seven scenarios), "participant" as a random intercept (to account for individual differences in baseline levels of AE and the non-independence of repeated observations within participants), and a first-order autoregressive [AR(1)] covariance structure to account for the correlation between repeated observations within participants. The model was estimated using Restricted Minimum Likelihood (REML) with the Kenward-Roger approximation for the degree of freedom. Estimated marginal means and Bonferroni-adjusted pairwise comparisons were computed to identify significant differences between scenarios.

For the sub-research question RQ1a, which addresses the trends in AE trajectories that can be identified, participant-level mean AE scores were examined across the seven scenarios. Individual longitudinal profiles were systematically examined and classified into four recurring trajectory patterns:

1. Overall increase
2. Relatively stable
3. Fluctuating (with no clear rise or fall)
4. Increase then decline

These categories served both to summarize patterns of AE development across participants and to inform subsequent sampling for qualitative analysis.

To address the second research question (how other constructs inform patterns of AE), a second LMM was fitted with agentic engagement (AE) as the dependent variable and "Scenario", "engagement" (EN), "self-efficacy" (SE), "intrinsic motivation" (IM), and "cognitive load" (CL) as fixed effects. "Participant" was again specified as a random intercept, and the same AR(1) covariance structure and REML estimation procedure were employed. Model fit was evaluated using the −2 Restricted Log Likelihood, Akaike Information Criterion (AIC), corrected Akaike Information Criterion (AICC), Bayesian Information Criterion (BIC), and Consistent Akaike Information Criterion (CAIC), with lower values indicating better model fit.

5.6.2. Qualitative Data

Qualitative data were used to address sub-research question RQ1b, which aimed to identify facilitators and obstacles underlying the agentic engagement trajectory categories established in RQ1a. To this end, a subset of post-scenario interviews was subjected to qualitative analysis. Given that the qualitative component was intended to provide explanatory rather than exploratory insights, a purposive sampling strategy was employed. Participants were selected based on their longitudinal AE trajectories, which were derived from mean AE scores per participant across scenarios, to ensure representation of the principal patterns of agentic engagement development (overall increase, relatively stable, fluctuating, and increase then decline). Because participants were selected to represent specific trajectory patterns within this study rather than to form a statistically representative sample, the qualitative findings are intended to provide explanatory insights into the identified trajectories and are not intended to be generalized to all learners participating in VR game-based language learning.

A deductive qualitative content analysis was cthen onducted. Rather than analyzing the entire interview dataset inductively, the analysis focused on identifying excerpts that helped explain the observed AE trajectories. Guided by the quantitative findings, interviews were examined for participant-reported facilitators and obstacles related to engagement, self-efficacy, intrinsic motivation, and cognitive load. Additional explanatory factors were noted when they recurred across multiple participants. The identified factors were then compared within and across trajectory categories.

6. Quantitative Results

6.1. Preliminary Analyses

6.1.1. Sample and Missing Data

The LMM included 25 participants and 170 observations across seven scenarios. Five observations were missing due to non-response at individual measurement occasions.

6.1.2. Reliability of the Questionnaire Scales

Prior to the main analyses, the internal consistency of each questionnaire scale was assessed using Cronbach's alpha. As shown in Table 1, all scales demonstrated acceptable to excellent reliability (α = .754–.915), indicating satisfactory internal consistency. Composite scores were subsequently computed for each construct and used in the quantitative analyses.

Table 1

Internal consistency of the questionnaire scales

Construct	*Number of items*	*Cronbach's α*
Agentic Engagement	5	.872
Engagement	4	.754
Intrinsic Motivation	5	.906
Self-efficacy	6	.915
Cognitive Load	10	.870

6.1.3. Descriptive Statistics

Descriptive statistics computing the mean AE score across the seven scenarios for all participants indicate that learners reported moderately high levels of agentic engagement throughout the intervention (overall M= 3.71, standard deviation (SD)= 0.66). Mean agentic engagement increased from 3.14 at scenario 0 to 4.01 at scenario 5 before decreasing slightly to 3.84 at scenario 6.

Table 2

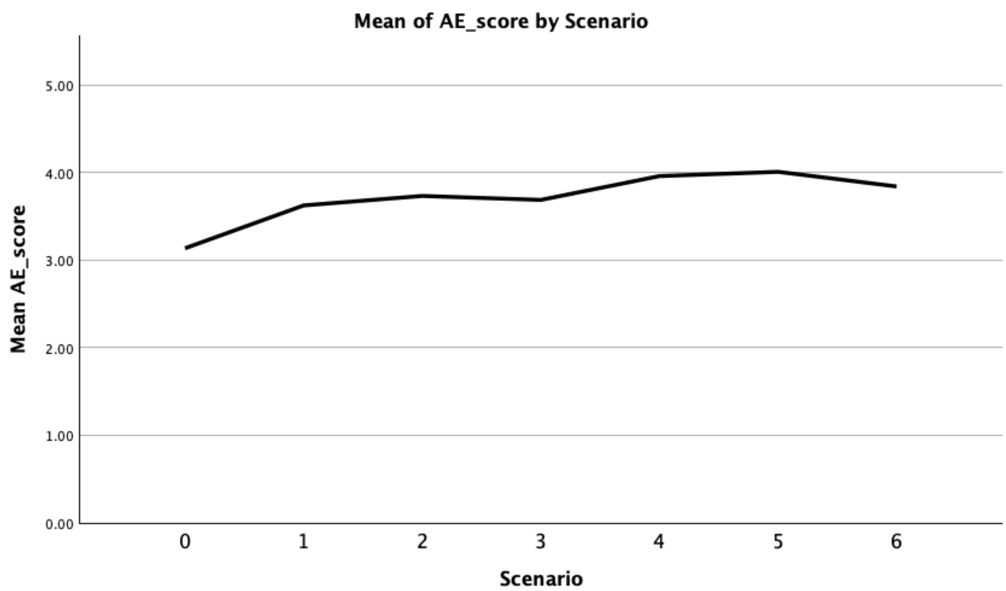
Descriptive statistics for agentic engagement across the seven scenarios

Scenario	*n*	*Mean*	*SD*
0	25	3.14	0.59
1	24	3.63	0.55
2	24	3.73	0.55

3	25	3.69	0.68
4	25	3.96	0.61
5	23	4.01	0.63
6	24	3.84	0.66
Overall	**170**	**3.71**	**0.66**

Figure 1

Descriptive statistics for agentic engagement across the seven scenarios



6.2. RQ1: Development of Agentic Engagement Over Time

As shown in Table 2 and Figure 1, mean agentic engagement generally increased throughout the intervention. Scores increased from 3.14 at Scenario 0 to a peak of 4.01 at scenario 5, before dropping slightly to 3.84 at Scenario 6. However, despite this slight decrease, AE remained higher at the end of the intervention than at the start, suggesting an overall increase trend across the seven scenarios.

Following this observation, a Linear Mixed Model was conducted to examine whether the observed differences across the seven scenarios were statistically significant. It revealed a significant main effect of Scenario (time) on agentic engagement, $F(6, 105.98) = 11.44$, $p < .001$, indicating that learners' agentic engagement differed significantly across the seven scenarios (Table 3).

Table 3
Type III tests of fixed effects for the effect of scenario on agentic engagement

Effect	*Numerator df*	*Denominator df*	*F*	*p*
Scenario	6	105.98	11.44	< .001

To determine where these differences occurred, estimated marginal means and Bonferroni-adjusted pairwise comparisons were examined. Using Scenario 6 as the reference category, parameter estimates indicated that agentic engagement at Scenario 0 was significantly lower than at Scenario 6 (Estimate = −0.703, SE = 0.120, $t = -5.85$, $p < .001$). Differences between Scenario 6 and Scenario 1-5 were not statistically significant.

Bonferroni-adjusted pairwise comparisons (Table 4) indicated that agentic engagement at Scenario 0 was significantly lower than at all subsequent scenarios (all $p \leq .002$). In addition, Scenario 5 was significantly higher in agentic engagement than Scenario 1 ($p = .038$). No other pairwise differences reached statistical significance after adjustment for multiple comparisons.

Table 4
Bonferroni-adjusted pairwise comparisons of agentic engagement across scenarios

Comparison	*Mean difference*	*SE*	p
Scenario 0 vs. Scenario 1	-0.482	0.116	.002
Scenario 0 vs. Scenario 2	-0.594	0.120	< .001
Scenario 0 vs. Scenario 3	-0.552	0.119	< .001
Scenario 0 vs. Scenario 4	-0.824	0.119	< .001
Scenario 0 vs. Scenario 5	-0.874	0.122	< .001
Scenario 0 vs. Scenario 6	-0.703	0.120	< .001
Scenario 1 vs. Scenario 5	-0.392	0.123	.038

6.2.1. RQ1a: Patterns of Agentic Engagement Trajectories

Following the identification of the overall trend in agentic engagement across the intervention, the analysis was extended to investigate individual differences in the temporal development of agentic engagement. To this end, descriptive statistics were computed by calculating the mean agentic engagement score for each participant at each of the seven scenarios. Visual inspection and systematic classification of these participant-level trajectories revealed four recurring patterns of agentic engagement development over the course of the intervention: (1) an overall increase, where AE scores exhibited a geenral upward trend across the intervention despite occasional dips, (2) relative stability of AE scores, notwithstanding minor fluctuations, (3) fluctuating AE scores that demonstrated repeated rise and falls without a clear directional trend, and (4) initial increase then decline, where AE scores increased during the initial and middle phases of the intervention but declined toward the end.

Table 5
Distribution of participants across identified agentic engagement trajectory patterns

Trajectory pattern	*Description*	*Participants (n)*
Overall increase	Scores generally increased across the intervention.	14
Relatively stable	Scores remained relatively constant across scenarios.	5
Fluctuating	Scores repeatedly rise and fall without a clear overall trend.	2
Initial increase then decline	Scores increased initially before declining in later scenarios.	4
Total		**25**

As shown in Table 5, the majority of participants (n=14) were classified into the “overall increase” trajectory, followed by “relatively stable” (n=5), “increase then decline” (n=4), and “fluctuating” (n=2). Notably, no participants exhibited an overall decrease in agentic engagement scores across the intervention.

These trajectories subsequently informed the purposive selection of participants for the qualitative analysis presented in RQ1b.

6.3. RQ2: Predictors of Agentic Engagement

A second Linear Mixed Model was conducted to examine whether intrinsic motivation (IM), cognitive load (CL), self-efficacy (SE), and engagement (EN) predicted learners’ agentic engagement as repeated measures (for each scenario). The analysis revealed significant variations across different measurements, $F(6, 109.48) = 6.46$, $p < .001$, self-efficacy, $F(1, 148.02) = 5.59$, $p = .019$, and engagement, $F(1, 158.63) = 9.95$, $p = .002$. In contrast, neither intrinsic motivation, $F(1,\ 145.34) = 0.15, p = .704$, nor cognitive load, $F(1,\ 156.67) = 1.26, p = .263$, significantly predicted agentic engagement.

Table 6
Fixed-effect estimates for predictors of agentic engagement

Predictor	b	*SE*	*df*	t	p	*95% CI*
Intrinsic motivation	0.033	0.088	145.34	0.38	.704	[−0.140, 0.206]
Cognitive load	−0.105	0.094	156.67	−1.12	.263	[−0.290, 0.080]
Self-efficacy	0.172	0.073	148.02	2.37	.019	[0.028, 0.316]
Engagement	0.325	0.103	158.63	3.15	.002	[0.122, 0.529]

Estimates of the fixed effects indicated that self-efficacy was a significant positive predictor of agentic engagement ($b = 0.172$, $SE = 0.073$, $t(148.02) = 2.37$, $p = .019$), indicating that higher self-efficacy was associated with higher levels of agentic engagement. Likewise, engagement emerged as a significant positive predictor of agentic engagement, $b = 0.325$, $SE = 0.103$, $t(158.63) = 3.15$, $p = .002$. In contrast, intrinsic motivation ($b = 0.033$, $p = .704$) and cognitive load ($b = −0.105$, $p = .263$) did not significantly predict agentic engagement after controlling for the other variables in the model.

7. Qualitative Results

Representative participants for each category were selected to help identify perceived facilitators or obstacles that may explain the observed patterns of agentic engagement development.

7.1. Overall Increase Trajectory

Data from three purposively selected participants illustrate how agentic engagement evolved alongside their experiences of confidence, motivation, psychological safety, cognitive load, and classroom engagement across the seven scenarios.

7.1.1. Self-Efficacy

The purposive sample of participants within the overall increase trajectory consistently described a gradual increase in confidence in their ability to communicate in German over the course of the intervention. Across the early scenarios, all three participants acknowledged limitations in their German vocabulary, which occasionally made it difficult to express themselves and contributed to feelings of uncertainty. For instance, P9 reports that “[it was more difficult at the beginning of the scenario] because there were not all the words I knew, and little by little, they came back a little" (S1). P11 and P22 similarly expressed their limitations with statements such as "my German level is not amazing… my vocabulary is lacking a little bit but for the most part it was okay" (P11, S2), and "I felt like I just could not come up with certain vocabulary" (P22, S2). These limitations caused the participants some levels of language-related anxiety.

Despite these initial difficulties, participants described becoming more confident as the scenarios progressed. P9 attributed this development to repeated practice of the same kind of task, explaining that "this time, it was easier than the others, I think… maybe because of training. We've done it several times, so maybe it came more easily" (S4), and later reported feeling that both their German proficiency and confidence in speaking German had improved (S5). Likewise, P22 felt that scenario 6 felt "slightly better than in past situations, because [they] were mostly describing things" (S6), meaning that the vocabulary and task had become more familiar and easier.

Overall, participants in the "overall increase" agentic engagement trajectory consistently described a progression from initial uncertainty due to lexical limitations and unfamiliar tasks toward greater confidence in communicating in German.

7.1.2. Intrinsic Motivation

Participants within the "overall increase" trajectory consistently described the VR activities as enjoyable and interesting throughout the intervention. Across interviews, they frequently emphasized that the scenarios were interesting because they resembled authentic situations that they could encounter in real life. Enjoyment remained consistently high across the seven scenarios, with no participant reporting a substantial decline in interest or willingness to participate.

P9 repeatedly characterized the activities as "really fun" across multiple scenarios. Similarly, P22 linked their motivation to the authenticity and personal relevance of the scenarios, for instance explaining that scenario 4 (campus visit) "increased [their] motivation because… it felt like [they] were really in a university in Germany" (S4).

Participants also highlighted that scenarios with clear objectives and personally meaningful topics were particularly motivating. Reflecting on the "day of departure" scenario (S3), P22 stated: "today, this task had a clearer goal. We actually had to look for something… it was quite interesting. It did increase my motivation to continue learning German" (S3). Likewise, the same participant felt particularly motivated by scenario 5 (ordering at a restaurant) because it felt like the vocabulary needed for this was actually useful in real life.

Overall, participants in this trajectory consistently described high levels of intrinsic motivation throughout the intervention, particularly when scenarios were perceived as authentic and meaningful for real-world communication.

7.1.3. Psychological Safety

The participant sample for this trajectory also consistently described the safe and supportive character of the VR environment. Across the interviews, they noted that interacting with familiar peers in a non-judgemental setting reduced feelings of nervousness or speaking anxiety and made them feel more comfortable using German, even when making mistakes or needing more time to construct their sentences.

P9 attributed this sense of comfort to working with a familiar classmate, explaining that "being with people that [they] know and not being afraid of making mistakes" (S5) made communication easier, while P11 felt comforted by the fact that they wouldn't be judged for mistakes (S4). P22 reported generally more anxiousness regarding speaking German, but noted that they felt comforted by the fact that "nothing could happen if [they] did not get the information" (S3), and that they were a bit less nervous than in class because they "could take [their] time for longer and ask for help" (S5).

Overall, the VR environment and the presence of familiar peers was seen as reducing communication anxiety and providing a supportive context in which the participants felt comfortable experimenting with the language, asking for assistance, and participating in interactions.

7.1.4. Cognitive Load

For this trajectory of "overall increase" of agentic engagement, participants generally perceived the cognitive demands of the VR activities as manageable throughout the intervention. Although there were some initial challenges related to VR equipment and environment, interface use, and unfamiliar task design, these difficulties were typically described as temporary and did not appear to prevent participants from completing the tasks.

P9 mentioned experiencing some initial difficulties with the VR controls in the early scenarios (S0-2), but reported that there was no part of the game that caused mental fatigue or overload (S2), suggesting that the overall cognitive demands of the VR game were manageable. P11 reported that the task instructions were clear and relatively easy to understand (S4), a perception that stayed true throughout the later scenarios.

P22 reported slightly greater cognitive challenges than the other participants, particularly during the earlier scenarios. For example, difficulties manipulating objects within the VR environment as well as some internet-related technological issues occasionally caused frustration. Nevertheless, these challenges appeared to lessen as the intervention progressed, and P22 reflected that "at first, I thought there could be too much information because the room looked quite complicated… but it turned out to be quite easy" (S3).

Overall, participants described initial challenges related primarily to the VR interface and unfamiliar task demands. However, as these obstacles generally diminished over time, the activities were increasingly percieved as clear, manageable, and cognitively accessible.

7.1.5. Classroom Engagement

Participants within the "overall increase" trajectory consistently reported being focused throughout the intervention. They reported experiencing little to no distractions and remaining attentive to the task, while also describing generally positive emotional experiences. In addition to maintaining attention throughout the activities, participants described actively thinking about how to communicate effectively in German. They reported planning what they wanted to say and how to say it, and adopting strategies when they encountered linguistic difficulties. Rather than disengaging when faced with challenges, they described asking for help or reformulating their sentences in order to continue the interaction.

Overall, the sampled participants for the "overall increase" trajectory described behavioral, emotional, and cognitive engagement throughout the intervention. They remained attentive, generally reported positive emotional experiences, and actively reflected on strategies that enabled them to continue participating in the communicative tasks.

7.1.6. Agentic Engagement

Across interviews, participants in this trajectory described multiple behaviors that reflect proactive contribution to the VR tasks, such as propsing ideas, asking for clarification, and persisting in communication when difficulties arose. They reported planning what they wanted to say and how to say it, and when they lacked specific vocabulary, they described taking active step to solve the issue such as reformulating or asking for help so that the interaction could continue. They also stated engaging with the task by expressing preferences or opinions and contributing actively. Overall, they portrayed themselves a willing to participate actively in the scenarios, to adapt their language use when needed, and to make use of available resources such as the NPC in order to contribute to task completion.

7.2. Relatively Stable Trajectory

Participants classified into the "relatively stable" agentic engagement trajectory exhibited moderate, consistent levels of AE across scenarios, with only minimal fluctuations after the initial orientation session in S0. Interview excerpts from two purposively sampled participants (P12 and P21) illustrate the facilitators (enjoyment, psychological safety, self-efficacy) and relatively manageable obstacles (language barrier, technical issues, task unfamiliarity) reported by participants whose agentic engagement scores followed a relatively stable trajectory.

7.2.1. Self Efficacy

Both participants reported either moderate and stable or increasing perceptions of their language ability across scenarios. P12 described early difficulty linked to limited perceived German proficiency, stating that "now it was hard for [them] because [their] German is not that good" (S1), which improved over the course of the scenarios: S4 was "in between easy and difficult", and S5 and 6 were overall easier.

P21 rated their fluency as moderate, self-evaluating it at six out of ten during S1, and expressed confidence in their ability to communicate efficiently, noting that "I was quite fluent in the sense that I could express what I wanted to say, but probably not gramatically" (S3). Both participants exhibited overall confidence in their ability to complete the communicative tasks.

P21 also highlighted how the use of question-asking and the presence of an NPC helped support comprehension and reduce nervousness, which supported engagement: "I guess if you ask questions, then you understand the meaning of what people are talking about. That increases your engagement with the game, because you understand and know how to react better to the scenario" (S4).

7.2.2. Intrinsic Motivation

The sampled participants overall described the scenarios as enjoyable and worth repeating, with some minor changes suggested. P12 emphasized fun as a motivator in multiple interviews, stating that “it’s fun and you want to do more” (S1). P21 noted that the novelty of the VR medium played a part in the enjoyment they felt because it was “kind of fun and new” (S1), while also indicating that some tasks felt too repetitive from classroom work and were therefore less interesting. However, even when enjoyment varied, P21 still reported feeling motivated by the process of task completion itself: “I might as well answer to the best of my ability whil I am still there” (S4).

Overall, participants consistently described the activities as enjoyable and motivating, although the defree of enjoyment varied depending on the scenario and task characteristics.

7.2.3. Psychological Safety

A strong facilitator for both participants was the perception that the environment is “low stakes” and non-judgemental. Both P12 and P21 reported feeling relaxed and not nervous, describing their anxiousness levels in the VR setting as generally similar or lower to their regular classroom environment. P21 explicitly contrasted the VR environment with real-life classroom environment, stating that “[they] felt like [they] would feel less nervous because [they] understand that this is an experiment, so it is not a real-life conversation. It is low stakes” (S0).

The presence of an NPC was also experienced as reassuring. P21 noted that they felt encouraged to speak more even if they made mistakes because “even if [they] did not remember a word… [the NPC] could help [them] with that, so [they] felt a bit better” (S1). P12 also expressed that the knowledge that they could ask for help made them feel more comfortable.

Moreover, P21 noted that the limited facial expressions and body language in the VR environment reduced fears of negative evaluation, making it easier to express themselves. This emotional safety was explicitly perceived to support consistent willingness to ask for clarification, express opinions, and propose changes.

7.2.4. Cognitive Load

Participants in the sample for this trajectory reported occasional increases in cognitive load associated mainly with the VR interface and task demands. P12 noted mental fatigue linked to the movements in the VR environment, and both P12 and P21 occasionally experienced dizziness from the headsets. P21 also described challenges related to lagging and inhterface issues that added to the mental effort, although they still reported focus and engagement. These cognitive factors introduced mild, recurring obstacles, but did not appear severe enough to substantially reduce agentic engagement or enjoyment and focus.

7.2.5. Classroom Engagement

Reports of attention and engagement during the game were generally consistent. Participants reported being mostly focused, despite some occasional distractions. P21 explained that S3, for instance, enhanced their focus and engagement, “because [they] were the one who had more information, [so they] basically had to guide [their partner] in the game” (S3). P21 also mentioned using strategies such as asking for help or feedback to manage task demands. These patterns generally indicated that engagement fluctuated in response to task features but remained within a moderate range overall.

7.2.6. Agentic Engagement

Consistent with their trajectory classification, participants demonstrated steady, moderate levels of agentic engagement. Agentic behaviors were primarily reflexted by participants’ willingness to ask questions, request assistance, propose improvements to the VR environment or task design, and actively contribute to task completion. P12 reported wishing that they could “swim in the water” (S6), and P21 proposed several task design enhancements that they would have liked to see, such as adding a notebook or glossary inside the game or adding interactive maps (S3, S4). Both participants reported asking the NPC for help or feedback, and trying to actively contribute to the task.

7.3. Fluctuating

Interview excerpts from two participants provide examples of learners whose agentic engagement varied more markedly across scenarios despite following the same overall intervention.

7.3.1. Self-Efficacy

For P10, self-efficacy was consistently described as high across all scenarios involving German. In scenarios 1-6, P10 repeatedly qualified the German level as “easy” and stated that they did not have to stop and think much before speaking, for instance saying “it was easy, maybe because I find it kind of easy to speak” (S1). This participant also reported that speaking in the VR environment felt the same as in real life, noting no difference in nervousness between VR and regular classroom environments. P10 indicates overall confidence in their speaking skills, even when vocabulary was lacking.

In contrast, P24 frequently reported lower or fluctuating confidence in their German, often highlighting language ability as a challenge for completing the task. They described their speech as “definitely not fluent” (S1) or even “not fluent at all” (S2), stating that “there were many moments when [they] could not think of a word, and then [they] had to formulate it in another way” (S1). Although their confidence grew a little in later scenarios, it still remained quite low: “it was not fluent, but overall, I could get my point across” (P24, S6). P24 later identified their language ability as the main challenge when trying to take initiative, reporting that they “did not think [they were] very successful, but [they] tried… but [they] could not really articulate it” (S6).

7.3.2. Intrinsic Motivation

For P10, intrinsic motivation was generally positive, with repeated references to enjoyment and willingness to repeat the activities. Across scenarios, P10 mainly credited task enjoyment for their motivation, repeating that “the situation was funny, and it was fun” (S2) and “the scenario was interesting” (S6), although they occasionally also replied “nothing” when asked what motivated them to actively participate.

P24 also reported high motivation, although they often cited personal relevance and learning opportunity as a motivator rather than the task being “fun” in and of itself. They described S0 (introduction) as interesting because it is a new technology, despite noting that “VR is not personally [their] favorite thing” (S0). They mentioned being motivated because of a personal interest in learning German, and thus seeing this experiment as an opportunity to learn.

7.3.3. Psychological Safety

Both participants reported little nervousness or anxiety, describing a more relaxed state while speaking German in VR. P10 repeatedly answererd “no” when asked if they felt nervous, explaining that being paired with someone they knew helped relax them: “the fact that it’s someone I know next to me, but also that I’m not speaking directly to them. It’s through a screen, so it’s good” (S1). P10 reported no situation in which they felt tense or anxious, and nted that did also did not feel particularly stressed in real-life communictive situations.

P24’s account is more nuanced. Although they also reported feeling relatively relaxed, they emphasized that the VR environment reduced their aniety compared with real-life classroom settings, though they acknowledged some ongoing discomfort about speaking German in general. In S1 (planning a trip), they stated that using VR “definitively felt easier” than real-life conversations, attributing this to being “isolated with a small group… [they] feel safe speaking with” (P24, S1), while in classroom environments “[the teacher] is starting at [them], and there are other people also looking…it feels like there is more pressure" (P24, S1). Overall, P24 reported feeling relaxed in the VR setting because they were there to learn.

7.3.4. Cognitive Load

Both participants indicated that the game scenarios and VR interface were easy to understand and mostly did not cause mental fatigue. P10 and P24 both reported some mild occasional dizziness due to the VR headset, but task difficulty and design were mostly considered clear and "just right". Although P10 did not report any significant changes in their perception of task difficulty, mostly describing it as "easy to understand" from S1 to S6, P24 did indicate improvement from S0 to S6. Task unfamiliarity was reported at the start (e.g. "I was not sure what to do" for S0), but this was not mentioned again as the scenarios progressed, and overall the tasks were described as "challenging, but just right" (S2, S3).

Across these interviews, both participants described using strategies such as simplifying sentences or using other ways to say what they wanted to express in order to manage moments when linguistic formulation felt effortful.

7.3.5. Classroom Engagement

Responses from both participants indicate overall consistent behavioral and cognitive engagement within the VR scenarios. They repeatedly report being focused and "very concentrated" (P10, S4) during the game and reported almost no distractions. P10 reported feeling "included in the discussion" and perceiving a need to "push [their] ideas forward" (S1).

P24 similarly described being focused and enjoying the game, and reported making an effort to understand and remember task information, as well as trying to keep an active eye on the storyline across the seven scenarios (e.g., in S6 (in the park): "I did forget the things we were looking for though. I was just picking up what I found and trying to think, 'is this what I need ? No, I need to find my watch'"). P24 also described using strategies to solve the tasks and overcome linguistical challenges.

Obstacles to engagement included the physical discomfort from the VR equipment (like nausea) and occasional task confusion in early scenarios, though these did not lead to disengagement.

7.3.6. Agentic Engagement

P10's perception describes multiple instances of taking initiative and actively contributing to the flow of the scenarios (e.g., in S2 (finding a job): "while we were in the kitchen, I asked what we had to do, and I was told that we had to put away the plates"), though they often reported not expressing preferences or needs beyond executing what the task asked of them. They described sometimes asking the NPC for help, mostly for vocabulary. At the same time, P10 frequently answered "nothing" when asked what motivated them to actively participate or bring ideas (S4), and often replied "no" when asked whether expressing opinions or preferences regarding task procedure or preferences had influenced their engagement.

P24's interviews contain more descriptions of agentic behaviors and perceived obstacles to agentic engagement. In S0, they reported trying "to do something" when other people were in the room but stated that they "did not know what [they] could have done" (P24, S0), and described abandoning attempts to use certain social features because they could not find the menu. In S1 (planning a trip) they indicated that they and their partner "took turns" in driving the conversation forward by suggesting ideas, asking questions, and offering suggestions, and P24 included specific examples such as initiating the self-introduction. For each scenario, P24 reports initiating actions and trying to go beyond what the task asks of them, asking questions and requesting feedback. Across scenarios, P24 repeatedly stated that they were motivated to "just learn German", and express occasional regrets about not asking more questions when they did not know how to say something: "for my learning experience personally, it would have been better if I had asked whenever I did not know how to express something" (S1), and "it would have helped if I had taken the initiative to [create a plan with partner on how to solve the task]" (S6).

7.4. Increase then Decline

In this trajectory, both sampled participants described factors related to self-efficacy, intrinsic motivation, psychological safety, cognitive load, and classroom engagement, as both facilitators and obstacles of agentic engagement.

7.4.1. Self-Efficacy

For P19, their self-efficacy and speaking confidence grew over the course of the scenarios. In early scenarios, they reported always having to think about grammar structure and word choice. In later scenarios, they reported their speech "becoming a bit more fluent" (P19, S5), but still reflecting afterwards on if their sentences were grammatically correct. They reported being able to form sentences more easily and faster in VR than in real-life conversations, and stated that VR practice would be useful for real-life German (for instance in S4 or S5). Overall, P19 described feeling growing confidence in their speaking abilities and their ability to solve the tasks.

P25 similarly reported hesitating a lot before speaking and low confidence, but by S6 their confidence had grown compared to the early scenarios. They also perceived the VR environment to be more conductive to speaking confidence, mentioning that they felt less pressure to be correct and feeling more encouraged to try, even if they made mistakes (S3).

7.4.2. Intrinsic Motivation

In S0 (orientation), P19 described the VR experience as “very interesting” and described VR as an interesting medium to use for learning. In S2-3, they also reported the tasks feeling “new and interesting” (S2), but noting that the tasks had no “direct or major influence” on their motivation to keep learning German (S3). This interest and enjoyment continued throughout the scenarios, with the participant genrally expressing that tasks where they could interact with the environment more freely were more motivating to them.

P25 reported some enjoyment and interest for the VR medium, but this sometimes dropped depending on the scenario. In S4 (university visit), they explained that “it was more interesting than the pas few ones. I guess it did increase my motivation a bit because I want to be able to talk more naturally in situations like that in real life” (P25, S4), showing that their personal motivation to learn German influences their interest in the task. In S6 (in the park) for example, they reported that the task was not motivating because there was no challenge.

7.4.3. Psychological Safety

Both P19 and P25 reported feeling less nervous in the VR environment than in real life. They cite low pressure and a small audience as reasons for this, which is similar to other trajectories as well. P19 reports that VR makes them “feel more confident” (S1-2) and less nervous to speak German than in real life. In later scenarios, they also described feeling relaxed because they were not scared to make mistakes, because it was not graded like in real-life classroom environment. They also reported feeling more relaxed because they could not see a real person in front of them and “felt like [they] are talking to characters” (S2). P25 mentioned similar reasons for feeling relaxed overall, noting that they usually feel “more anxious” in real life (S4).

7.4.4. Cognitive Load

P19 reported that the VR interface sometimes caused them to feel mentally exhausted and overwhelmed, although this decreased as scenarios progressed and they became more familiar with the VR environment. This mental load was mostly attributed to the movements and unfamiliar commands of the VR environment, but some tasks were also challenging to understand: “initially, when the brief was given, I felt like there was too much information. It was a lot of words, and that was a bit heavy to understand and to figure out” (S1). Later tasks were generally described as easier to understand for P19.

P25 did not report any mental fatigue due to the VR controls after S0, except for a little bit of dizziness in the last scenario. They generally described the task demands as easy, noting that it sometimes felt too easy for them.

7.4.5. Classroom Engagement

P25 reported being generally focused, with little to no distractions or interruptions. They reported that referring to the questions on the board during S1(planning a trip), for example, helped them restate what had been discussed, and that they made active efforts to remember discussion points with their partner.

P19 had a more nuanced report, mentioning that they did occasionally become distracted or feel discomfort. In S1, they were distracted by objects lying around, and in S3 they reported becoming distracted due to the lagging that sent them in and out of the “room”, which “took away from the feeling” (S3) and hindered their experience.

7.4.6. Agentic Engagement

Both P19 and P25 reported exploring the VR environment beyond what was asked of them and trying to actively participate in the conversation. For P19, early interviews show frequent reports of taking initiative, for example “[trying] to look around and catalogue what was in each room, and understand what the layout was like, outside of the instructions that we were receiving” (S0), or expressing personal opinions instead of just “passively listening” (S2).

For P25, early interviews likewise indicate agentically engaged behaviors. In the oprientation, they report taking intiative “in terms of exploring the environment, getting used to the controlsm and seeing what [they] could do inside the VR environment” (S0). In S1 and S2 they also reported trying to take oinitiative to lead the conversation, but their descriptions of agentic behavior decreased in later scenarios, for instance explaining that “there were not really many opportunities for me to contribute in the first place” (P25, S6) and that they did not really try to express preferences or opinions.

7.5. Potential Facilitators and Obstacles

Across the four trajectory types for agentic engagement evolution, several recurring potential facilitators and obstacles were identified. Although their relative prominence differs across participants, a number of common factors emerged repeatedly throughout the interview data. Table 7 summarizes these factors below.

Table 7
Summary of potential facilitators and obstacles of agentic engagement identified across the qualitative analysis

Factor	*Potential facilitators*	*Potential obstacles*
Self-efficacy	Growing or sufficient confidence in German speaking ability; familiarity with vocabulary and communicative tasks; confidence to communicate despite linguistic imperfections.	Limited vocabulary, language barriers, and unfamiliar tasks reducing confidence, particularly in the early scenarios.
Intrinsic motivation	Enjoyable, authentic, and goal-oriented scenarios; personally relevant topics; novelty of VR; opportunities to learn German in meaningful contexts.	Repetitive or classroom-like tasks; insufficient challenge; scenarios perceived as less personally relevant or less interesting.
Psychological safety	Low-stakes, non-judgemental VR environment; familiar peers; supportive NPC; feeling comfortable asking for help, making mistakes, and experimenting with the language.	No major recurring obstacles identified, although some participants continued to experience general anxiety about speaking German despite the supportive environment.
Cognitive load	Clear instructions; manageable task demands; increasing familiarity with VR controls and task format; smooth technical performance.	Initial unfamiliarity with VR navigation and task procedures; technical issues (lag, dizziness, movement difficulties); increased mental effort caused by interface-related problems.
Classroom engagement	Sustained attention and concentration; use of communication and problem-solving strategies; active collaboration	Occasional distractions; missed opportunities to take initiative; reduced engagement when tasks were perceived as repetitive, overly simple,

	with peers; emotional engagement with the task.	or offered limited opportunities for contribution.
Task characteristics	Authentic, realistic scenarios; clear objectives; opportunities for exploration and interaction; meaningful communicative tasks.	Narrow task scope; limited interaction possibilities; insufficient opportunities for proactive contribution; tasks perceived as too easy or lacking challenge.

Overall, the qualitative findings indicate that participants consistently identified growing self-efficacy, intrinsic motivation, and sustained classroom engagement as factors associated with greater willingness to participate actively in the VR activities. At the same time, lexical limitations, technical issues, unfamiliar task formats, and restrictive task designs emerged as potential obstacles that participants perceived. These findings provide explanatory insights into the quantitative patterns of agentic engagement observed across the intervention.

8. Integrated Discussion

8.1. Overall Development of Agentic Engagement (RQ1)

At the group level, agentic engagement across the VR scenarios was generally maintained at moderate-to-high levels and showed a statistically significant increase over time. The largest gain in agentic engagement occurred between the orientation (S0) and the first VR scenario (S1), after which scores continued to rise more gradually before reaching a plateau and showing a very slight decrease in the final session (S6). Despite this positive overall trajectory, considerable individual variability remained, which indicates that not all learners followed the same developmental course and that the group pattern may conceal important intra-individual differences explored later through the trajectory analysis in RQ1a & RQ1b.

Qualitative accounts across participants help explain why the overall trend is upward. Learners frequently reported increasing confidence in their ability to cope with the German language VR tasks, both linguistically and from a point of view of task familiarity, which suggests that repeated successful participation may have operated as a series of mastery experiences that strengthened self-efficacy in German language tasks. This growing confidence may have been further solidified by the psychologically safe character of the VR environment, with the majority of participants describing it as a “safe space” to try and make mistakes without fear of grading, judgement, or pressure due to the presence of a large audience.

Another recurring factor is sustained intrinsic motivation. Many students described the VR scenarios as fun, enjoyable, meaningful, and interesting, particularly when the tasks were experienced as authentic and goal-oriented. Some students highlighted that the VR game format was particularly enjoyable and that the scenarios were fun and interesting, while others rather mentioned relevance to personal goals of learning German and wanting to become able to better communicate in German as reasons why they felt motivated, while some even mentioned feeling motivated by the task process itself. Regardless of the particular reason, most reported high intrinsic motivation. Finally, another non-negligible factor identified in the qualitative data is cognitive load. Most participants reported manageable levels of extraneous cognitive load, rarely reporting feeling overwhelmed, but still noting occasional discomfort due to VR logistics or unfamiliar task procedure, especially in early scenarios.

Interestingly, while intrinsic motivation and cognitive load emerged as recurring themes in participants’ interviews, neither significantly predicted agentic engagement in the quantitative model. This may suggest that these constructs may function more as enabling conditions rather than directs predictors of agentic engagement. Learners consistently described enjoying the VR experience and experiencing manageable levels of cognitive load, but variation in these constructs alone did not explain differences in agentic engagement evolution across scenarios. In contrast, self-efficacy emerged as a both statistically significant predictor and a recurring qualitative factor.

These findings align with and extend existing theorization of agentic engagement. Reeve’s work conceptualizes agentic engagement as students’ proactive contributions that personalize learning, and shows that such engagement can increase over time (Reeve, 2013; Reeve & Jang, 2022). The present pattern, an initial jump after orientation followed by a smaller but still positive trend, corroborates this view: once learners begin to act agentically engaged in the game (by asking questions, suggesting actions, or expressing opinions and preferences, for example), they help create a more supportive learning environment for themselves, which may then sustains their willingness to continue acting agentically (Patall, 2024; Reeve et al., 2004). At the same time, the observed plateau and slight decline toward the end of the sequence bring nuance to the expectation of a simple linear increase, highlighting that agentic engagement remains contingent on perceived affordances and is not guaranteed to keep rising just because it rose in previous scenarios (Kahu, 2013; Li & Xue, 2023; Wong & Liem, 2022).

The central role of self-efficacy is shown in both quantitative and qualitative data of this study. Quantitatively, self-efficacy significantly predicted agentic engagement, while qualitative data show that students repeatedly described growing confidence as facilitating proactive participation. This strongly resonates with Bandura's self-efficacy theory, which identifies mastery experiences as the most powerful source of efficacy beliefs, and links higher efficacy to greater persistence and willingness to face challenges (Bandura, 1994, 2001, 2006b). In this study, each successful completion of a VR scenario (i.e. when learners managed to communicate and solve tasks) can be interpreted as a mastery experience that reinforces their sense that they can function in German in this environment, thereby supporting higher levels of agentic engagement in subsequent scenarios (Bandura, 1994; Raoofi et al., 2012).

With respect to the VR language learning literature, the overall pattern of increased and sustained agentic engagement in this study corroborates findings that immersive VR tends to enhance motivation, confidence, and engagement, while also adding a more longitudinal perspective (Chen et al., 2024; Huang et al., 2021; Wu et al., 2020). Systematic reviews and empirical studies have shown that VR environments typically tend to boost learner motivation and engagement and support more autonomous and interactive participation (Huang et al., 2021; Wang et al., 2019; Yang et al., 2010). The significant increase from S0 to S1 and the generally high levels of agentic engagement across the scenarios are rather consistent with these reports of VR as a motivating, engaging, and confidence-building context for language use (Chen et al., 2024; Huang et al., 2021; Yu, 2021). However, the later plateau and slight decrease, together with individual variation, contributes a nuance: while VR can foster agentic engagement, the maintenance and further growth over time depend on how many factors, such as the task design, perceived challenges, and interaction opportunities, continue to feed learners' self-efficacy, intrinsic motivation, and cognitive load.

To summarize, the first research question aiming to identifying the overall agentic engagement trend across all scenarios can be answered as follows: in this VR game-based L2 context, agentic engagement develops positively at the group level, with a pronounced initial increase and generally high subsequent levels, but this development is condition-dependent rather than uniformly linear, which is in line with complex, context-sensitive views of engagement and learner agency (Hiver et al., 2023; Larsen-Freeman, 2019; Wong & Liem, 2022). The integration of quantitative trends with learner accounts shows that increases in self-efficacy, sustained intrinsic motivation, low anxiety, and manageable cognitive load jointly support the upward trajectory in ways that are broadly consistent with Reeve's conceptualization of agentic engagement, Bandura's self-efficacy theory, and existing VR engagement research, while also highlighting the limits of these positive effects when task and medium affordances stop offering or hinder opportunities for agentic action (Bandura, 1994, 2001; Huang et al., 2021; Patall, 2024; Reeve, 2013; Ryan & Deci, 2000).

8.2. Understanding Different Trajectories of Agentic Engagement (RQ1a & RQ1b)

8.2.1. Multiple Trajectories Rather Than a Single Pattern

While agentic engagement increased at the group level, learners did not develop uniformly. Four distinct developmental trajectories were identified through quantitative analysis (RQ1a): overall increase of AE scores, relatively stable AE scores, fluctuating AE scores, and then increase then decline of AE scores. The examination of qualitative data for RQ1b suggests that these trajectories were associated with a combination of facilitators and obstacles of agentic engagement, such as self-efficacy, intrinsic motivation, cognitive load, psychological safety, and the affordances provided by individual scenarios.

The identification of several patterns (trajectories) of agentic engagement evolution across learners within the same intervention suggests that individual variability may partly reflect differences in how the learners perceived and acted on the affordances of *The Travel Kit* over time. From a CDST perspective, such divergence is expected because development emerges from the interaction of multiple learner- and context-related factors rather than from a single causal mechanism (De Bot, 2017; Hiver et al., 2023; Larsen-Freeman, 2019). Consequently, variability is not viewed as statistical "noise" but as a central feature of language learning and development (Larsen-Freeman, 2019).

This multi-trajectory pattern is consistent with broader engagement research, which conceptualizes engagement as a dynamic and context-sensitive construct rather than static. Kahu's (2013) framework emphasizes how engagement arises from the ongoing interaction between student characteristics and the learning environment, leading to shifting patterns over time even within the same educational context. Wong and Liem (2022) similarly argue that student engagement should be understood as a temporally extended process that fluctuates with changing contextual features, supports, and constraints, rather than a fixed learner attribute. Empirical and conceptual work on engagement further conceptualizes engagement as a multidimensional construct that is responsive to contextual influences, reinforcing the idea that the same instructional design can yield various engagement trajectories across learners (Fredricks et al., 2004; Korhonen et al., 2019; Li & Xue, 2023). Recent CDST-oriented work in language education and agency also argues that engagement and agency evolve through moment-to-moment adaptations to tasks and learning environment, which produces different individual trajectories (Larsen-Freeman, 2019).

Although the same general facilitators and obstacles were identified across trajectories, the qualitative findings suggest that the trajectories differed in the way these factors combined and evolved over time. Learners in the overall increase trajectory tended to report accumulating mastery experiences and increasing confidence, whereas fluctuating or declining trajectories were more often associated with temporary increases in cognitive load or inconsistent perception of challenge. This suggests that no single factor determined the developmental pathway, but rather, different configurations of interacting learner and contextual factors shaped individual trajectories.

8.2.2. Differences Across Trajectories Reflecting Interaction of Learner and Contextual Factors

Learners following an overall increase trajectory generally described a convergence of conditions that may have supported the progressively stronger agentic engagement evolution across scenarios. They reported increasing confidence in their ability to speak German and handle the VR-based tasks, with repeated successful communication experiences that may have functioned as mastery experiences, making subsequent scenarios feel easier and more manageable (Bandura, 1994). This group highlighted enjoyable, authentic, and goal-oriented tasks and noted that the game-like format kept them interested and willing to take initiative, for example by suggesting actions or experimenting with their linguistic abilities, as reflected in the qualitative interviews (Gee, 2003; Reinhardt, 2019). Cognitive demands were typically described as effortful but manageable, which is consistent with Cognitive Load Theory, which proposes that learning is facilitated when unnecessary extraneous load is minimized while sufficient cognitive challenge is maintained (Leppink et al, 2013; van Merriënboer & Sweller, 2005). Under these conditions, learners were able not only to complete the tasks but also to coshape the interaction in ways that reflected their own preferences and ideas, displaying the proactive qualities associated with agentic engagement (Reeve, 2013; Reeve & Jang, 2022).

In contrast, participants showing fluctuating or increase-then-decline trajectories more often reported mixed or shifting conditions that may have limited their ability or willingness to sustain agentic engagement. While they also described episodes of successful communication or enjoyment and interest, they reported more repeated cognitive load in the form of technical difficulties or dizziness due to the VR equipment, which disrupted immersion and may sometimes have led to lower self-reported agentic engagement (Hamilton et al., 2021; Makransky & Petersen, 2021). The participants highlighted overall more varying task perceptions: some scenarios were experienced as interactive and interesting, while others were described as repetitive, less interesting, or as offering less opportunities to act agentically. Such differences in perceived affordances and interactivity echo VR and game-based learning research showing that engagement is sensitive to design features such as meaningful choices, challenge, and opportunities for exploration (Hamilton et al., 2021; Hung et al., 2018; Przybylski et al., 2010).

The relatively stable trajectory occupies a middle ground, where learners tended to start the intervention already moderately confident in their German abilities and reported consistent, but not dramatically changing, levels of enjoyment and motivation across scenarios. They experienced relatively few major barriers: technical issues were less frequent and distracting than for the fluctuating and increase-then-decrease groups, and task demands were mostly judged as appropriate rather than confusing or over/underwhelming. These findings suggest that learners experienced an appropriate balance between challenge and support, resulting in relatively stable levels of agentic engagement throughout the intervention. Because these learners entered the intervention with moderate confidence and encountered relatively few substantial barriers, they appeared to have less room for dramatic change than participants in the other categories.

Notably, perceptions of the VR environment as psychologically safe were remarkably consistent across all trajectories. Rather than differentiating learners' developmental pathways, this appears to have constituted a common foundation that enabled participation across the intervention, while differences in self-efficacy, task affordances, and cognitive load more clearly distinguished the trajectories. Importantly, no single facilitator or obstacle appeared sufficient to determine a participant's developmental pathway. Instead, the qualitative findings suggest that different combinations of learner- and context-related factors may have influenced the appearance of different patterns of agentic engagement across the intervention.

Together, these four trajectories illustrate that agentic engagement in this VR game-based L2 context is better understood as an evolving, context-responsive process than as a stable learner characteristic. This interpretation aligns with engagement frameworks that conceptualize engagement as dynamic, multidimensional, and sensitive to shifting contextual features, rather than a fixed trait (Fredricks et al., 2004; Kahu, 2013; Wong & Liem, 2022). CDST-oriented work in language education similarly emphasizes that agency is manifested differently over time as learners interact with changing task designs, interlocutors, and learning conditions, leading to diverse developmental pathways even within a single intervention (Hiver et al., 2023; Larsen-Freeman, 2019). In this sense, the four trajectories in the present study do not represent anomalies around a single norm, but they exemplify the range of ways in which agentic engagement can emerge, stabilize, fluctuate, or decline as learners and environments co-evolve within this particular group of participants.

While the trajectory analysis helps explain why learners followed different developmental pathways, it does not indicate the relative contribution of the individual learner-related constructs examined quantitatively. The following section therefore integrates the quantitative and qualitative findings for self-efficacy, classroom engagement, intrinsic motivation, and cognitive load in order to examine how each construct contributed to the evolution of agentic engagement throughout the intervention.

8.3. Role of the Four Constructs in Agentic Engagement Evolution (RQ2)

While the first research question examined the overall evolution of agentic engagement across the seven scenarios, and the research sub-questions RQ1a & RQ1b identified more precise patterns of agentic engagement evolution within the participant group (RQ1a) and identified potential facilitators and obstacles and how they may help explain those patterns (RQ1b), the second research question aims to examine the relationship between the four constructs and agentic engagement, and to what extent they shape patterns of agentic evolution over time.

Across the VR intervention, the four focal constructs besides agentic engagement—self-efficacy, intrinsic motivation, classroom engagement (behavioral, emotional, cognitive), and cognitive load—played roles in how agentic engagement evolved for the participants.

Quantitatively, self-efficacy and classroom engagement emerged as significant predictors of agentic engagement over time, whereas intrinsic motivation and cognitive load did not. Qualitatively, however, all four constructs were repeatedly mentioned in learners' interviews in relation to agentically engaged behaviors, suggesting partial convergence between the two strands of data. This pattern is consistent with contemporary engagement and CDST perspectives, which view engagement as an emergent result of multiple factors rather than any single factor alone (Fredricks et al., 2004; Hiver et al., 2023; Kahu, 2013; Larsen-Freeman, 2019).

8.3.1. Convergence and Divergence Between Quantitative and Qualitative Findings

There is strong convergence between the quantitative and qualitative data for self-efficacy and classroom engagement in this study. Both constructs were statistically associated with higher AE and more positive AE trajectories, and learners connected feeling confident and cognitively invested with acting more agentically in the VR tasks. By contrast, intrinsic motivation and cognitive load did not appear as significant unique predictors of AE in the quantitative findings, yet they were still frequently highlighted in interviews as conditions that either supported or constrained learners' willingness and ability to act agentically.

Self-efficacy showed the clearest convergence across methods in the present study. Quantitatively, higher self-efficacy was significantly associated with higher levels of AE and with more favorable AE trajectories over time. Qualitatively, the sampled participants across all four trajectory types described changes in their confidence to speak German and handle VR-based tasks as central to their agentically engaged behaviors. This pattern converges closely with Bandura's self-efficacy theory and its mastery experiences, which are a powerful source of efficacy beliefs, which link higher self-efficacy to greater effort, persistence, and willingness to face challenges (Bandura, 1994, 2001, 2006b). In the VR context of this study, repeated successful completion of scenarios (communicating meaning, solving problems, repairing misunderstandings in German) can be interpreted as mastery experiences that strengthened learners' beliefs that they could function effectively in this environment. In scenarios where self-efficacy increased, learners also generally report more proactive behaviors (Reeve, 2013; Reeve & Jang, 2022). This interpretation is consistent with L2-specific research showing that self-efficacy may predict language learning behaviors and outcomes, including willingness to communicate and sustained engagement with tasks (Raoofi et al., 2012; Yang & Gan, 2024).

Classroom engagement also showed convergence between the quantitative and qualitative findings. Quantitatively, higher levels of engagement such as effort, concentration, and the adoption of cognitive strategies, were associated with higher AE and more positive AE trajectories. Qualitatively, learners who described staying focused and investing mental effort to understand and navigate the scenarios also generally reported more frequent instances of taking initiative and shaping the interaction, although it must be noted that this was to a lesser degree than self-efficacy, and that some participants with lower agentic engagement or self-efficacy sometimes still reported being focused and adopting strategies to a certain extent. This pattern, when present, reflects multidimensional models of engagement that conceptualize behavioral, emotional, cognitive, and agentic engagement as interrelated facets of a broader engagement construct (Mameli & Passini, 2017, 2019; Reeve, 2013; Reeve & Tseng, 2011). It also aligns with work suggesting that higher general engagement often co-occurs with more agentic forms of participation, particularly in environments that combine autonomy support by a guiding figure with clear structure (Jang et al., 2010; Reeve & Jang, 2022).

Intrinsic motivation shows a more pronounced convergence-divergence pattern between quantitative and qualitative data. In the quantitative data analysis, IM did not emerge as a significant unique predictor of AE once self-efficacy and classroom engagement were included. However, interview data indicated that interest, enjoyment, curiosity, and perceived meaningfulness of the VR scenarios were central in many learners' own explanations of why they were willing to "do more" and act agentically. This pattern is in line with self-determination theory, which posits that intrinsic motivation thrives when basic psychological needs for autonomy, competence, and relatedness are supported, which is strongly associated with deeper engagement and persistence (Deci & Ryan, 1985, 2000; Ryahn & Deci, 2017, 2020). The game-like and immersive design of the VR tasks in this study appears to have supported intrinsic motivation for many learners, particularly when scenarios were experienced as authentic, challenging, and goal-oriented. At the same time, the non-significant unique effect of intrinsic motivation in the quantitative model suggests that its influence may have been perhaps mediated or shared with more close constructs such as self-efficacy and overall classroom engagement. This interpretation is supported by prior work indicating that intrinsic motivation often energizes engagement indirectly and can be difficult to detect as an independent predictor once related variables are statistically controlled (Przybylski et al., 2010; Ryan et al., 2006).

This situation illustrates the value of the explanatory mixed-methods design employed by this study. While the quantitative analysis identified the constructs that uniquely predicted agentic engagement after statistical adjustment, the qualitative findings revealed additional experiential factors that participants themselves perceived as important in shaping their willingness to act agentically.

Cognitive load exhibits a similar pattern in the present study. Statistically, it did not uniquely predict AE trajectories once the other constructs were entered into the models. However, qualitative data shows that participants frequently commented on task difficulty, mental effort, and the balance between challenge and manageability, as well as extraneous factors such as technical affordances and the presentation of task information. Some scenarios were described as cognitively challenging but manageable by some sampled participants, which was often associated with feeling focused and willing to take initiative. Other scenarios were experiences as a tad overwhelming, for example due to linguistic unfamiliarity or VR equipment issues, which were sometimes associated with a shift toward a slightly more passive attitude. Conversely, some participants also reported finding the tasks too easy or repetitive, which was then linked to boredom and reduced inclination to act agentically. These accounts are consistent with Cognitive Load Theory, which holds that learning and performance are optimized when intrinsic and extraneous load are balanced so that working memory resources can be devoted to germane processing (i.e. the mental processing involved in meaningful learning), whereas excessive or insufficient load can undermine engagement (Leppink et al., 2013; Sweller, 1994; van Merriënboer & Sweller, 2005). In immersive VR environments, the balance of cognitive load is particularly delicate because high presence and sensory input may simultaneously increase processing demands and enhance motivation (Makransky & Petersen, 2021; Poupard et al., 2025). In this study, cognitive load therefore appears to function primarily as a boundary condition for agentic engagement: when load was too high or too low, learners tended to retreat to minimal participation, whereas moderate, manageable levels of challenge supported the enactment of agentic behaviors, even if this role may not have appeared as a strong independent predictor in the statistical models.

8.3.2. Integration

Overall, the mixed-methods findings indicate a differentiated but interactive role for the four constructs in the evolution of agentic engagement over time in this specific VR context. Quantitatively, self-efficacy and classroom engagement emerged as the two key statistical drivers for agentic engagement evolution over time, explaining unique variances in participants' AE trajectories, which was also reflected in the qualitative data. Qualitatively, intrinsic motivation and cognitive load appeared alongside those two constructs as potentially enabling or constraining conditions that shaped when and how learners could translate their confidence and general engagement into agentic action, even though this was not reflected in the quantitative data. This configuration is not contradictory, and it aligns with CDST-informed accounts of language learner development, which emphasize that outcomes emerge from the interaction of multiple, sometimes overlapping subsystems rather than from linear, additive effects (De Bot, 2017; Larsen-Freeman, 2019). It is also in line with contemporary engagement frameworks that highlight the dynamic, context-sensitive nature of engagement (Kahu, 2013; Wong & Liem, 2022). From this perspective, self-efficacy and classroom engagement can be understood as core drivers of agentic engagement change over time, while intrinsic motivation and cognitive load may be understood to operate as contextual and experiential conditions that enable or constrain the realization of agentic potential in VR language learning. Over time, the interplay of all four constructs may help explain why learners with similar starting points and identical exposure to the VR tasks still followed different agentic engagement trajectories in the present study.

8.4. Limitations

Several limitations of the present study should be acknowledged when interpreting the findings.

First, classroom engagement was measured using a reduced version of the classroom engagement inventory (CEI), which may not fully capture the breadth of behavioral, emotional, and cognitive engagement discussed in engagement literature (Fredricks et al., 2004; Wang et al., 2014). Using a shortened version of the scale is practical and acceptable in applied settings, but it can reduce reliability and limit comparability with studies that employe the full instrument or other validated multi-dimensional engagement measures. This was done with the intention of reducing the already long, multi-scale questionnaires, in order to minimize the burden for participants.

Second, the study adopted a strongly process-oriented focus on the development of agentic engagement over time, rather than on language learning outcomes. As a result, the analyses cannot determine whether observed trajectories translate into measurable gains in German proficiency.

Third, the findings are limited within a specific context: VR-based, game-based L2 German learning using a game designed for a single, relatively constrained purpose. Although this design allowed for precise tracking of agentic engagement patterns within a coherent set of scenarios, it also means that the results may not transfer to other target languages, age groups, institutions, or VR/game designs with different goals, mechanics, or levels of immersion (Huang et al., 2021; Jensen & Konradsen, 2017; Wang et al., 2019). Generalizations beyond this particular VR game and setting should therefore be made with caution or avoided.

Fourth, this study relied primarily on self-reported measures (questionnaires and post-hoc interviews) to assess the different variables examined. While self-report is widely used and accepted in engagement and motivation research and can capture internal states that are difficult to observe directly, it is susceptible to biases and discrepancies between perceived and enacted behavior (Fredricks et al., 2016; Wong & Liem, 2022). In some cases, participants described agentic behaviors in interviews that were not observed to the same extent during gameplay, suggesting that observational or trace-based measures could have provided valuable complementary evidence.

Fifth, the qualitative component employed purposive sampling of participants representing each agentic engagement trajectory type, which is appropriate for generating explanatory insights but not for statistical generalization. The narratives and explanatory patterns identified in the qualitative analysis are therefore best understood as theoretically informative and context-bound, rather than as representative of all learners in VR-based language learning (Kahu, 2013; Hiver et al., 2023).

Lastly, the total sample size of 25 participants is acceptable for an in-depth, context-specific mixed-methods study, but remains relatively small for quantitative modelling. The limited N constrains statistical power, especially for detecting smaller effects or more complex interaction patterns between constructs. This may partly explain why some variables (e.g. intrinsic motivation and cognitive load) did not emerge as significant unique predictors. The primary aim was to model and understand trajectories within this specific group and context, but future work with larger samples could test the strength and generalizability of the identified patterns.

8.5. Implications for Future Research

Future research could address these limitations in several ways.

First, studies might use the full CEI scale or other validated multi-dimensional engagement scales, potentially alongside agentic engagement measures, to allow for more distinctions among the different engagement types and to facilitate comparison across contexts.

Second, future work should integrate process-oriented engagement measures with language outcome indicators, such as pre/post-tests of German proficiency, performance-based speaking assessments, or longitudinal curricular grades. This would make it possible to examine whether particular AE trajectories are associated with differential gains in L2 performance and to test more directly the hypothesized links between AE and learning outcomes (Reeve, 2013; Ryan & Deci, 2020).

Third, to enhance generalizability, future research could replicate and extend this design in other languages, educational levels, cultures, and VR/game configurations. Comparative work across different VR designs and degrees of gamification would help clarify which design features most strongly support agentic engagement and whether similar trajectories emerge in other contexts (Gee, 2003; Huang et al., 2021; Reinhardt, 2010).

Fourth, future research should incorporate additional data sources beyond self-report, such as systematic video-based observation, interaction transcripts, or in-game behavioral logs. Combining these modalities with self-report would allow researchers to triangulate perceived and enacted agentic behaviors.

Fifth, while purposive sampling is appropriate for explanatory qualitative work, larger-scale studies could use stratified or random sampling strategies within VR language programs to examine the prevalence of different AE trajectories and their correlates in more diverse populations. Such designs would complement the present explanatory findings with stronger, more general claims about these patterns and the conditions under which they are likely to occur.

Lastly, future research would benefit from larger sample sizes and more measurement occasions. This would allow for more exploration of reciprocal or causal relations between the constructs. Such work would also be well-suited to CDST approaches that treat agentic engagement as an evolving process shaped by the ongoing interaction of learner and contextual factors, and not as a stable trait.

9. Conclusion

9.1. Summary of the Study

This thesis was motivated by the observation that, despite growing interests in immersive virtual reality for language learning and game-based approaches, there is still very little longitudinal research on how agentic engagement develops over time in VR game-based language learning (GBLL) contexts (Huang et al., 2021; Hung et al., 2018; Patall, 2024; Reeve, 2013; Yu, 2021). Existing studies on VR and digital game-based L2 learning have generally focused on short-term outcomes, motivation, or general engagement, leaving the temporal dynamics of agentic engagement largely unexplored (Huang et al., 2021; Reeve & Tseng, 2011; Wang et al., 2019; Xu et al., 2020). It is within this context that the present study aims to examine how agentic engagement develops across a series of VR Germa-speaking scenarios, how different developmental trajectories can be characterized, and how key motivational and cognitive constructs help explain these patterns.

The purpose of this study was to investigate the development of agentic engagement in a VR game-based L2 German course, to identify distinct agentic engagement trajectories across repeated scenarios, and to explore how and to what extent self-efficacy, intrinsic motivation, classroom engagement, and cognitive load relate to and help explain these developmental patterns. To address this purpose, the study adopted an explanatory mixed-methods design in which quantitative analyses of longitudinal survey data was followed by qualitative interviews with purposively selected participants representing different agentic engagement trajectories. The core research questions focused on: (1) how agentic engagement develops over the intervention as a whole, (2) what distinct developmental trajectories can be identified and what facilitators or obstacles can be observed that may help explain those trajectories, and (3) to what extent self-efficacy, intrinsic motivation, classroom engagement, and cognitive load are implicated in those trajectories from both quantitative and qualitative perspectives.

The quantitative findings indicated that, at the group level, agentic engagement increased significantly across the intervention, with the largest gains occurring between the orientation session (S0) and the first game scenario (S1), followed by a more gradual rise and eventual plateau/slight decline in later sessions. Further statistical analyses revealed four distinct developmental trajectories—overall increase, relatively stable, fluctuating, and increase-then-decline—showing that learners did not follow a single uniform pattern of agentic engagement over time, although the majority of participants belonged to the "overall increase" category, explaining the general trend. Self-efficacy and classroom engagement emerged as significant predictors of agentic engagement, explaining unique variance in learners' agentic engagement levels and trajectories. Intrinsic motivation and cognitive load, on the other hand, did not show unique statistical effects once the other predictors were included.

The qualitative findings complemented and enriched these patterns by suggesting that confidence, task affordances, intrinsic motivation, cognitive load, and psychological safety shaped learners' experiences of agentic engagement in VR. Data suggested that growing confidence and repeated successful communication fostered agentic behaviors, that authentic and interactive tasks encouraged initiative, and that manageable cognitive demands supported experimentation. It also showed that technical issues, low or overly high challenge level, or limited interaction possibilities may have constrained participants' willingness to extend beyond minimal participation. Together, the mixed-methods results depict agentic engagement in this VR GBLL context as a dynamic, heterogeneous process that is shaped by the interplay of learner-related and contextual factors over time.

9.2. Main Contributions

Theoretical Contributions

The first main contribution of this thesis is theoretical, extending the concept of agentic engagement into the under-explored context of VR game-based language learning (Huang et al., 2021; Patall, 2024; Reeve, 2013; Reeve & Jang, 2022). Previous work on agentic engagement has largely focused on traditional classroom settings or general technology-enhanced environments, with little attention to fully immersive, game-like VR spaces where learners interact with both virtual environments and human partners (Bond & Bedenlier, 2019; Reeve & Tseng, 2011). By documenting how learners proactively try to shape tasks and interactions in a VR German-speaking setting, this thesis shows that core ideas from agentic engagement theory are applicable in VR GBLL, while also being modulated by unique affordances and constraints of immersion (Reeve, 2013; Reeve et al., 2022; Reinhardt, 2019).

A second theoretical contribution lies in providing a longitudinal perspective on agentic engagement, which has relatively rarely been studied as a process unfolding across multiple sessions (Patall, 2024; Reeve, 2013). The identification of four distinct trajectories demonstrates that agentic engagement is not a static learner trait but develops in diverse ways even when learners share the same instructional context (De Bot, 2017; Hiver et al., 2023).

This multi-trajectory pattern supports a dynamic understanding of engagement consistent with CDST, in which variability and non-linearity are treated as central features of development rather than statistical noise (Hiver et al., 2023; Larsen-Freeman, 2019).

Third, the thesis contributes to theorizing the multi-construct ecology of agentic engagement by clarifying how self-efficacy, intrinsic motivation, classroom engagement, and cognitive load jointly shape agentic engagement over time. The quantitative results position self-efficacy and classroom engagement as core correlates of agentic engagement, which aligns with Bandura's view of self-efficacy as a key determinant of effort and persistence, and also aligns with multidimensional views of engagement that integrate behavioral, emotional, cognitive, and agentic components (Bandura, 1994, 2001; Fredricks et al., 2004; Mameli & Passini, 2017; Reeve, 2013; Reeve & Tseng, 2011). The qualitative findings highlight intrinsic motivation and cognitive load as enabling and constraining conditions that influence when and how learners can translate their confidence and general engagement into agentic behaviors, which resonates with Self Determination Theory and Cognitive Load Theory (Deci & Ryan, 1985, 2000; Makransky & Petersen, 2021; Sweller, 1994; van Merriënboer & Sweller, 2005). Together these results suggest a layered model of AE in VR GBLL, in which multiple motivational and cognitive subsystems interact to produce diverse engagement trajectories.

Methodological Contributions

Methodologically, this thesis offers a concrete example of an explanatory mixed-methods design tailored to the study of engagement trajectories in VR language learning. By first analyzing longitudinal quantitative data and then using those results to guide purposive qualitative sampling and analysis, the study shows how mixed-methods designs can be used to both identify and explain complex developmental patterns in engagement. This approach adds to calls in engagement research for designs that move beyond cross-sectional snapshots toward more process-oriented, longitudinal understandings (Fredricks et al., 2016; Wong & Liem, 2022).

More specifically, the present paper demonstrates the value of trajectory-oriented quantitative techniques such as linear mixed models for capturing individual variability in engagement over time. These methods allowed the study to move beyond mean-level trends and to uncover more precise understandings of agentic engagement trajectories within a relatively small sample, illustrating how even small longitudinal datasets can show valuable process insights with the right analysis methods.

Finally, the study also illustrates a concrete way of combining established scales with context-specific interview protocols in VR research. This demonstrates how validated instruments from general educational psychology can be adapted and integrated into VR language learning research , while still leaving room for participants' own meanings and experiences to emerge.

Practical Contributions

The findings of this study have several implications for language teachers, VR designers, and curriculum planners who wish to foster agentic engagement in VR game-based L2 learning. For language teachers, the study suggests that designing tasks that feel authentic, meaningful, and slightly challenging, can support intrinsic motivation and create an environment in which students feel free and encouraged to act agentically (Deci & Ryan, 2000; Dörnyei, 2019; Reinhardt, 2019). Providing opportunities for learners to ask questions, make choices, and propose actions rather than just following scripted instructions, appears particularly important for sustaining agentic engagement over time (Reeve, 2013; Reeve & Tseng, 2011).

For VR and game designers, the results highlight the importance of task affordances and interactivity for agentic engagement. Scenarios that offer multiple ways to complete tasks, many opportunities for exploration, and responsive environments or characters seems more conducive to agentic engagement than tightly scripted low-interaction tasks (Gee, 2005; Huang et al., 2021; Hung et al., 2018). At the same time, designers need to balance challenge and cognitive load to ensure that the tasks are neither overwhelming nor boring, in line with principles from Cognitive Load Theory and immersive learning research (Makransky & Petersen, 2021; Poupard et al., 2025; Sweller, 1994; van Merriënboer & Sweller, 2005).

For curriculum designers and program leaders, this study underscores the value of creating psychologically safe environments in which learners feel comfortable experimenting, making mistakes, and trying out new language without fear of negative evaluation. In this way, the thesis presents concrete directions for designing VR-GBLL environments that motivate learners and cultivate their capacity and willingness to act agentically.

9.3. Closing statement

Overall, this thesis demonstrates that agentic engagement in VR game-based language learning is neither static nor uniformly experienced across learners, even when they share the same tasks, tools, and instructional setting. Instead, agentic engagement emerges as a dynamic, heterogeneous process that develops through the ongoing interaction between learner-related factors (such as self-efficacy, motivation, and prior experiences) and contextual factors (such as task design, affordances for interaction, cognitive demands, and psychological safety). Recognizing agentic engagement as a process rather than a fixed characteristic highlights the importance of designing VR GBLL environments that engage learners in the moment, and also provide sustained, meaningful opportunities for them to actively shape their own learning trajectories over time.

References

Acland, D. J. (2020). An Investigation of Flow Theory in an Online Game. *Review of Behavioral Economics, 7,* 317-336. DOI 10.1561/105.00000127

Andersen, M.S., Makransky, G. (2021). The Validation and Further Development of the Multidimensional Cognitive Load Scale for Physical and Online Lectures (MCLS-POL). *Frontiers in Psychology, 12*, 1-11. https://doi.org/10.3389/fpsyg.2021.642084

Astin, A. W. (1999). Student involvement: A Developmental Theory for Higher Education. *Journal of College Student Development, 40*(5), 518-529. (Original work published in 1984).

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of Human Behavior*, 4, 71-81. Academic Press.

Bandura, A. (2001). Social Cognitive Theory : An Agentic Perspective. *Annual Review of Psychology, 52*, 1-26.

Bandura, A. (2006a). Guide for Constructing Self-Efficacy Scales. *Self-Efficacy Beliefs of Adolescents,* 307-337. https://doi.org/10.1146/annurev.psych.52.1.1

Bandura, A. (2006b). Toward a Psychology of Human Agency. *Perspectives on Psychological Science, 1*(2), 164-180. https://www.jstor.org/stable/40212163

Baumeister, R. F., & Leary, M. R. (1995). The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. *Psychological Bulletin, 117*(3), 497-529. https://doi.org/10.1037/0033-2909.117.3.497

Bond, M., & Bedenlier, S. (2019). Facilitating Student Engagement Through Educational Technology: Towards a Conceptual Framework. *Journal of Interactive Media in Education, 2019*(1), 11, 1-14. https://doi.org/10.5334/jime.528

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping Research in Student Engagement and Educational Technology in Higher Education: A Systemic Evidence Map. *International Journal of Educational Technology in Higher Education, 17*(2), 1-30. https://doi.org/10.1186/s41239-019-0176-8

Buntins, K., Kerres, M., Heinemann, A. (2021). A Scoping Review of Research Instruments for Measuring Student Engagement in Technology Enhanced Learning in Higher Education. *International Journal of Educational Research*, *2*, 1-12. https://doi.org/10.1016/j.ijedro.2021.100099

Chen, J., Fu, Z., Liu, H., & Wang, J. (2024). Effectiveness of Virtual Reality on Learning Engagement: A Meta-Analysis. *International Journal of Web-Based Learning and Teaching Technologies, 19*(1), 1-14. https://doi.org/10.4018/IJWLTT.334849

Chiu, T. K. F., Lin, T-J., & Lonka, K. (2021). Motivating Online Learning: The Challenges of COVID-19 and Beyond. *Asia-Pacific Education Researcher, 30*(3), 187-190. https://doi.org/10.1007/s40299-021-00566-w

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper & Row.

De Bot, K. (2017) Complexity Theory and Dynamic Systems Theory: Same or Different ? In L. Ortega & Z. Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 51-58). John Benjamins. https://doi.org/10.1075/lllt.48

Deci, E. L. (1971). Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology, 18*(1), 105-115. https://doi.org/10.1037/h0030644

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin, 125*(6), 627-668. Https://doi.org/10.1037/00332909.125.6.627

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press.

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

Deci, E. L., & Ryan, R. M. (2012). Self-Determination Theory. In P. A. M Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (Vol. 1, pp. 416-437). SAGE Publications.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). https://doi.org/10.1145/2181037.2181040

Dörnyei, Z. (2019). Task Motivation: What Makes an L2 Task Engaging? In Z. (E.) Wen & M. J. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan* (pp. 53–66). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.13.04dor

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student Engagement, Context, and Adjustment: Addressing Definitional, Measurement, and Methodological Issues. *Learning and Instruction*, *43*, 1-4. https://doi.org/10.1016/j.learninstruc.2016.002

Garzón, J., Pavón, J., & Baldiris, S. (2019). Systematic Review and Meta-Analysis of Augmented Reality in Educational Settings. *Virtual Reality, 23*(4), 447–459. doi:10.1007/s10055-019-00379-9

Gee, J. P. (2003). *What Video Games Have to Teach Us About Learning and Literacy*. Palgrave Macmillan.

Gee, J. P. (2005). Learning by design: Good Video Games as Learning Machines. *E-Learning and Digital Media, 2*(1), 5–16. https://doi.org/10.2304/elea.2005.2.1.5

Hamilton, D., Mckechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive Virtual Reality as a Pedagogical Tool in Education: A Systematic Literature Review of Quantitative Learning Outcomes and Experimental Design. *Journal of Computer Eduation,*. *8*, 1-22. https://doi.org/10.1007/s40692-020-00169-2

Han, J., Zheng, Q., & Ding, Y. (2021). Lost in Virtual Reality? Cognitive Load in High Immersive VR Environments. *Journal of Advances in Information Technology, 12*(4), 302-310. doi: 10.12720/jait.12.4.302-310

Hiver, P., Larsen-Freeman, D., Al-Hoorie, A. H., & Lowie, W. (2023). Complex Dynamic Systems and Language Education: A Sampling of Current Research–Editorial. *International Journal of Complexity in Education, 4*(1), 1-8. DOI: https://doi.org/10.26262/ijce.v4i1.9476

Huang, X., Zou, D., Cheng, G., & Xie, H. (2021). A systematic review of AR and VR enhanced language learning. *Sustainability, 13*(9), 4639. doi:10.3390/su13094639

Hung, H.-T., Yang, J. C., Hwang, G.-J., Chu, H.-C., & Wang, C.-C. (2018). A Scoping Review of Research on Digital Game-Based Language Learning. *Computers & Education, 126*, 89–104. https://doi.org/10.1016/j.compedu.2018.07.001

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging Students in Learning Activities: It Is Not Autonomy Support or Structure but Autonomy Support and Structure. *Journal of Educational Psychology, 102*(3), 588-600. Https://doi.org/10.1037/a0019682

Jensen, L., & Konradsen, F. (2017). A Review of the Use of Virtual Reality Head-Mounted Displays in Education and Training. *Education and Information Technologies, 23*(4), 1515–1529. doi:10.1007/s10639-017-9676-0

Kahu, E. R. (2013). Framing Student Engagement in Higher Education. *Studies in Higher Education, 38*(5), 758-773. doi:10.1080/03075079.2011.598505

Kassab, S.E., Al-Eraky, M., El-Sayed, W., Hamdy, H., & Schmidt, H. (2023). Measurement of Student Engagement in Health Professions Education: A Review of Literature. *BMC Medical Education, 23*, 354. https://doi.org/10.1186/s12909-023-04344-8

Kearsley, G., & Shneiderman, B. (1998). Engagement Theory: A Framework for Technology-Based Teaching and Learning. *Educational Technology, 38*(5), 20–23. https://www.jstor.org/stable/44428478

Korhonen, V., Mattsson, M., Inkinen, M., & Toom, A. (2019). Understanding the Multidimensional Nature of Student Engagement During the First Year of Higher

Education. *Frontiers in Psychology*, *10*(1056), 15. https://doi.org/10.3389/fpsyg.2019.01056

Larsen-Freeman, D. (2019). On Language Learner Agency: A Complex Dynamic Systems Theory Perspective. *The Modern Language Journal, 103*(1), 61-79. https://doi.org/10.1111/modl.12536

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate Peripheral Participation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an Instrument for Measuring Different Types of Cognitive Load. *Behavior Research Methods, 45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Li, Y., & Lerner, R. M. (2013). Interrelations of Behavioral, Emotional, and Cognitive School Engagement in High School Students. *J. Youth Adolescence, 42*, 20-32. DOI 10.1007/s10964-012-9857-5

Li, J., & Xue, E. (2023). Dynamic Interaction between Student Learning Behaviour and Learning Environment: Meta-Analysis of Student Engagement and its Influencing Factors. *Behavioral Sciences*, *13*(1), 59. doi:10.3390/bs13010059 PMID:36661631

Li, J., van der Spek, E. D., Feijs, L. Wang, F., & Hu, J. (2017). Augmented Reality Games for Learning: A Literature Review. In N. Streitz & P. Markopoulos (Eds*.), Distributed, Ambient, and Pervasive Interactions: 5th International Conference, DAPI 2017, held as part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II* (pp. 612-626). Springer. DOI: 10.1007/978-3-319-58697-7_46

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding Immersive Virtual Reality to a Science Lab Simulation causes more Presence but less Learning. *Learning and Instruction, 60,* 225–236. doi:10.1016/j. learninstruc.2017.12.007

Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive

Virtual Reality. *Educational Psychology Review (2021), 33*, 937–958. https://doi.org/10.1007/s10648-020-09586-2

Mallek, F., Mazhar, T., Shah, S. F. A., Ghadi, Y. Y., & Hamam, H. (2024). A Review on Cultivating Effective Learning: Synthesizing Educational Theories and Virtual Reality for Enhanced Educational Experiences. *PeerJ Computer Science, 10*, e2000. https://doi.org/10.7717/peerj-cs.2000

Mameli, C., & Passini, S. (2017). Measuring Four-Dimensional Engagement in School: A Validation of the Student Engagement Scale and of the Agentic Engagement Scale. *TPM – Testing, Psychometrics, Methodology in Applied Psychology*, 24(4), 527-541. https://doi.org/10.4473/TPM24.4.4

Mameli, C., & Passini, S. (2019). Development and Validation of an Enlarged Version of the Student Agentic Engagement Scale. *Journal of Psychoeducational Assessment, 37*(4), 450–463. DOI: https://doi.org/10.1177/0734282918757849

Maples-Keller, J. L., Bunnell, B. E., Kim, S. J., & Rothbaum, B. O. (2017). The Use of Virtual Reality Technology in the Treatment of Anxiety and Other Psychiatric Disorders. *Harvard Review of Psychiatry, 25*(3), 103-113.

Martin, A. J. (2007). Examining a Multidimensional Model of Student Motivation and Engagement Using a Construct Validation Approach. *British Journal of Educational Psychology, 77*, 413-440. DOI:10.1348/000709906X118036

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the Effect of Pre-Training when Learning Through Immersive Virtual Reality and Video: A Media and Methods Experiment. *Computers & Education*, *140*. https://doi.org/10.1016/j.compedu.2019.103603

Nkomo, L. M., Daniel, B. K., & Butson, R. J. (2021). Synthesis of Student Engagement with Digital Technologies: A Systematic Review of the Literature. *International Journal of Educational Technology in Higher Education, 18*(1), 34. https://doi.org/10.1186/s41239-021-00270-1

Patall, E. A. (2024). Agentic Engagement: Transcending Passive Motivation. *Motivation Science*, 10(3), 222-233. https://doi.org/10.1037/mot000332

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan.

Poupard, M., Larrue, F., Sauzéon, H., & Tricot, A. (2025). A Systematic Review of Immersive Technologies for Education: Effects of Cognitive Load and Curiosity State on Learning Performance. *British Journal of Educational Technology, 56*(1), pp.5-41. https://doi.org/10.1111/bjet.13503

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A Motivational Model of Video Game Engagement. *Review of General Psychology, 14*(2), 154-166. https://doi.org/10.1037/a0019440

Raoofi, S., Tan, B. H., & Chan, S. H. (2012). Self-Efficacy in Second/Foreign Language Learning Contexts. *English Language Teaching, 5*(11), 60-73. http://dx.doi.org/10.5539/elt.v5n11p60

Reeve, J. (2013). How Students create Motivationally Supportive Learning Environments for Themselves: The Concept of Agentic Engagement. *Journal of Educational Psychology*, 105(3), 579-595. https://doi.org/10.1037/a0032690

Reeve, J., & Halusic, M. (2009). How K-12 Teachers Can Put Self-Determination Theory Principles into Practice. *Theory and Research in Education, 7*(2), 145-154. https://doi.org/10.1177/1477878509104319

Reeve, J., & Jang, H. (2022). Agentic Engagement. In A. L. Reschly & S. L. Christenson (Eds.), *Handbook of Research on Student Engagement* (pp. 95–107). Springer.

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing Students' Engagement by Increasing Teachers' Autonomy Support. *Motivation and Emotion, 28*(2), 147-169. https://doi.org/10.1023/B:MOEM.0000032312.95499.6f

Reeve, J., Jang, H.-R., Shin, S. H., Ahn, J. S., Matos, L., & Gargurevich, R. (2022). When Students Show Some Initiative: Two Experiments on the Benefit of Greater Agentic Engagement. *Learning and Instruction*, 80, 101564. https://doi.org/10.1016/j.learninstruc.2021.101564

Reeve, J., & Tseng, C. M. (2011). Agency as a Fourth Aspect of Students' Engagement During Learning Activities. *Contemporary Educational Psychology, 36*(4), 257-267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Rigby, C. S., & Ryan, R. M. (2011). *Glued to Games: How Video Games Draw Us in and Hold Us Spellbound*. Praeger.

Reinhardt, J. (2019). *Gameful Second and Foreign Language Teaching and Learning: Theory, Research, and Practice.* Palgrave-Macmillan.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist, 55*(1), 68-78. https://doi.org/10.1037//0003-D66X.55.1.68

Ryan, R. M., & Deci, E. L. (2017). Self-determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. Guilford Publications. https://www.guilford.com/books/Self-Determination_theory/Ryan-Deci/9781462528769

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and Extrinsic Motivation from a Self-Determination Theory Perspective: Definitions, Theory, Practices, and Future Directions. *Contemporary Educational Psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion, 30*, 344-360. DOI 10.1007/s11031-006-9051-8

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student Engagement in High School Classrooms from the Perspective of Flow Theory. *School Psychology Quarterly, 18(*2), 158–176. https://doi.org/10.1521/scpq.18.2.158.21860

Sikandar, M. H., Mao, M.e., Riaz, A. (2025). The Role of Motivation in Second Language Learning: A Systematic Review of Empirical Studies. *Pakistan Social Sciences Review, 9*(2), 322–336. https://doi.org/10.35484/pssr.2025(9-II)24

Sweller, J. (1994). Cognitive Load Theory, Learning Difficulty, and Instructional Design. *Learning and Instruction*, *4*, 295-312. https://doi.org/10.1016/0959-4752(94)90003-5

Thi, L. D. (2025). Game-Based Learning in Improving Student Engagement. *International Journal of Research and Innovation in Social Science (IJRISS)*, *9*(5), 5116–5123. https://dx.doi.org/10.47772/IJRISS.2025.905000397

Thorne, S. L., Black, R. W., & Sykes, J. M. (2009). Second Language Use, Socialization, and Learning in Internet Interest Communities and Online Gaming. *The Modern Language Journal, 93*(S1), 802–821. https://doi.org/10.1111/j.1540-4781.2009.00974.x

Van Merriënboer, J. J. G., & Sweller, J. 2005. Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review, 17*(2), 147-178. https://doi.org/10.1007/s10648-005-3951-0.

Vansteekiste, M., & Ryan, R. M. (2013). On Psychological Growth and Vulnerability: Basic Psychological Need Satisfaction and Need Frustration as a Unifying Principle. *Journal of Psychotherapy Integration, 23*(3), 263-280. DOI: 10.1037/a0032359

Vansteenkiste, M., Ryan, R .M., & Soenens, B. (2020). Basic Psychological Need Theory: Advancements, Critical Themes, and Future Directions. *Motivation and Emotion, 44*, 1-31. https://doi.org/10.1007/s11031-019-09818-1

Wang, C-P., Lan, Y-J., Tseng, W-T., Lin, Y-T., & Gupta, K. C-L. (2019). On the Effects of 3D Virtual Worlds in Language Learning: A Meta-Analysis. *Computer Assisted Learning, 33*(1), 1-25. https://doi.org/10.1080/09588221.2019.1598444

Wang, D., Peng, Y., Haddouk, L., Vayatis, N., Vidal, P-P. (2025). Assessing Virtual Reality Presence through Physiological Measures: A Comprehensive Review. *Frontiers in Virtual Reality*, *6,* 1530770. https://doi.org/10.3389/frvir.2025.1530770

Wang, Z., Bergin, C., & Bergin, D. A. (2014). Measuring Engagement in Fourth to Twelfth Grade Classrooms: The Classroom Engagement Inventory. *School Psychology Quarterly*, *29*(4), 517-535. https://doi.org/10.1037/spq0000050

Wong, Z. Y., & Liem, G. A. D. (2022). Student Engagement: Current State of the Construct, Conceptual Refinement, and Future Research Directions. *Educational Psychology Review*, *34*(1), 107–138. https:// doi.org/10.1007/s10648-021-09628-3

Wu, X. (2022). Motivation in Second Language Acquisition: A Bibliometric Analysis between 2000 and 2021. *Frontiers in Psychology, 13*, 1032316. doi: 10.3389/fpsyg.2022.1032316

Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of Immersive Virtual Reality using Head-Mounted Displays on Learning Performance: A Meta-Analysis. *British Journal of Educational Technology, 51*(6), 1991–2005. doi:10.1111/bjet.13023

Yang, J. C., Chen, C. H., & Chang Jeng, M. (2010). Integrating Video-Capture Virtual Reality Technology into a Physically Interactive Learning Environment for English Learning. *Computers & Education, 55*(3), 1346–1356. doi:10.1016/j.compedu.2010.06.005

Yang, K., & Gan, Z. (2024). A Systematic Review of Reading Self-Efficacy in Second or Foreign Language Learning. *System, 125*, 103436. https://doi.org/10.1016/j.system.2024.103436

Yousefi, M., and Mahmoodi, M. H. (2022). The L2 Motivational Self-System: A Meta-Analysis Approach. *International Journal of Applie. Linguisics*. *32*, 274–294. doi: 10.1111/ ijal.12416

Yu, Z. (2021). A Meta-Analysis of the Effect of Virtual Reality Technology Use in Education. *Interactive Learning Environments, 31*, 4956-4976. doi:10.1080/10494820.2021.1989466

Xu, Z., Chen, Z., Eutsler, L., Geng, Z., & Kogut, A. (2020). A Scoping Review of Digital Game-Based Technology on English Language Learning. *Educational Technology Research and Development*, *68*(3), 877–904. https://www.jstor.org/stable/48727469

Xu, X., Shi, Z. Bos, N. A., & Wu, H. (2023). Student Engagement and Learning Outcomes: An Empirical Study Applying a Four-Dimensional Framework. *Medical Education Online, 28*(1), 2268347. https://doi.org/10.1080/10872981.2023.2268347

Zareian, G., & Jodaei, H. (2015). Motivation in Second Language Acquisition: A State of the Art Article. *International Journal of Social Sciences & Education,5*(2), 295-308.