Leveraging Artificial Intelligence in Digital Game-Based Learning to Enhance Medical Diagnostic Skills

Vladislav Ilin

# Abstract

Diagnostic errors significantly impact patient outcomes, yet training doctors through the apprenticeship model is challenging due to time constraints, curriculum design, and other contributing factors. This project proposes the development of a serious game aimed at improving medical diagnostic skills for New Zealand’s second-year doctors (PGY2 ). The intended game would scaffold learning and facilitate flow through engaging narratives, aligning educational objectives with gameplay immersion. The work examines the core components of traditional game design in contrast to the principles of **Digital Game-Based Learning (DGBL)**. A key challenge in this domain arises from the tension between games designed primarily for entertainment and those focused on achieving educational outcomes. In DGBL, this often leads to a dichotomy where player engagement may not necessarily translate to effective learning, while learning-focused activities may reduce player immersion. However, **Large Language Models (LLMs)** offer a promising solution by integrating real-time feedback and scaffolding to enhance learning and enjoyment.

# 1. Problem Statement

Training doctors to diagnose accurately is critical and challenging ([Koelewijn et a](#vu5e4gftw2fj)l., 2024; [Morgan et al](#178sdhhxyv91)., 2024; [Olson & Graber](#w748xjgthryp), 2020; [Ten Cate & Scheele](#kpjmw3x6djw6), 2007). Diagnostic errors often stem from medical students struggling to transition from theoretical knowledge into actual practice ([Yanagita et al](#hz0ne36czvvr)., 2023; [Hawkins et al](#n14ccpsr2yrf)., 2021; [Monrouxe et al](#2fph1973xpcx)., 2018; [Picton et al](#n4aqhcib2uwb)., 2022). Diagnostic errors, compounded by time constraints ([Scott](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.sqk3sbq60i92), 2007) and curriculum design issues ([Graber](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.mkr45ahcm0vx) et al., 2022), are a global problem that severely impacts patient outcomes ([Cresswell et al](#l5jbxg6oy4jl)., 2013; [National Academies of Sciences, Engineering, and Medicine](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c4t56r83m5y5), 2015; [Singh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.f2uj0yinmlt1)., 2017; [Makary & Daniel](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.7lol3b5v20px), 2016).

While serious games have demonstrated potential as training tools in the medical field ([Sharifzadeh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c1pm6ui9yn72)., 2020; [Wang et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.xdur2vqy6pd8)., 2022; [Graafland et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.l2a2s7raouyo)., 2012; [Xu et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.x0vmy6da6c4o)., 2023; [Abd-alrazaq et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.uov0sxmedypz)., 2022; [Do et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.shxm4cadhfu4)., 2023; [Bijl](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.9wfh7zrpasxk) et al., 2024), they remain underutilized due to the demands and costs of development, as well as varying requirements and practices across institutions ([Gorbanev et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.aqbjc259voud)., 2018; [Olgers et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.2rdsbupefuzr)., 2021).

I propose Contra-Indication, a serious game designed to enhance diagnostic reasoning. Through immersive, case-based scenarios that simulate real-world complexity and time-sensitive decision-making, Contra-Indication aims to improve diagnostic accuracy while keeping learners engaged. The project proposes to leverage AI-driven large language models (LLMs) to provide real-time personalized feedback and dynamically adapt to individual learning needs. By integrating digital game-based learning (DGBL) with LLMs, this project presents an innovative approach to reducing diagnostic errors and improving patient outcomes.

## 1.1 What: Diagnostic Errors

Accurate medical diagnosis is central to healthcare, as it directs treatment decisions and ensures patient safety ([Olson & Graber](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.d2zqmp6rjh3p), 2020). Diagnostic errors, which include missed, incorrect, or delayed diagnoses, occur when a timely and accurate explanation of a patient's health issue is not established or communicated effectively ([National Academies of Sciences, Engineering, and Medicine](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c4t56r83m5y5), 2015). These errors can result in adverse health outcomes, psychological distress, and financial costs.

While New Zealand lacks comprehensive national statistics on misdiagnosis, data from the Accident Compensation Corporation (ACC) indicates cancer misdiagnosis has nearly doubled in five years ([Cook,](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.o1yk7efz6nw1) 2022). This trend mirrors increasing diagnosis problems in Western healthcare systems ([Cresswell et al](#l5jbxg6oy4jl)., 2013). In the United States, diagnostic errors are the leading cause of malpractice claims across medical specialties ([Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.m80xzoy37ulg)., 2018). Studies show that annually, one in twenty patients in ambulatory settings experiences a diagnostic error ([Singh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.f2uj0yinmlt1)., 2017), and an estimated 40,000 to 80,000 hospitalized patients die due to diagnostic mistakes ([Makary & Daniel](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.7lol3b5v20px), 2016). Similarly, in the UK, since 2020, negligence claims have been rising**.** In 2024 alone, the NHS paid £2.8 billionin clinical negligence claims, with over 25% allocated to legal expenses ([NHS Resolutions](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.6ogh58jylbct), 2024).

The causes of diagnostic errors are multifaceted, as the diagnostic process is inherently uncertain, complicated, and nuanced ([Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.m80xzoy37ulg)., 2018), with multiple valid diagnostic paths often existing for a single case ([Olson & Graber](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.d2zqmp6rjh3p), 2020). Infections, trauma, and cancer are among the top areas for error as these are notoriously difficult to diagnose, as symptoms often appear as something else. Without appropriate training and knowledge, doctors can easily misdiagnose the underlying causes ([Singh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.f2uj0yinmlt1)., 2017). One of the contributing factors is the decline in basic clinical skills among medical students and residents ([Faustinella & Jacobs](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.cill1zg0l52q), 2018). This decline stems from deficiencies in early medical education and extends into continuing professional development for practicing physicians ([Singh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.f2uj0yinmlt1)., 2017). For instance, [Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.m80xzoy37ulg)., (2018) reveal that 84% of clerkship directors find students entering clinical clerkships with inadequate or only a fair grasp of key clinical reasoning concepts, while 57% highlight a lack of formal training on these critical skills during the preclinical years.. Given the profound impact of diagnostic errors, improving medical diagnosis accuracy requires innovative education approaches, particularly in training environments that reflect real-world complexity and uncertainty.

A small preliminary research project was performed to address this gap. The study got ethical approval both by the Capital & Coast District Health Board (ID: 2025-CCHV-390, 17-02-2025) and VUW Research Ethics approval 2024/HE000359. As part of the preliminary study, interviews were conducted with two specialists: Dr. Brad Peckler, an Emergency Medicine Specialist and President of the New Zealand Association for Simulation in Healthcare (NZASH), and Dr. George Bax, an Intensive Care Unit Specialist and Senior Clinical Lecturer. Their insights highlighted critical gaps in medical training that contribute to diagnostic errors, including incorrect drug prescriptions, poor consultation management, and unnecessary testing—factors that can delay treatment or cause patient harm. Based on the interviews and the literature review five learning outcomes were identified to be integrated into Contra-Indication: clinical thinking, task prioritization, medical knowledge, policy comprehension, and communication.

## 1.2 Who: Junior Doctors

This project proposes the development of a serious game to improve diagnostic skills among second year (PGY2) doctors, who will be referred to as “**junior doctors**” in this proposal. According to [Bindra et al](#oqi2ecepi397)., 2022, in New Zealand, becoming a doctor typically involves completing a six-year Bachelor of Medicine program focusing on foundational medical sciences and clinical skills. After graduation, medical students enter a two-year internship (PGY1 & PGY2), working under supervision in various specialties. Upon completing this phase, they receive general registration with the Medical Council of New Zealand (MCNZ) and choose between pursuing general practice or specialist training.

During PGY1, junior doctors typically work in small teams, shadowing senior colleagues and performing clerical tasks. By PGY2, they are expected to diagnose, prescribe medicine, and manage patients more independently. However, according to Dr. Peckler and Dr. Bax, this transition often proves challenging as increased responsibility, coupled with varying levels of supervision, can lead to decision-making difficulties that result in clinical errors. Their observations are supported by findings from [Bindra et al](#oqi2ecepi397)., 2022, who surveyed junior doctors in New Zealand, concluding that many junior doctors felt underprepared and undertrained for the challenges of real world medicine.

Since New Zealand is relatively small, few academic studies have applied DGBL to train junior doctors. However, New Zealand's apprenticeship structure closely mirrors the UK and the US, allowing us to draw parallels between the academic research in these medical systems. The UK's system of first and second-year doctors (FY1 & FY2) follow a similar trajectory, with FY1 doctors primarily observing and assisting, while FY2 doctors take on greater clinical responsibility before specializing ([Monrouxe et al](#2fph1973xpcx)., 2018). Likewise, in the US, medical graduates complete a one-year internship before entering residency, during which their clinical independence progressively increases ([Crimson Education NZ](#1xvinnsn2rkz), 2023). Given these similarities, insights from UK and US studies inform on training gaps and strategies to improve diagnostic skills for junior doctors in New Zealand.

Despite its importance, medical education often fails to prepare students to diagnose accurately. Research highlights cognitive and system-related training gaps, clinical reasoning, communication, and teamwork ([Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.mkr45ahcm0vx)., 2022). Addressing these deficiencies through innovative, real-world training approaches is crucial to improving diagnostic accuracy and patient outcomes ([Ruedinger et al.](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.b70z7vutw6gt), 2017; [Ten Cate & Scheele](#kpjmw3x6djw6), 2007; [Monrouxe et al](#2fph1973xpcx)., 2018).

## 1.3 Why: Gaps in Medical Training

### 1.3.1 Challenges in Curriculum and Assessment

A significant barrier to developing diagnostic skills is the absence of curricula and assessments focused on diagnostic reasoning ([Olson & Graber](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.d2zqmp6rjh3p), 2020; [Ruedinger et al.](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.b70z7vutw6gt), 2017; [Rencic et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.jc98guylt63f)., 2017;). Most medical training follows the apprenticeship model, where students acquire diagnostic skills through clinical experience rather than formal instruction ([Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.mkr45ahcm0vx)., 2022; [Bindra et al](#oqi2ecepi397)., 2022; [Ten Cate & Scheele](#kpjmw3x6djw6), 2007). For instance, many U.S. medical schools do not offer dedicated courses on diagnosis, instead relying on clinical rotations to fill this gap ([Rencic et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.jc98guylt63f)., 2017). However, experts argue that fundamental diagnostic concepts, such as the threshold approach to clinical decision-making ([Pauker & Kassirer](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.14x4nns7mlz7), 1980), medical test interpretation ([Morgan et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.r510rahbdioa)., 2024), and recognizing cognitive biases ([Reilly et al](#ummdmqc5ckpi)., 2013; [Tversky & Kahneman](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.bcokq8o4yyw5), 1974) should be taught before students enter clinical practice ([Bindra et al](#oqi2ecepi397)., 2022; [Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.mkr45ahcm0vx)., 2022; [Ruedinger et al.](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.b70z7vutw6gt), 2017; [National Academies of Sciences, Engineering, and Medicine](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c4t56r83m5y5), 2015; [Rencic et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.jc98guylt63f)., 2017).

Dr. Peckler (2025) and Dr. Bax (2025) noted that students transitioning into junior doctor roles frequently struggle with essential skills, including effective communication with colleagues and patients ([Bindra et al](#oqi2ecepi397)., 2022; [Monrouxe et al](#2fph1973xpcx)., 2018; [Hawkins et al](#n14ccpsr2yrf)., 2021). These skills are difficult to teach in academic settings because their application varies across institutions and specializations ([Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.m80xzoy37ulg)., 2018). For instance, theoretical knowledge learned in medical school often does not translate well to the high-pressure environment of an emergency department, where factors like time constraints and cognitive biases complicate decision-making. Junior doctors are expected to develop these skills during their apprenticeship. Still, this model has limitations, including inconsistent training due to reliance on individual mentors ([Wimmers et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.qfgusntose91)., 2006) where feedback is often reactive, focusing on errors rather than proactively guiding structured learning (Peckler, 2025).

Dr. Peckler (2025) emphasized fostering critical thinking, an essential part of clinical thinking. During their academic studies, many assessments, such as multiple-choice tests, prioritize content knowledge and single-best answers, failing to reflect the complexity of real-world situations ([Graber et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.mkr45ahcm0vx)., 2022; [Ten Cate & Scheele](#kpjmw3x6djw6), 2007). In contrast, serious games offers flexibility that allows for process-based assessment, focusing on how students develop and refine differential diagnoses, prioritize information, and adapt their reasoning ([Koelewijn et a](#vu5e4gftw2fj)l., 2024; [Chon et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.ocmi3qgf3ix)., 2019). By shifting from static knowledge recall to clinical thinking, serious games mightbetter prepare junior doctors for real-life medical challenges.

Serious games' adaptability to learning objectives and individual needs makes it a compelling tool for improving curriculum and assessment. Case studies of serious games boast enhanced skill development ([Buajeeb et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.x6z8feeuvhxz)., 2023) and sustained learner motivation ([Zairi et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.nhegzt5fiyyo)., 2022). Research on clinical diagnostic support systems ([Yanagita et al](#hz0ne36czvvr)., 2023) and game-based learning platforms ([Sharifzadeh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c1pm6ui9yn72)., 2020) claim to improve diagnostic accuracy among medical students. Furthermore, patient dialogue simulation has proven effective in honing essential medical communication skills ([Atkins](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.mh1vcljug5vt), 2018). By shifting the focus of training from passive content recall to active, process-based learning, *Contra-Indication* has the potential to bridge the gaps in current medical training. It not only supports the development of diagnostic reasoning but also prepares junior doctors for the dynamic, high-pressure environments they will face in clinical practice, offering a promising pathway to enhancing both competence and confidence in medical professionals.

### 1.3.2 Time Constraints in Medical Training

Time constraints significantly hinder students’ ability to develop diagnostic accuracy, a critical skill for medical practice ([Singh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.f2uj0yinmlt1)., 2017). Academic research underscores the importance of frequent, high-quality supervision to enhance clinical reasoning and decision-making ([Wimmers et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.qfgusntose91)., 2006; [Ten Cate & Scheele](#kpjmw3x6djw6), 2007; [Monrouxe et al](#2fph1973xpcx)., 2018). However, the time pressures of clinical schedules often limit opportunities for students to engage in the in-depth learning needed to refine diagnostic skills ([Scott](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.sqk3sbq60i92), 2007). As Dr. Peckler (2025) noted, junior doctors experience these limitations most acutely. During their first year, they primarily shadow senior physicians and work with pre-diagnosed cases, which limits their opportunities to develop independent diagnostic reasoning. By their second year, however, they transition into high-pressure environments such as the emergency department, where they are expected to diagnose, prescribe, and treat patients with minimal direct oversight (Peckler, 2025; Bax, 2025). This sudden transition can be overwhelming and underscores the need for structured, proactive learning strategies that provide sufficient diagnostic practice before assuming full responsibility ([Bindra et al](#oqi2ecepi397)., 2022; [Milwood](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.tzq5rjb25q29), 2014; [Picton et al](#n4aqhcib2uwb)., 2022; [Rencic et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.jc98guylt63f)., 2017; [Reilly et al](#ummdmqc5ckpi)., 2013; [Ten Cate & Scheele](#kpjmw3x6djw6), 2007; [Monrouxe et al](#2fph1973xpcx)., 2018; [Hawkins et al](#n14ccpsr2yrf)., 2021).

Simulation-based training has long been recognized as an effective method for teaching skills rarely encountered in clinical practice or cannot be safely practiced on actual patients ([Koelewijn et a](#vu5e4gftw2fj)l[.](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.jxqhaq4zsenv), 2024; [Chon et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.ocmi3qgf3ix)., 2019). DGBL expands on this approach by offering interactive, immersive environments replicating the complexity of real clinical settings. Through DGBL, medical students can engage in diagnostic exercises that simulate experiences that are often difficult to obtain within the constraints of traditional training models.

Another key area that is affected by time constraints is communication. Dr. Peckler and Dr. Bax highlighted cases where junior doctors struggled with communication, often providing excessive or disorganized information to colleagues. Effective communication is especially critical in high-pressure settings like the emergency department and intensive care unit, where timely and accurate information can directly impact patient outcomes. Due to inexperience and a tendency to be overly cautious, junior doctors often over-explain rather than distill patient information into the most relevant details (Peckler, 2025; Bax, 2025). Talking with patients and their families is another essential skill ([National Academies of Sciences, Engineering, and Medicine](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c4t56r83m5y5), 2015). Dr. Bax (2025) explained that while junior doctors attend lectures on best communication practices, practicing having actual patient conversations would solidify their learning. DGBL platforms can simulate doctor-to-doctor and doctor-to-patient interactions, preparing them for real-life clinical demands.

By integrating structured, scenario-based learning, DGBL can allow students to practice diagnosing and managing complex cases in a risk-free setting. This approach mitigates the impact of time constraints and fosters essential skills such as task prioritization, clinical knowledge application, adherence to protocols, and effective communication—learning outcomes that Dr. Peckler (2025) emphasized as crucial for junior doctors. By bridging the gap between theoretical learning and real-world application, *Contra-Indication* has the potential to address the time constraints of medical training and ensure that students are better prepared to navigate the complexities of medical practice.

### 1.3.3 Fear of Mistakes

Beyond time constraints, anxiety about errors often hinders the development of diagnostic skills in medical students. Junior doctors frequently struggle with a profound fear of mistakes, which can slow their diagnostic growth ([Millwood](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.tzq5rjb25q29), 2014). This fear is intensified by the high stakes of real-world medical decision-making, where errors can have serious consequences. Dr. Bax (2025) highlighted that junior doctors often fear asking for help and may fail to escalate critical issues when they are out of their depth. Compounding the problem, students rarely encounter structured opportunities to make diagnostic errors in low-pressure environments designed for learning ([Picton et al](#n4aqhcib2uwb)., 2022). Their training consists of cautionary tales and case studies. When mistakes are made, they are discussed privately or via a tribunal if the error leads to serious consequences (Bax, 2025). As a result, many hesitate to make independent decisions, frequently delaying diagnoses or seeking unnecessary confirmation from senior colleagues—behaviors that can disrupt workflow efficiency and compromise patient care (Peckler, 2025).

One of the benefits of DGBL is that it can provide a safe and controlled environment for learners to practice without fearing real-world repercussions ([Bolstad & McDowall](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.y9jpr1o57qxg), 2019). This method promises to bridge a gap in medical education by encouraging students to learn from errors and view them as part of cognitive growth ([Reilly et al](#ummdmqc5ckpi)., 2013; [Singh et al](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.f2uj0yinmlt1)., 2017) and promote a culture where mistakes are embraced as opportunities for development ([Milwood](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.tzq5rjb25q29), 2014). Mitigating the fear of errors also reduces stress ([Hewitt & Stubbs](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.sw6l3trl9sx4), 2017) and creates a more productive learning experience ([Slussareff et a](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.1sclu5xlu31w)l., 2016). The potential benefits of DGBL align with guidelines to improve diagnostic training from World Health Organization ([Cresswell et al](#l5jbxg6oy4jl)., 2013) and the US [National Academies of Sciences, Engineering, and Medicine](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c4t56r83m5y5), (2015).

Improving the diagnostic process is not just possible but essential. It is a moral, professional, and public health imperative ([Cresswell et al](#l5jbxg6oy4jl)., 2013; [National Academies of Sciences, Engineering, and Medicine](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.c4t56r83m5y5), 2015). By addressing key gaps in medical education, including curriculum deficiencies, time constraints, and the fear of making mistakes, DGBL may offer a promising solution to improving diagnostic training. By integrating realistic and contextually relevant simulations, *Contra-Indication* mightenhance diagnostic accuracy while helping students develop the confidence and resilience required for high-pressure clinical environments. However, further investigation is needed to understand the impact and effectiveness of DGBL in these areas.

# 

# 2. Games

A fundamental part of this process involves understanding the key components that contribute to the success of a game. [Salen﻿ & ﻿Zimmerman](#8ume7880vgu0) (2004, p. 83) succinctly define games as “a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome.” Although not all games fit neatly within this definition, understanding these foundational elements are essential to develop successful games.

## 2.1 Game Mechanics

Game mechanics refer to the formal systems that define rules and gameplay. Rules delineate what players can and cannot do. According to [Plass et al](#kix.8wuxyrxwbr40). (2011, p. 11), "game mechanics are therefore a means to guide the player into particular behaviors by constraining the space of possible plans to attain goals."

## 2.2 Engagement

Engagement is crucial for a game's success. Players will likely continue playing only if they find the game enjoyable ([Moore](#6taarrlg1t5r), 2016; [Fullerton](#l2l1gr2giok6), 2024). [Schuytema](#emu6006fxkf5) (2007, p. 8) identifies four elements of enjoyment:

* **Receptiveness** refers to the player's mindset, as a distracted or frustrated individual is unlikely to enjoy playing.
* **Expectations** represent the balance between previous experiences and the current game; meeting or surpassing expectations leads to enjoyment.
* **Personal preference**s are entirely subjective yet play a significant role in shaping which games players enjoy.
* **Flow** denotes the balance between the game’s challenge and the player’s skill and is essential to engagement. A game that is too easy causes boredom, while one that is too difficult results in frustration ([Csikszentmihalyi](#6hijuk3kfsuj), 2014). An individual in a flow state is fully engaged and may lose track of time or ignore physical needs.

While receptiveness and preferences are beyond the designer's control, meeting expectations and maintaining the difficulty balance are crucial aspects of game design. Section 5.4 delves deeper into Contra Indication game mechanics and how these will enhance learning outcomes and player experience.

## 2.3 Conflict and Victory

In many games, conflict is central to the experience. Achieving objectives results in a quantifiable outcome, measuring and often comparing the player's progress to other scores. In such games, meeting goals is integral to attaining victory (e.g., solving a puzzle, racing, capturing the flag, defeating the boss). Opponents can be other players, computers, a previous score set by the player, or the clock. However, conflict or victory is not essential to all games. For example, many games incorporate simulations, world building or interactive stories, breaking away from the traditional win-lose paradigm ([Tsai & Tsai,](#7j6dpk8l508d) 2020).

## 2.4 Video Games

Video games are a sub-genre of games characterized by input/output information (e.g., display, control modes), multiplayer aspects (e.g., nature of the co-player, online gaming techniques), and in-game content (e.g., narrative, challenge) ([Caroux](#k8eo07g11gvu), 2023). Video games are played on electronic devices such as computers, consoles, or mobile devices. The platform type determines input/output mechanisms (e.g., mouse and keyboard vs. console controller) and influences the game's complexity and graphical demands (e.g., disk storage and video card capabilities). Most electronic game devices can connect to the internet, enabling multiplayer scenarios that are impossible in analog games.

Video games as a genre are incredibly successful. For example, in 2022, "League of Legends" had 13,687,409 players (Demirkol, 2024). The global commercial video game industry reached $227.6 billion in 2023 (Ballhaus et al., 2024). At their best, video games generate engagement and interactivity at a level unparalleled by other media. Given the success of video games as entertainment, their application as learning tools holds immense potential.

## 2.5 Digital game-based learning (DGBL)

DGBL integrates gameplay with educational objectives to enhance academic performance (Hu et al., 2024). Its primary goal is achieving specific learning goals (Mayer, 2014), not entertainment (Clark et al., 2016). This distinction poses a challenge because traditional games are solely for entertainment, whereas educational tools are only sometimes perceived as engaging. However, this does not imply that games designed for learning must be dull; many effectively combine learning with engagement (Bolstad & McDowall, 2019).

DGBL has been extensively integrated into various educational fields, including science ([Tsai & Tsai](#7j6dpk8l508d), 2020), mathematics ([Fadda et al](#awx1i2bcrcuu)., 2022), STEM ([Wang](#ql2qopminbu5) et al. et al., 2022, [Gui et al](#ep5u464ko7jn)., 2023), English language instruction ([Thompson & Gillern](#gk45n8h5fae6), 2020) as well as specialized areas like language acquisition and software engineering ([Cai et al](#p8h4ri9k4ar6)., 2022). Numerous studies have examined the influence of specific game elements, including narrative ([Jackson et al](#kxiap3awy6qk)., 2018), gamification ([Majuri et al](#1tjbx8gslrj)., 2018), competition ([Chen et al.](#4sfxlcvr9ow4), 2020), and the comparative effects of game versus non-game learning environments ([Clark et al](#1qu9pocyv90d)., 2016). These studies have examined how these elements impact creativity ([Blanco-Herrera et al](https://docs.google.com/document/d/1bvdoQnTjfNV3XI7Avggb26F0XlVG66pTFFiqeRzW5cQ/edit#bookmark=id.4gyn8u38jh0r)., 2019), cognitive abilities ([Sauce et al](https://docs.google.com/document/d/1bvdoQnTjfNV3XI7Avggb26F0XlVG66pTFFiqeRzW5cQ/edit#bookmark=id.ghl4kxdohfj)., 2022; [Chaarani et al.](https://docs.google.com/document/d/1bvdoQnTjfNV3XI7Avggb26F0XlVG66pTFFiqeRzW5cQ/edit#bookmark=id.bz3d7cat2g2), 2022), problem-solving skills, and academic performance ([Adachi & Willoughby](https://docs.google.com/document/d/1bvdoQnTjfNV3XI7Avggb26F0XlVG66pTFFiqeRzW5cQ/edit#bookmark=id.byuw8twu7pr9), 2013; [Bejjanki et al](https://docs.google.com/document/d/1bvdoQnTjfNV3XI7Avggb26F0XlVG66pTFFiqeRzW5cQ/edit#bookmark=id.2b2hpwgqsx5c)., 2014), as well as student engagement, motivation, and concentration ([Cai et al](#p8h4ri9k4ar6)., 2022; [Lei et al.,](#kix.t4ihyav9xkd9) 2022).

### 2.5.1 Serious Games vs. Gamification

[Landers](#kix.qmahiul2npbz) (2014) differentiates DGBL into two categories: serious games and gamification. Serious games are designed with the explicit goal of developing specific skills or knowledge. For example, flight simulators are serious games that train pilot competencies. In contrast, gamification incorporates game elements into non-game contexts to influence behaviors that enhance learning outcomes ([Majuri et al](#1tjbx8gslrj)., 2018). While gamification does not directly target learning, it aims to modify behaviors that, in turn, improve educational results. Therefore, effective gamification design is challenging as it must ensure that the targeted behaviors positively impact learning ([Landers](#kix.qmahiul2npbz), 2014).

## 2.6 DGBL Challenges

Despite the academic attention given to educational game development, poor game design remains one of the primary factors behind the failure of educational games ([Wang et al](#ql2qopminbu5)., 2020; [Tsai & Tsai,](#7j6dpk8l508d) 2020; [Gui et al](#ep5u464ko7jn)., 2023; [Clark et al.](#1qu9pocyv90d), 2016; [Hu et al](#ivq88tcnkm9a)., 2024). The design and evaluation of DGBL tools present considerable challenges due to the inherent complexity of video game development. The creation of video games necessitates diverse skills, including programming, graphic design, narrative development, and user experience design ([Schuytema, 2007](#emu6006fxkf5)). These competencies are particularly crucial in academic research focused on DGBL effectiveness. Numerous studies have documented instances where educational games have failed to achieve their intended outcomes due to deficiencies in design. For example, [Jackson et al.](#kxiap3awy6qk) (2018) conducted a meta-review of 130 narrative-driven game studies, revealing that nearly half of the research teams lacked professional game designers, and only two studies employed dedicated writers. The authors emphasize the importance of possessing appropriate skill sets during development.

Moreover, DGBL tools are designed to achieve specific educational outcomes ([Slussareff et al](#ff9tv2p48w61)., 2016). These outcomes can vary widely depending on the academic goals, encompassing satisfaction, affective responses, cognitive processes, effort, attitudes, social interactions, psychological states or traits, performance, engagement, and physical activity ([Majuri et al.](#1tjbx8gslrj), 2018). [Dörner et al](#nztyxq9r4e4z). (2016) categorized these outcomes into three primary groups: cognitive (e.g., knowledge acquisition), affective (e.g., engagement, motivation, enjoyment), and skill-based. Therefore, effectively designing and evaluating DGBL tools is critical to ensure they meet the intended educational objectives.

In serious games, measuring outcomes is often straightforward, as game mechanics are designed to target acquiring the desired skills or knowledge. In contrast, gamified content presents a more complex challenge as gamified elements motivate users in behaviors that may lead to desired results. Therefore, evaluating the effectiveness of game mechanics is crucial to ensure a high-quality user experience, as ineffective mechanics or monotonous gameplay can lead to disengagement ([Plass et al](#kix.8wuxyrxwbr40)., 2011). Additionally, it is vital to confirm that increased motivation translates into actual learning outcomes in gamification. For example, while a compelling narrative might engage students, more is needed to ensure skill acquisition.

Measuring motivation presents challenges, as it is influenced by various personal, social, and cultural factors ([Jackson et al.](#kxiap3awy6qk), 2018; [Dörner et al](#nztyxq9r4e4z)., 2016). For instance, while leaderboards may effectively motivate certain players, they may demotivate others less inclined toward competitive mechanics. There are numerous considerations that must be addressed in order for DGBL to be effective. Adding game elements to a curriculum without appropriate design is unlikely to enhance learning ([Hu et al](#ivq88tcnkm9a)., 2024; [Fadda et al](#awx1i2bcrcuu)., 2022; [Ilin](#9jkegyplepb7), 2020). Evaluating individual game mechanics and their impact on specific outcomes is essential, as failing to do so impedes establishing causal relationships. Failing to do so makes it difficult to establish causal relationships that lead to the outcomes ([Landers](#kix.qmahiul2npbz), 2014). However, absence of standardized methods for measuring effectiveness also hampers comparing and validating results across studies ([Jackson et al](#kxiap3awy6qk)., 2018). Chapter 5 will address these challenges and outline the methodologies to tackle them.

### 2.6.1 Personal Preferences

Play styles significantly influence engagement in DGBL. Play styles refer to individual preferences for specific game mechanics and are integral to how players interact with a game ([Karpinskyj et al](#93f8011hcbmm)., 2014). While games often blend multiple mechanics, researchers have categorized various genres outlined in [Table 1](#kix.1hqm73t3p0pv).

**Table** 1: Game Genres ([Gui](#ep5u464ko7jn) et al., 2023, p5)

| **Game type** | **Description** |
| --- | --- |
| Role-playing | Playing one or more characters and moves in the virtual world |
| Strategy | Making strategic deployment through the application of game rules and plan |
| Simulation | Training in a realistic game environment |
| Construction | Accessing, using, and managing resources in an open-game environment |
| Adventure | Focusing on episodic and exploratory interactions such as exploring the unknown and solving puzzles |
| Action | Making quick movements and quick reactions in the game |
| Puzzle | Solving puzzles through logical thinking |

Effective educational game design ensures that gameplay elements align with learners' needs, affecting their overall gaming experience and satisfaction ([Fullerton](#l2l1gr2giok6), 2024; [Slussareff et a](#ff9tv2p48w61)l., 2016). Equally important are learning preferences, which affect how players absorb and apply information. [Honey and Mumford](#zer51s4d3kgm) (1992) attempted to categorize learning preferences into four styles but were criticized for oversimplifying a complex phenomenon ([Reynolds](#l6knn3dw54h7), 1997). Rather than adhering to rigid learning style classifications, modern frameworks like Universal Design for Learning (UDL) emphasize flexibility, offering multiple ways for players to engage with content ([Rose & Meyer](#wy88pgq5i2ji), 2007). This approach acknowledges that no single game can cater to every preference but can provide options that make learning more accessible and meaningful ([Smith](#vm5f3ymtljbm), 2012).

DGBL represents a transformative educational approach, equipping learners with essential skills and knowledge for navigating complex, real-world scenarios. The success of DGBL hinges on its ability to merge engaging gameplay with educational objectives. These elements are inherently interconnected as game mechanics form the structural backbone supporting engagement and learning activities. Chapter 5 covers the Contra Indication strategy of leveraging game mechanics and engagement strategies, to reach the learning objectives.

# 

# 3. DGBL in Medical Education

In medical education, systematic reviews have shown that gamification can significantly enhance medical students' learning outcomes across multiple domains ([Salehi et al](#35gkalho12uv)., 2023; [Xu et al](#5ulvlntydghf)., 2023). Gamified platforms in medical education have been shown to increase student engagement and promote decision-making ([Krishnamurthy et al.](#efp72flubcgv), 2022; [Do et al](#3htqe7lpesg)., 2023). Common gamified elements employed in medical educational settings include badges awarded for module completion, points for correct answers, and leaderboards ([Al-Rayes et al](#uwknjysmizmc)., 2022). Moreover, gamified quizzes and interactive assessments facilitate continuous learning and provide mechanisms for tracking progress ([Öz & Ordu](#kix.pzemx7n0szig), 2021). Additionally, gamified case studies engage students by requiring them to navigate complex patient diagnoses and management tasks within competitive or reward-based frameworks ([Do et al](#3htqe7lpesg)., 2023).

Serious games in medical education have demonstrated considerable potential to improve learning outcomes, technical skills and clinical reasoning across diverse medical disciplines ([Caserman et al](#81a3otko6rqm)., 2020; [Sharifzadeh et al](#5zyq46tlq595)., 2020; [Olgers et al](#90uc64um2bo)., 2021; [Koelewijn et al.](#vu5e4gftw2fj), 2024; [Wang et al](#ql2qopminbu5)., 2022; [Zairi et al](#59ne8isskbz8)., 2022; [Graafland et al](#e118l27iden4)., 2012; [Wang et al](#32a796ynk3jm)., 2016; [Xu et al](#5ulvlntydghf)., 2023; [Morgan et al](#178sdhhxyv91)., 2024; [Abd-alrazaq et al](#kix.hegqro1dkgxo)., 2022). For instance, simulation-based games replicating real-world medical scenarios—such as surgeries, emergency care, and patient interactions—allow students to practice decision-making and procedural skills in a risk-free environment. The screen-based touch surgery simulator utilizes cognitive task analysis to simulate, plan, and practice surgical procedures ([Bunogerane et al.](#jxkbcxpu45jc), 2018). Use case studies show significant skill improvement in junior doctors compared to their peers learning from textbooks ([Nehme et al](#jfb1p72whex3)., 2016). The game's internal assessment tools use metrics to give immediate feedback and record the learner's progress ([Bunogerane et al.](#jxkbcxpu45jc), 2018). In the same vein, interacting with SimMan, a life-size mannequin with realistic sounds and pulse, showed a significant increase in students' test scores in both technical proficiencies and non-technical competencies ([Swamy et al](#fe8bbr7yjrfs)., 2013).

Enjoyment is a key factor in learning. Septris, an online sepsis recognition and management game, improved knowledge and was enjoyable ([Evans et al](#fo1i1yhpz3c4)., 2015). Although students playing InsuOnline only scored slightly better than the control group, they reported enjoying the game more than traditional lectures. They considered the tool as a more effective way of learning ([Diehl et al](#l2p68baz95go)., 2015). Serious games in medicine are not only limited to simulation and digital format. [Hage et al](#a2m9bpy4dwh7). (2023) developed a card game based on “Happy Families” as a learning tool for musculoskeletal discipline. GridlockED is a board game successfully taught procedures and prioritization in the emergency department ([Brar et al](#dulsj0szohw9)., 2021).

Despite these successes, serious games remain underutilized in mainstream medical education, often due to limited availability and lack of integration into formal curricula ([Gorbanev et al](#f5j3ss9jxg3z)., 2018; [Olgers et al](#90uc64um2bo)., 2021). This gap underscores the need for broader adoption and innovative applications to leverage the potential of digital learning tools.

## 3.1 Addressing Diagnostic Challenges with DGBL

This section presents several case studies demonstrating the effectiveness of Digital Game-Based Learning (DGBL) in diagnostic training, particularly in emergency medicine. Below are two examples of serious games that provide interactive, immersive environments where learners can refine their clinical reasoning by making diagnostic and treatment decisions in simulated high-stakes scenarios.

### 3.1.1 EMERGE: A Case Study in Emergency Medicine Training

[Chon et al](#pasly7pfa999). (2019) examined the impact of EMERGE, a serious game designed to improve declarative and procedural knowledge in emergency medicine. EMERGE offers a highly interactive environment where students can freely navigate a virtual emergency department, interact with a digital mentor, and make independent diagnostic and treatment decisions. The game immerses students in realistic clinical scenarios, requiring them to assess incoming patient cases, gather medical histories, order diagnostic tests, and determine appropriate treatment plans. These decisions have direct consequences, as virtual patients respond dynamically to interventions, with medications producing both therapeutic and side effects. One of EMERGE’s distinguishing features is its high-fidelity graphics, contributing to an authentic emergency department atmosphere. The game promotes critical thinking and problem-solving by allowing students to explore different diagnostic pathways and treatment options.

Findings from [Chon et al](#pasly7pfa999). (2019) indicate that EMERGE significantly improved students’ declarative knowledge, enhancing their ability to recall and apply theoretical concepts. Student feedback was positive, with participants describing the game as engaging, user-friendly, and a valuable learning tool. However, despite its strengths, the study found no significant improvement in procedural knowledge, which is crucial for clinical practice. This finding highlights a key limitation of simulation-based learning in its current form—while practical for reinforcing theoretical knowledge, it may not fully translate into valuable skills. This gap suggests opportunities to explore ways to improve transfer between virtual learning and real-world practice.

### 3.1.2 PediatricSim: Advancing Pediatric Emergency Training

Like EMERGE, PediatricSim is a serious game designed for medical students and residents, focused explicitly on pediatric emergency medicine. According to [Gerard et al](#h1tbzxiiu8p9). (2018), training healthcare professionals in pediatric emergency medicine is often limited, and exposure to specific critical scenarios is inconsistent, making it difficult for junior doctors to develop expertise through clinical experience alone.

The game is designed to train and assess medical students and residents in seven pediatric emergency scenarios. Developed using the Unity game engine, PediatricSim is a three-dimensional, first-person, single-player simulation accessible via computer download or web browser. The game places users as code leaders who must assess and stabilize critically ill pediatric patients by directing a team of virtual healthcare providers.

While PediatricSim shares high-fidelity graphics and a realistic environment with EMERGE, it emphasizes immersion and realism by minimizing visual cues and requiring players to interact with virtual patients in a way that mirrors real clinical encounters. For example, instead of selecting actions from a predefined menu, players must click on body parts to access diagnostic and treatment options. Audio and text-based feedback (e.g., heart and lung sounds, skin temperature, and pulse strength) provide real-time assessments, and players can order laboratory tests, monitor vitals, and administer treatments.

A unique aspect of PediatricSim is its two play modes. The first is Tutorial Mode, which provides real-time guidance and feedback for incorrect actions, helping players learn as they progress. Non-Tutorial Mode – Removes hints, requiring players to rely on their clinical knowledge and decision-making skills. Additionally, PediatricSim features a case log, which records all player actions and response times. This allows for detailed performance analysis, making it a helpful tool for self-assessment and faculty-led evaluation.

According to [Gerard et al](#h1tbzxiiu8p9). (2018), there was a strong positive correlation between PediatricSim performance and written test scores, supporting the game's effectiveness as a teaching tool. The study concluded that serious gaming can improve decision-making in pediatric emergencies, reinforcing its value in medical education.

Simulations like EMERGE and PediatricSim provide valuable lessons for developing tools tailored to medical training, where clinical reasoning and procedural skills are critical. By integrating realistic decision-making scenarios with immersive, hands-on components, serious games can transform how medical students and trainees prepare for high-stakes situations. The following section discusses how AI could augment and strengthen DGBL.

# 4. AI

Artificial Intelligence (AI) refers to computer systems capable of performing tasks that typically require human intelligence, such as pattern recognition, decision-making, and learning ([Cardona et a](#20tf2kfudt8c)l., 2023). The term "artificial intelligence" was coined by John McCarthy in 1956, marking the start of a field that evolved significantly over the decades. Early progress was slow due to technological limitations, but by the 1990s, advancements in hardware enabled AI to perform tasks like handwriting recognition, with applications in areas such as mail sorting ([Urmeneta & Romero](#kix.oiqkzr5kfvgh), 2024). A breakthrough came in 2009 with the rise of graphics processing units (GPUs), which provided the computational power necessary for deep learning. This shift culminated in the 2012 ImageNet Challenge, where deep learning models outperformed previous benchmarks in image recognition ([Gershgorn](#nycatq6tdjaf), 2017).

By 2015, deep learning had become a standard methodology in various domains, including image and speech recognition, facial recognition, and language translation. This widespread adoption was fueled by the abundance of online data and increased computational resources. The introduction of transformer-based models in 2017 marked a significant breakthrough in AI. These models improved AI systems' ability to understand context and perform "self-supervised learning" on large datasets that did not require explicit labeling ([Urmeneta & Romero](#kix.oiqkzr5kfvgh), 2024). The release of GPT-3.5 in late 2022 represented another substantial leap in AI capabilities. Based on this model, ChatGPT garnered widespread acclaim for its ability to generate diverse content, marking a transition from recognition-focused applications to generative tasks ([Bubeck et al](#slm3q064jn0c)., 2023).

Currently, AI varies based on its learning approach. Machine learning enables systems to improve through data analysis, deep learning leverages neural networks to recognize complex patterns, and natural language processing facilitates human-like communication. Applications of these technologies include computer vision, autonomous systems, and AI-driven content generation, all of which continue to push the boundaries of automation and intelligence ([Urmeneta & Romero](#kix.oiqkzr5kfvgh), 2024).

Enhanced computational power propelled AI beyond simple pattern recognition, enabling machine learning systems to analyze data and make predictions and decisions, automating complex processes. Deep learning applications in natural language processing, media recognition, content generation, and the development of autonomous vehicles exemplify this advancement. The evolution of AI presents unique opportunities for DGBL particularly in medical education, where AI-driven simulations can enhance diagnostic training and clinical reasoning.

## 4.1 AI in DGBL

EMERGE and PediatricSim have demonstrated effectiveness in medical education; however, gaps remain in understanding how digital game-based learning (DGBL) translates to real-world clinical decision-making, particularly regarding transference, retention, and sustained user engagement. Emerging technologies, particularly AI, offer new opportunities to enhance learning experiences in serious games ([Abd-alrazaq et al](#kix.hegqro1dkgxo)., 2022). Drawing on [Urmeneta & Romero’s](#kix.oiqkzr5kfvgh) (2024) book on AI applications in education and the [US Office of Educational Technology’s](#20tf2kfudt8c) (2023) report on AI and future learning, this section synthesizes key areas where AI mightstrengthen medical training games.

### 4.1.1 Adaptive Learning & Personalization

AI can analyze student performance patterns and tailor gameplay by adjusting difficulty levels, providing hints, and modifying case complexity to optimize learning. In PediatricSim, AI can offer personalized feedback based on the case log, delivering customized hints that help players focus on their weak areas. Similarly, AI enables dynamic difficulty adjustment by automatically scaling the complexity of cases based on player performance, eliminating the need for a separate tutorial mode. In EMERGE, AI-driven virtual tutors could replace preprogrammed mentors, offering context-specific explanations and adaptive guidance tailored to individual learning needs. Additionally, AI could enhance automated feedback and assessment in both EMERGE and PediatricSim by identifying knowledge gaps and providing targeted recommendations for improvement.

### 4.1.2. Natural Language Processing (NLP) for Interaction

NLP can facilitate more realistic medical interactions within the game environment. In PediatricSim, there is currently no patient dialogue, while EMERGE relies on scripted responses. AI-powered conversational agents could introduce dynamic, interactive dialogues, enhancing diagnostic reasoning and decision-making. Instead of relying on rigid menu-based inputs, AI-driven text and voice analysis could allow players to issue spoken or written commands, making interactions more natural and immersive.

### 4.1.3. Content Generation

AI can dynamically generate content within the game to enhance engagement and variability in training scenarios. PediatricSim and EMERGE currently rely on scripted interactions, which limits case diversity. AI-powered procedural case generation could create unique, randomized patient scenarios based on real-world epidemiological data, ensuring a more comprehensive learning experience. Furthermore, AI-generated narratives could enable adaptive storylines that evolve based on player decisions, making gameplay more engaging and personalized.

### 4.1.4. Data-Driven Insights

AI can analyze gameplay data to provide valuable insights that enhance learning outcomes and medical research. Through behavioral analytics, AI can track and assess player behavior, identifying patterns, measuring engagement, and evaluating learning progress. Additionally, AI can dynamically adjust the game’s challenges to align with a player's predicted interests and learning needs. The system can also modify rewards and incentives based on player progress and preferences, maintaining motivation and engagement. AI-driven learning analytics can generate detailed reports on player performance, knowledge acquisition, and areas requiring further attention. By leveraging longitudinal gameplay data, AI can help identify strengths, weaknesses, and opportunities for remediation, providing valuable insights for medical education.

### 4.1.5. Social Interaction and Collaboration

AI can enhance team-based medical training by simulating realistic collaborative environments. AI enables multidisciplinary practice in an interactive setting by simulating virtual team members such as nurses and paramedics. Additionally, AI can facilitate social learning by analyzing peer interactions and supporting knowledge-sharing mechanisms that improve teamwork and communication skills.

### 4.1.6. Accessibility

AI can improve the accessibility of educational games, making them more inclusive for diverse learners. Through assistive technologies, AI can support features such as voice commands, real-time translations, and visual adjustments, catering to players with different needs. Moreover, AI-driven customization can tailor learning environments to accommodate individual preferences and requirements, ensuring an equitable and personalized educational experience.

## 4.2 Shortcomings with using AI

### 4.2.1 Data and Privacy Concerns

Given our intention to use existing AI models, awareness of what data is being collected raises significant ethical concerns, particularly regarding privacy and data security ([Urmeneta & Romero](#kix.oiqkzr5kfvgh), 2024). To mitigate these concerns, service providers will be selected that offer transparent data collection rules and limit their collection to essential information not used for marketing or training purposes. Transparency with CCHV will be maintained by clearly communicating which entities can access the data and how it is used. Informed consent will be central to the process, ensuring learners understand the data collection and processing scope, with opt-in and opt-out mechanisms. Additionally, all collected data will be stored securely, adequately encrypted, and, whenever possible, anonymized to safeguard personal information.

### 4.2.2 Bias and Fairness

AI systems frequently inherit biases from their training data, which can skew learning outcomes ([Sabzalieva](#q7t98h6pela7), 2023). This could affect content delivery and assessments and disproportionately impact specific student demographics ([Kasneci et al.,](#kix.2adyvm6dph5i) 2023). To address this, AI models will be trained on diverse datasets or, if pre-trained models are used, that developers have implemented measures to minimize bias. Before deployment, comprehensive audits of AI algorithms will be conducted to identify and mitigate any existing biases. Notably, a "human-in-the-loop" system will be maintained to allow educators to intervene when biases are detected in learning outcomes ([Mollick & Mollick](#jyhipqdqhn7w), 2023). For instance, conducting regular bias audits with tools like IBM’s AI Fairness 360 toolkit can provide a structured way to identify and address biases in AI models. Regular human oversight ensures that bias does not persist unchecked.

### 4.2.3 Cognitive Bias

AI systems may reinforce cognitive biases, potentially affecting how learners process information or interact with the learning environment ([Sabzalieva](#q7t98h6pela7), 2023). This is particularly critical in medical education, where cognitive biases could lead to flawed decision-making ([Tversky & Kahneman](#ap1qxgmssngb), 1974). Educating students and developers on the risks of cognitive bias in AI and mitigation strategies will be essential. Specific algorithms can be developed to detect and correct cognitive biases in AI-driven learning paths. For medical education, bias-detection systems like Google’s What-If tool could be used to analyze how the AI interacts with learners across different cases. This could prevent incorrect diagnostic reinforcement and support unbiased learning experiences.

### 4.2.4 Gender and Diversity Considerations

AI systems can unintentionally perpetuate gender stereotypes or marginalize diverse learners by delivering content reflecting biases in their training data ([Baidoo-Anu & Ansah](#smffbjbmphs6), 2023). To counteract this, the development process will incorporate gender-neutral and culturally inclusive language and scenarios. User testing will be conducted across a diverse range of students to ensure AI responses are appropriate for all demographics. Incorporating case studies from various cultures and ensuring gender-neutral language throughout the curriculum are vital actions. Tools like Microsoft’s Inclusive Design Toolkit can guide developers through this process.

### 4.2.5 Academic integrity

AI systems, including generative models, can produce “hallucinations” or incorrect outputs, leading to potential misinformation in educational content ([Bubeck et al](#slm3q064jn0c)., 2023). To prevent this, human review processes will be incorporated to verify the accuracy of AI-generated content. Secondary AI models could also detect and flag potential hallucinations, ensuring only accurate information is delivered. Integrating fact-checking algorithms such as ClaimBuster can help verify the accuracy of AI-generated content before it reaches students, or by training students on how to use such tools ([Mollick & Mollick](#jyhipqdqhn7w), 2023). Implementing human oversight in review stages further strengthens content reliability ([Cardona et a](#20tf2kfudt8c)l., 2023).

### 4.2.6 Technical Complexity

Developing AI-powered DGBL solutions demands significant technical expertise and resources, posing challenges, particularly for small-scale academic projects. This project aims to create a feasible proof of concept within the project's scope. Pre-built, modular AI solutions will be leveraged to reduce the need for in-house development. Partnerships with internal and external university resources may be necessary if additional support is required. Collaborating with open-source initiatives like TensorFlow or PyTorch, which have robust community support, can significantly reduce the complexity of the technical development process. Furthermore, seeking support through platforms like Google AI Impact Challenge can provide resources and visibility for educational AI innovations. Should the proof of concept prove successful, applications for funding from government or NGO bodies focused on educational innovation will be pursued to expand the project.

## 4.3 The Future of AI in Medical DGBL

Integrating AI into medical learning games promises to create personalized and adaptive educational experiences tailored to individual learners ([Lee et al](#tqcygv2jfi1q)., 2024). Futurists highlight AI’s ability to dynamically adjust game content and difficulty levels, enhancing engagement and skill acquisition ([Haleem et al](#qnkigad83ery)., 2022; [Baidoo-Anu & Ansah](#smffbjbmphs6), 2023). However, AI is not a guaranteed solution. [Goh et al](#jn5g1r8iw3yz)., (2024) found that access to large language models (LLMs) did not significantly improve physicians’ diagnostic reasoning compared to conventional resources, while [Tolks et al](#d83zt8pwsuv2). (2024) noted that empirical evidence on AI’s impact in serious health games remains limited, with most studies being small-scale, non-randomized pilots. While AI applications may seem futuristic, [Kung et al](#ax1xlq3ji4ma). (2023) demonstrated its potential when ChatGPT passed the U.S. Medical Licensing Exam without specialized training or reinforcement. [Li et al](#kix.lfye8xwb4br8). (2023) explored AI’s recognition of human facial expressions, laying the groundwork for more dynamic, personalized learning experiences. Users with basic knowledge of prompt engineering can customize AI personas as tutors or interactive game characters ([Mollick & Mollick](#jyhipqdqhn7w), 2023; [Sabzalieva](#q7t98h6pela7), 2023). In gaming, [Tao et al](#rhtxz7mfk5v7). (2024) successfully integrated AI to create highly realistic scenarios, [Ashby et al](#kix.5lnr1696d27s). (2023) developed diverse and authentic characters, and [Shao et al](#7rpfhhsjime6). (2024) used LLMs to generate contextually appropriate responses in role-playing games. With proper implementation, AI can bridge the gap between entertainment and education ([Bubeck et al](#slm3q064jn0c)., 2023; [Abd-alrazaq et al](#kix.hegqro1dkgxo)., 2022; [Kasneci et al.,](#kix.2adyvm6dph5i) 2023).

# 

# 5. Theories & Methods

This study aims to address a key research question:

How can Digital Game-Based Learning, enhanced with Large Language Models, improve clinical reasoning in junior doctors?

To address this, the main question is broken down into key sub-questions that explore different aspects of the learning process, the impact of game dynamics, and the usefulness of Large Language Models in scaffolding and engagement:

* How do junior doctors interact with various game elements ([Caroux](#k8eo07g11gvu), 2023)?
* How do these game elements impact learning, engagement, and enjoyment?
* How are AI-driven features utilized, and to what extent were they perceived as beneficial?

To investigate these questions, this study proposes the development of a **serious game** called **Contra-Indication**, designed to improve diagnostic accuracy among junior doctors. The game aims to enhance **clinical reasoning (LO 1)** by guiding players to apply **medical knowledge (LO 2)** in identifying threats, prioritizing patients **(LO 3)**, and adhering to medical protocols **(LO 4)**. Effective communication through **ISBAR (LO 5)** is essential for success, while failure triggers **probability-based challenges (LO 6)** that reinforce medical testing skills. Integrating these elements, the game aims to bridge theoretical learning with clinical practice.

Using the Evidence-Centered Design Framework ([Mislevy, Almond, & Lukas](#rhko02g0kl7m), 2003) the six LOs were aligned with specific game dynamics and assessment metrics to evaluate the game's effectiveness in promoting learning and engagement. [**Table 2**](#hntqb3do26l7) outlines the relationship between the learning outcomes, evidence models, game mechanics, and task structures, ensuring the game design aligns with its educational goals.

This chapter is separated into four sections. The first two sections cover theories of learning and motivation that underpin game feature choices. The third section outlines methods for measuring learning and motivation. The final section showcases the proposed gameplay, via a three-level game structure and a case study using **s**epsis to illustrate the complexity of medical decision-making.

**Table 2: ECD Framework for Contra-Indication**

| **Learning Outcome** | **Evidence Model** How to measure? | **Game Mechanic** | **Task Model** |
| --- | --- | --- | --- |
| 1 Clinical Reasoning | Game analytics and observation | Sorting out confusing, incomplete or irrelevant information to narrow down the possible options | 1) consider the situation, 2) collect cues, 3) process information, 4) understand the problem, 5) set goals, 6) implement interventions, 7) evaluate outcomes, 8) reflect & learn. ([Koelewijn et al.](#vu5e4gftw2fj), 2024) |
| 2 Prioritization | Game analytics and observation. Data from a test developed in partnership with CCHV | Choosing between wait, test, treat, consult (ask for help), or discharge in the correct sequence | Identifying which patients need help urgently |
| 3 Clinical Decision-Making | Game analytics and observation. Data from a test developed in partnership with CCHV | Choosing between wait, test, treat, consult (ask for help), or discharge in the correct sequence | Selecting the appropriate medical procedures, tests, and tools to carry out tasks |
| 4 Policy Comprehension | Game analytics and observation. Data from a test developed in partnership with CCHV | Choosing between wait, test, treat, consult (ask for help), or discharge in the correct sequence | Choosing the Best Course of Action: hospital policy, ethical guidelines, and when to ask for help |
| 5 communication | Game analytics on the ISBAR | To check out a patient from the ED, the player must debrief using ISBAR | Deliver concise and relevant information in language specific to that medical profession |
| 6 probability | Game analytics on their ability to estimate probability quickly ([Pauker & Kassirer](#27606nwmngl), 1980) | Calculating X and Y coordinates to find the weak point | Use pretest probability with Sensitivity and Specificity values to determine post-test probability of disease for negative/positive results |

## 5.1 Learning

The primary goal of this project is to meet the learning objectives set forth by CCHV. To evaluate learning effectiveness, this study will employ a three-stage assessment framework: (1) a pre-test to establish baseline knowledge, (2) a post-test to measure immediate learning gains, and (3) a follow-up assessment to evaluate retention over time. Senior CCHV doctors will oversee the design of the evaluation tool, ensuring it accurately measures cognitive and skill-based gains for the LO listed in Table 2.

Learning is a complex and multifaceted process. [Schunk](#mpayf3y2mtqq), (2019) defines learning as a sustained change in behavior resulting from practice, which involves three key components: (1) repeated engagement with content, (2) measurable behavioral change, and (3) the persistence of that change over time. While this definition is widely accepted, the most effective methods for fostering learning remain a topic of academic debate. Various learning theories emphasize different cognitive and social mechanisms. This study applies scaffolding and the Zone of Proximal Development (ZPD) as theoretical frameworks to examine how digital game-based learning (DGBL) enhances diagnostic accuracy in junior doctors.

[Vygotsky](#e3kem1y6q6sw) (1978) concept of ZPD describes the gap between what learners can achieve independently and what they can accomplish with guided support. Scaffolding, a constructivist instructional strategy, provides structured assistance to bridge this gap, gradually withdrawing support as learners gain proficiency ([Schunk](#mpayf3y2mtqq), 2019). A synthesis of scaffolding in DGBL ([Cai et al](#p8h4ri9k4ar6)., 2022) highlights five key strategies:

1. **Exposition** – Direct instructional content embedded within the game.
2. **Coaching** – Context-sensitive hints and real-time feedback mechanisms.
3. **Collaboration** – Role-based interaction with peers or AI that incorporates cooperative problem-solving.
4. **Reflection** – Opportunities for learners to evaluate and refine their decision-making.
5. **Debriefing** – Post-game discussions reinforcing key takeaways and learning outcomes.

Scaffolding strategies enhance learning by providing targeted guidance at critical points ([Locke & Latham](#kix.k8427ulrj85y), 2002). A meta-analysis examining the effect of scaffolding in DGBL in 49 primary studies shows that scaffolding in DGBL could effectively improve learning ([Cai et al](#p8h4ri9k4ar6)., 2022). However, research on how different scaffolding mechanisms influence long-term learning outcomes in digital game environments remains limited ([Koelewijn et al.](#vu5e4gftw2fj), 2024; [Cai et al](#p8h4ri9k4ar6)., 2022). Addressing this gap, this study will investigate how scaffolding strategies in Contra-Indication impact junior doctors' learning behaviors and diagnostic accuracy over time ([Caroux](#k8eo07g11gvu), 2023).

*Contra-Indication* will integrate Large Language Models (LLMs) as interactive non-player characters (NPCs) that provide coaching, collaboration, and reflective dialogue to enhance scaffolding and engagement. Section 5.4 elaborates on AI’s role in game design. Drawing from research demonstrating LLMs capacity for automatic question generation ([Lee et al](#tqcygv2jfi1q)., 2024), and adaptive learning content ([Mollick & Mollick](#jyhipqdqhn7w), 2023), this study will explore the use of Prompt Engineering techniques. However, given LLMs’ tendency to produce inaccurate or misleading information ("hallucinations"), the study will consider implementing Retrieval-Augmented Generation (RAG) if necessary. RAG improves accuracy by retrieving external medical training data, ensuring responses remain contextually relevant and factually reliable ([Martinea](#6vsrz8kep70k), 2021).

*Contra-Indication* aims to bridge the gap between theoretical learning and clinical practice by integrating scaffolding strategies and AI-driven support. The game will foster diagnostic accuracy, critical thinking, and decision-making skills in junior doctors through structured guidance, real-time feedback, and adaptive AI interactions. This approach aligns with established learning theories and contributes to the evolution of DGBL in healthcare education.

## 5.2 Motivation & Engagement

Motivation and engagement are closely related concepts, both essential for effective learning. Motivation is the internal drive that compels individuals to engage in specific activities ([Ryan & Deci,](#kix.s9l2ap6nke64) 2010) while engagement refers to the level of focus, interest, and enjoyment experienced during an activity ([Fadda et al](#awx1i2bcrcuu)., 2022). Research suggests that motivation is critical in driving engagement in educational games, leading to improved learning outcomes ([Hamari et al](#i7y49juya8ti)., 2016; [Jackson et al](#kxiap3awy6qk)., 2018; [Naul & Liu](#uv74pui62ngl), 2020). Motivation fuels engagement in behaviors aligned with educational goals ([Landers](#kix.qmahiul2npbz), 2014; [Van Roy & Zaman](#3rgestojyg7), 2018; [Ryan & Deci,](#kix.s9l2ap6nke64) 2010).

### 5.2.1 Motivation Types

Motivation can be categorized into extrinsic motivation, where individuals engage in activities for external rewards, and intrinsic motivation, where engagement stems from inherent enjoyment and personal satisfaction ([Ryan & Deci,](#kix.s9l2ap6nke64) 2010). Designing educational games that foster intrinsic motivation is challenging, requiring compelling gameplay mechanics that align with learning objectives. As a result, developers often incorporate extrinsic motivation strategies to maintain player engagement.

Extrinsic engagement strategies are often grounded in operant conditioning. [Majuri et al.](#1tjbx8gslrj) (2018) and [Dori](#kix.37z9mf2bh69z) (2015) identify strategies based on positive reinforcement (e.g., rewards for daily engagement) and fear of missing out (FOMO) (e.g., limited-time events that create a sense of urgency). Other standard methods include repetitive designs, endless leveling-up progress, frequent reminders, and pressure from in-game interactions. Online games, in particular, leverage multiplayer mechanics such as achievement sharing, cooperative play, and team-based incentives to drive engagement. While extrinsic rewards effectively capture initial interest, over-reliance on them can lead to diminished learning outcomes. Long-term engagement is best sustained through intrinsic motivation, which is closely tied to game design quality ([Ryan & Deci,](#kix.s9l2ap6nke64) 2010; [Van Roy & Zaman](#3rgestojyg7), 2018).

In educational games, intrinsic motivation is more complex as it requires better game design and engagement mechanics that result in learning. The flow theory provides a framework for optimizing player engagement in DGBL, demonstrating positive learning and motivational outcomes ([Hamari et al](#i7y49juya8ti)., 2016; [Zairi et al](#59ne8isskbz8)., 2022). Flow occurs when there is a balance between challenge and skill ([Csikszentmihalyi](#6hijuk3kfsuj), 2014), aligning with Vygotsky’s ZPD, where scaffolding balances frustration and boredom to maintain engagement and optimize learning outcomes.

AI can facilitate flow by balancing difficulty and adjusting to the player level. For example, [Li et al](#kix.lfye8xwb4br8)., (2023) used a game-based learning system that recognized player emotions via facial expressions. This allows AI to gauge frustration and satisfaction in real time and take action to maintain engagement. Another area where AI can manage flow is narrative. Ashby et al. (2023) introduced a novel framework for procedural content generation that allowed the personalisation of meaningful interactions between players and AI. Similarly, [Shao et al](#7rpfhhsjime6)., (2024) developed LLM agents based on a specific person to customise the interaction.

It is important to note that players do not respond equally to the same motivational strategies ([Van Roy & Zaman](https://docs.google.com/document/d/1BgGbxgc32H2qi9urz_DYNW7o0lNxpRspRc5yztfc_WE/edit?tab=t.9lmhji8c1uue#bookmark=kix.mj5dq8olqnrn), 2018). *Contra-Indication* will incorporate user experience (UX) best practices to enhance engagement ([Lei et al.,](#kix.t4ihyav9xkd9) 2022; [Salloum et al.,](#jaak8qldp3za) 2019). Additionally, AI-driven scaffolding such as adaptive levels ([Bijl](#yg06irclpmcz) et al., 2024), and contextual hints ([Cai et al](#p8h4ri9k4ar6)., 2022) can accommodate varied learning styles ([Fullerton](#l2l1gr2giok6), 2024). By aligning game mechanics with different playstyles, *Contra-Indication* aims to provide insights into effective engagement strategies for a broader audience ([Schuytema](#emu6006fxkf5), 2007).

## 5.3 Methodology & Data Collection

To evaluate the effectiveness of Contra-Indication, this study will adopt a mixed-methods approach, addressing three key dimensions: usefulness for learning, engagement, and enjoyment. The data collection process will incorporate:

**Testing Cognitive and Skill Based Outcomes  
The** CCHV team is invested in partnering to develop an assessment to establish evaluation metrics. The study will involve two consecutive cohorts of junior doctors, assessed before and after the course. The first cohort will be a control group, while the second will use *Contra-Indication*. A comparative analysis will measure the game's impact on knowledge acquisition and clinical decision-making skills.

**In-Game Analytics for Skill Based and Affective Outcomes:**Following [Ilin](#37yrasrsw429) (2022), in-game analytics will track player interactions, frequency of use, duration, and task completion rates to provide quantitative insights into engagement and learning patterns ([Bijl](#yg06irclpmcz) et al., 2024).

**Post-Level Surveys for Affective Outcomes**Adapting [Gerard et al](#h1tbzxiiu8p9). (2018) surveys from research on PediatricSim, this study will implement a five-point Likert scale to assess user perceptions ofengagement, enjoyment, and training effectiveness after each game level. The researchers established high validity for their survey, making it a precise tool to measure these variables.

**Semi-Structured Interviews for Affective Outcomes:**Once sufficient quantitative data is collected, semi-structured interviews will be conducted based on [Schunk](#mpayf3y2mtqq)’s (2019) advice to employ qualitative data to explore the underlying meaning behind quantitative patterns.

This study aims to comprehensively evaluate how Contra Indication supports junior doctors in developing diagnostic accuracy by combining quantitative analytics with qualitative insights.

## 5.4 Proposed Game Design

In Contra-Indication players will use their medical knowledge toinvestigate the causes of a strange disease and thereby save people from becoming (cute) monsters. Contra-Indication’s dynamics will bebased on RPG style, with several mini-games that will targetspecific skills. The game aims to train junior doctors to diagnose and treat the patients using clinical thinking, task prioritization, medical knowledge, policy comprehension, and communication. If patients are left untreated or the wrong decisions are made, their condition deteriorates until they become cute monsters, triggering a probability game as a last attempt to save them. See [Table 2](#hntqb3do26l7) for learning outcome and game dynamics.

### 5.4.1 Level 1.

The first chapter introduces the player to the game world and its dynamics. It is slow-paced to familiarize the player with the gameplay while exploring the plot. The narrative part of the game is skippable, allowing the player to focus only on learning. An in-game journal is available for players who skip but want to revisit the plot later.

The player joins their supervisor in the ICU ward. The supervisor guides the player through his actions, introducing the game dynamics of wait, test, treat, consult (ask for help), or discharge. Throughout the tutorial, the supervisor asks the player procedural questions and provides differential diagnosis (LO 1), giving advice and corrections. Prompt-based LLM will be used to create the NPC’s personality and RAG if the NPC’s feedback does not meet the expectations. In this level, the players are also introduced to the probability mini game when asked to estimate the accuracy of lab results. As part of this tutorial, the NPC supervisor tests the player's knowledge of different cognitive biases.

### 5.4.2 Level 2.

In the second chapter, the player goes on a mission to uncover the source of the disease. This level uses detective-style investigation given various clues that test their medical knowledge (LO 2) by narrowing down the choices and eliminating wrong options. The game will provide hints or tutorials that gradually step players through narrowing down choices.

As the narrative progresses, the player visits various locations, looking at how the apocalypse began. The game will provide diverse clues that require players to connect different medical concepts. The player will also review several case studies that will allow them to learn from others' mistakes. After the level wraps, the game will provide detailed feedback after every significant decision. Highlight what the player did well and how they can improve. This chapter will test their knowledge of task prioritization (LO 3), policy comprehension (LO 4), and communication (LO 5).

The player is also introduced to the probability game to save the patient (LO 6), where the percent of transformation (pretest probability) vs. weak spot location (Sensitivity and Specificity) by calculating the x and y axes. This is a time-sensitive task where the player must calculate quickly.

### 5.4.3 Level 3.

In this level, players must manage multiple patient cases, testing their ability to prioritize (LO 3), adhere to hospital policies (LO 4), and communicate effectively (LO 5) in high-pressure situations. The level takes place in the emergency department. The chapter starts slowly but quickly picks up the pace as more and more patients enter the ward. The mechanics will simulate resource scarcity (e.g., limited testing and medical supplies) that will tie into task prioritization. The player has to make rapid decisions (LO 1) and prioritize, monitor, and test. Once diagnosed, they must effectively communicate (LO 5) this information via ISBAR to discharge the patient. An LLM will be trained to recognize appropriate formats of ISBAR submissions and give feedback on how the input can be improved.

Patients who deteriorate due to a diagnostic error become cute monsters, activating the probability game (LO 6). The game gets progressively more challenging. At the end of the level, an LLM will be employed to allow for reflection. Data from the game will be fed to the LLM, which will analyze the players' mistakes and ask the player to reflect on what could have been done differently and what they learned. This feature encourages replayability and helps players reinforce the learning objectives.

### 5.4.4 Tying Gameplay to Clinical Thinking

To illustrate the game's functions, below is a case-based scenario using sepsis, a notoriously difficult-to-diagnose condition. According to Drs. Peckler and Bax, sepsis identification remains one of the most challenging yet vital skills for junior doctors, as prompt recognition can significantly impact patient outcomes in critical care. In medical literature sepsis is identified as the leading cause of admissions to Intensive Care Units (ICU) worldwide ([Marik](#3upehztz9776), 2014) and one of the most common preventable causes of death ([Duncan et al](#xuis5w1attul)., 2021), resulting in more deaths annually than breast, colorectal, and prostate cancers combined ([Suspected Sepsis: Recognition, Diagnosis and Early Management](#mkvctzod23yk), 2024). Despite its prevalence, sepsis is challenging to diagnose due to its complex presentation, emphasizing the need for advanced diagnostic training for medical professionals.

Sepsis is a condition when the immune system, in response to an infection, begins to damage the body’s tissues ([Zhang et al](#9sagjidvkf74)., 2016). Infectious agents, such as bacteria, viruses, fungi, or parasites, can trigger this response, making diagnosis difficult when the infection source is not readily identifiable ([Sweeney et al](#vdq11pt733tr)., 2019). Moreover, the inflammatory response may also be induced by non-infectious factors, such as post-surgical changes, blood clots, or autoimmune conditions, further complicating the diagnostic process ([Sweeney et al](#vdq11pt733tr)., 2019).

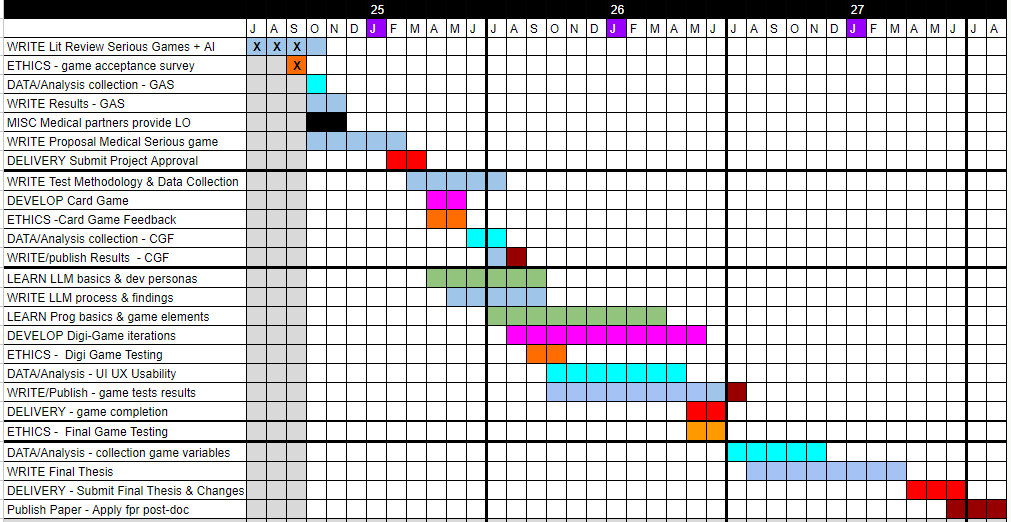
Early identification of sepsis is essential for reducing severe complications and mortality rates ([Zhang et al](#9sagjidvkf74)., 2016; [Marik](#3upehztz9776), 2014; [Seymour et al](#89c80dy9vndm)., 2016). Current clinical guidelines recommend initiating treatment within sixty minutes of diagnosis ([Sweeney et al](#vdq11pt733tr)., 2019); however, the non-specific nature of sepsis symptoms often delays recognition ([Suspected Sepsis: Recognition, Diagnosis and Early Management](#mkvctzod23yk), 2024). Compounding this challenge, no gold standard exists for diagnosing sepsis [(](https://www.zotero.org/google-docs/?Dh3BuS)[Zhang et al](#9sagjidvkf74)., 2016; [Marik](#3upehztz9776), 2014; [Suspected Sepsis: Recognition, Diagnosis and Early Management](#mkvctzod23yk), 2024). Blood tests are considered one of the most reliable methods for detecting infection. Most require two to three days for results and are limited by low sensitivity ([Marik](#3upehztz9776), 2014), with approximately 20–30% of sepsis cases returning negative results ([Seymour et al](#89c80dy9vndm)., 2016). Diagnostic accuracy can also be influenced by several variables, including the volume of blood drawn, the timing of the sample, prior antibiotic use, and the presence of specific organisms ([Sweeney et al](#vdq11pt733tr)., 2019). Furthermore, many pathogens—including certain bacteria, viruses, and fungi—are undetectable with conventional culture techniques, and there is a growing prevalence of sepsis cases caused by atypical organisms ([Duncan et al](#xuis5w1attul)., 2021). Due to time constraints, doctors often rely on biomarkers like white blood cell count, C-reactive protein (CRP), and procalcitonin (PCT) to help identify infections. However, the accuracy of these biomarkers is uncertain, as many lack sensitivity and specificity. ([Duncan et al](#xuis5w1attul)., 2021). Alternative diagnostic tools, such as the Sequential Organ Failure Assessment (SOFA), are valuable for predicting sepsis severity in ICU settings. For use outside the ICU, the quick SOFA (qSOFA) is a more straightforward tool that aids in rapidly assessing suspected sepsis without extensive lab tests ([Seymour et al](#89c80dy9vndm)., 2016).

In the game, players will encounter sepsis among other deadly conditions. The game dynamics allow players to monitor key parameters such as respiration, oxygen saturation, blood pressure, pulse, consciousness, and temperature. For example, a patient arrives with fever, low blood pressure, and confusion. The player must decide whether to treat, test, wait, or consult (ask for help). Their decisions affect the patient's condition, triggering alarms if they worsen. Successful early recognition and treatment of sepsis enable players to progress using ISBAR communication.

# 

# 6. Timeline & Stages

This project is structured into four key phases, each crucial in developing, testing, and finalizing a serious medical game alongside academic publications and a thesis. These phases are designed to build upon each other, ensuring a well-rounded and methodologically sound progression from theoretical exploration to practical application. Below is a breakdown of each stage:



## 6.1 Phase 1 - Literature Review & Proposal

| **2024 - 2025** | J | A | S | O | N | D | J | F | M |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| WRITE Lit Review Serious Games + AI | **X** | **X** | **X** |  |  |  |  |  |  |
| ETHICS - Doctor Interview |  |  | **X** |  |  |  |  |  |  |
| DATA/Analysis of Doctor Interview |  |  | X | X |  |  |  |  |  |
| WRITE Results - Doctor Interview |  |  |  | X | X |  |  |  |  |
| WRITE Proposal Medical Serious game |  |  |  | X | X | X | X | X |  |
| DELIVERY Submit Project Approval |  |  |  |  |  |  |  | X |  |

### 6.1.1 Literature Review: Serious Games and AI

The foundation of this project is an extensive literature review on serious games and artificial intelligence, focusing on medical education. This stage aims to identify trends, challenges, and knowledge gaps, laying a foundation for the game’s development. AI’s potential to enhance gameplay through personalized learning experiences, decision-making support, and adaptive feedback will be a focal point. Additionally, this phase will highlight how serious games can improve student engagement, knowledge retention, and problem-solving abilities.

### 6.1.2 Learning Outcomes and Game Mechanics

This study has received approval from the Capital & Coast District Health Board (ID: 2025-CCHV-390, 17-02-2025) and the university ethics committee (junior doctor survey: 2024/HE000153, senior doctor interviews: 2024/HE000359), ensuring adherence to ethical guidelines. the project's Learning Outcomes (LO) were established via interviews with Dr Peckler and Dr. Bax, 1 Clinical Reasoning, 2 Prioritization, 3 Clinical Decision-Making, 4 Policy Comprehension, 5 communication, and 6 probability. These are essential in defining the educational objectives and metrics for evaluating game success.

## 

## 6.2 Phase 2 - Testing Methodology & Pilot

| **2024 - 2025** | J | A | S | O | N | D | J | F | M | A | M | J | J | A |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| WRITE Test Methodology & Data Collection |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DEVELOP Card Game |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ETHICS -Card Game Feedback |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DATA/Analysis collection - CGF |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WRITE/publish Results - CGF |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Phase 2 begins with formulating a comprehensive test methodology in partnership with CCHV. Section 5.3 outlined the evaluation procedures for the game's effectiveness. At this stage, the best tools will be selected to measure the variables, including user engagement, AI performance, and qualitative factors such as user satisfaction and perceived usefulness.

A card game will be a prototype to test specific game mechanics and educational outcomes described in [Table 2](#hntqb3do26l7). A card game allows for easier testing and feedback gathering vs digital prototypes.During the development of the card game, another application for ethics approval will be submitted to collect feedback on the prototype.The game will be tested with a sample of participants who fit the selection criteria. Feedback from participants will be collected through structured interviews, surveys, and observational data. This feedback is crucial for refining the game's mechanical and educational elements. Particular focus will be given to user engagement, the clarity of game mechanics, and the effectiveness of the game content in reaching the LO.

The feedback data will be systematically analyzed to identify patterns and areas that require improvement. Quantitative and qualitative data will be considered, emphasizing how the game mechanics impact educational outcomes. The insights gained from this analysis will inform the development of the digital version of the game, ensuring that it aligns with user expectations while meeting its educational goals. By refining the card game based on this data, the research ensures that the final product is well-grounded in both theory and practice.The results from the card game feedback analysis will be documented and published. This publication will share insights into the card game's educational value and user experience, providing a foundation for the subsequent development of the game.

## 

## 6.3 Phase 3 - Digital Game Development & Testing

| **2025 - 2026** | J | A | S | O | N | D | J | F | M | A | M | J | J | A | S | O | N | D | J | F | M | A | M | J | J |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| LEARN LLM character design |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WRITE LLM process & findings |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LEARN Game Programming |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DEVELOP Digi-Game |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ETHICS - Digi Game Testing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DATA/Analysis - UI & UX |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WRITE/Publish - test results |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DELIVERY - game completion |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ETHICS - Final Game Testing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

### 6.3.1 Learn Large Language Model (LLM) Basics

Understanding the basic architecture and functionality of large language models (LLMs) is crucial to improving educational games' accuracy and relevance. This stage involves learning how LLMs process information, generate responses, and adapt to different contexts.

Once integrated, the implementation process will be documented, detailing how the LLMs were trained and applied, the methods used to develop AI personas, and the challenges faced during integration. The report will also evaluate the effectiveness of LLM-powered features in improving educational content and gameplay.

### 6.3.2 Learn Game Programming and Game Elements

This stage involves learning a game development environment, such as Godot., and design the digital game based on what was learned from testing the initial card game.

As the digital game nears completion, an application for ethical approval will be sought to ensure that testing follows established guidelines. This application for ethical clearance will cover key aspects such as data privacy, AI behavior, and user interaction. Specific attention will be given to how AI-driven features function within the game and whether they respect user autonomy, particularly in high-stakes educational environments like medical training. The ethical review will also assess the fairness of AI responses, ensuring they do not introduce biases or adversely affect learners based on their backgrounds or skill levels.

### 6.3.3 Data Collection and Analysis – UI/UX Usability

During the pilot testing of the digital game, data will be collected focusing on UI/UX usability. This analysis will focus on how users interact with the game, the ease of navigation, and how intuitive the educational content is. The results will guide adjustments to the game design in iterations, with each version incorporating improvements based on feedback from previous tests to balance engaging gameplay and effective learning outcomes, ensuring that the game is intuitive, user-friendly, and aligned with its learning objectives. The findings will be compiled and published once the digital game tests are completed. This publication will report on the game's usability and educational impact, focusing on how AI and LLMs contribute to the learning experience. The paper will explore how users interacted with the game, the role of AI in providing feedback, and the overall effectiveness of the game as a tool for medical education. It will also evaluate the UI/UX usability data, assessing how game design facilitated or hindered learning.

## 6.4 Phase 4 - Data Collection/Analysis and Final Thesis

| **2026 - 2027** | J | A | S | O | N | D | J | F | M | A | M | J | J | A |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| DATA/Analysis - product testing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WRITE Final Thesis |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DELIVERY - Submit Final Thesis |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Publish Paper - Apply for postdoc |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

### 6.4.1 Delivery – Game Completion

Upon integrating all feedback and making the necessary final adjustments, the digital serious game will reach completion. The final product will harmonize ethical considerations, AI enhancements, and gameplay refinements to produce an educational tool specifically designed for medical training. The game will have undergone iterative development, incorporating input from users and experts, ensuring it engages learners and meets its educational objectives.

Once the game is finalized, it will undergo comprehensive testing with a target audience, likely including medical students and professionals. This testing phase is essential to validate that the game achieves the intended learning outcomes, such as improving diagnostic accuracy, decision-making skills, and knowledge retention. Furthermore, the final product will be assessed for scalability, allowing for potential adaptation to other fields within medical training or even broader educational contexts. This completion phase marks the culmination of all the research, design, and development efforts, providing a ready-to-use, AI-powered educational tool.

### 6.4.2 Write Final Thesis

Following the completion of the digital serious game, the final thesis will be written. This thesis will offer a comprehensive account of the research process, from the initial literature review to game development, testing, and analysis. It will include detailed sections on the theoretical foundations of serious games and AI, the methodologies employed in card and digital game testing, and the ethical frameworks that guided the research. Additionally, the thesis will provide an extensive analysis of the data collected throughout the project, showcasing the game’s impact on medical education.

The thesis will also address the challenges faced during development, including integrating AI and LLMs, and how these challenges were overcome to create a functional educational tool. Special attention will be paid to how the findings contribute to the broader academic conversation surrounding AI-enhanced learning tools, serious games, and medical education.

Once the thesis is complete, it will be submitted for academic review. This review process may identify areas requiring revision to ensure the thesis meets academic standards. All necessary revisions will be addressed before the final submission.

### 6.4.3 Publishing and Post-Doctoral Application

After submitting the final thesis, the next step involves disseminating the research findings through academic publications. The key results, particularly those related to the educational efficacy of AI-enhanced serious games, will be shared in peer-reviewed journals. These publications will contribute valuable insights to the fields of AI in education, serious games, and medical training, inviting further research and dialogue.

Following publication, the researcher will apply for a post-doctoral position to continue exploring the intersection of AI and serious games in education. The post-doctoral phase will likely expand on the current research, developing more advanced medical AI learning tools or exploring other educational fields where serious games can have a transformative impact. The post-doctoral position will offer an opportunity to refine the integration of AI-driven learning environments further, aiming to advance the use of technology in education across various disciplines.

# 7. Conclusion

**Artificial Intelligence (AI)** and **Digital Game-Based Learning (DGBL)** offer a transformative opportunity for enhancing medical education, particularly in developing diagnostic skills among junior doctors ([Abd-alrazaq et al](#kix.hegqro1dkgxo)., 2022; [Morgan et al](#178sdhhxyv91)., 2024). This proposal explores the intersection of serious games, AI-driven adaptive learning, and medical training, emphasizing how these technologies can address key challenges such as critical thinking, task prioritization, policy comprehension, and communication. By leveraging AI, Contra-Indication seeks to create an engaging and immersive learning environment that dynamically adjusts to player performance, providing real-time feedback and scaffolding to optimize the learning experience.

Motivation and engagement are central to the success of DGBL, existing research highlights gaps in the understanding of how game mechanics align with learning outcomes. This study seeks to bridge those gaps by investigating how specific game features—such as role-playing, strategic decision-making, and simulation—enhance educational effectiveness. Additionally, personal learning preferences are explored, recognizing that not all learners engage with educational games similarly. Understanding these differences is crucial for designing adaptive and inclusive learning experiences that cater to diverse playstyles and medical training needs.

The success of DGBL depends on its ability to motivate and teach more effectively than traditional methods ([Ilin](#37yrasrsw429), 2022). Contra-Indication aims to achieve this by incorporating a player-centered design strategy based on user research, iterative testing, and adaptable mechanics ([Plass et al](#kix.8wuxyrxwbr40)., 2011). Based on the feedback, the game design will be refined ensuring that play style preferences and learning needs are effectively met ([Smith](#vm5f3ymtljbm), 2012) as well as well-designed learning content supported by game mechanics ([Clark et al.](#1qu9pocyv90d), 2016; [Dörner et al](#nztyxq9r4e4z)., 2016).

Ultimately, this research intends to contribute to the growing field of DGBL by demonstrating how serious games can be used to improve diagnostic accuracy and clinical decision-making. By refining AI-driven mechanics, exploring the effect of game features, and expanding the application of serious games, this study seeks to advance medical training and AI-powered education.

# Bibliography

Abd-alrazaq, A., Abuelezz, I., Hassan, A., AlSammarraie, A., Alhuwail, D., Irshaidat, S., Abu Serhan, H., Ahmed, A., Alabed Alrazak, S., & Househ, M. (2022). Artificial Intelligence–Driven Serious Games in Health Care: Scoping Review. JMIR Serious Games, 10(4), e39840. <https://doi.org/10.2196/39840>

Al-Rayes, S., Al Yaqoub, F. A., Alfayez, A., Alsalman, D., Alanezi, F., Alyousef, S., ... & Alanzi, T. M. (2022). Gaming elements, applications, and challenges of gamification in healthcare. Informatics in Medicine Unlocked, 31, 100974. <https://doi.org/10.1016/j.imu.2022.100974>

Ashby, T., Webb, B. K., Knapp, G., Searle, J., & Fulda, N. (2023). Personalized quest and dialogue generation in role-playing games: A knowledge graph-and language model-based approach. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1-20). <https://doi.org/10.1145/3544548.3581441>

Atkins, S. (2019). Assessing health professionals’ communication through role-play: An interactional analysis of simulated versus actual general practice consultations. Discourse Studies, 21(2), 109-134. <https://doi.org/10.1177/1461445618802659>

Baidoo-Anu, D., & Ansah, L. O. (2023). *Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning*. DOI:[10.61969/jai.1337500](http://dx.doi.org/10.61969/jai.1337500)

Ballhaus W., Chow W., Rivet E. & Spiegel B. (2024). Perspectives: Global Entertainment & Media Outlook 2024–2028. PwC. Retrieved July 25, 2024, from <https://www.pwc.com/gx/en/issues/business-model-reinvention/outlook/insights-and-perspectives.html>

Bijl, A., Veldkamp, B. P., Wools, S., & De Klerk, S. (2024). Serious games in high-stakes assessment contexts: A systematic literature review into the game design principles for valid game-based performance assessment. *Educational Technology Research and Development*.<https://doi.org/10.1007/s11423-024-10362-0>

Bindra, V., Chao, P., Srinivasa, S., & Koea, J. (2022). Are they ready? A survey of postgraduate year 1 and 2 surgical house officers. The New Zealand medical journal, 135(1565), 23–30. <https://doi.org/10.26635/6965.5872>

Bolstad, R., & McDowall, S. (2019). Games, Gamification, and Game Design for Learning: Innovative Practice and Possibilities in New Zealand Schools. New Zealand Council for Educational Research. PO Box 3237, Wellington 6140 New Zealand. <https://www.nzcer.org.nz/research/publications/games-gamification-and-game-design-learning-innovative-practice-and>

Brar, G., Lambert, S., Huang, S., Dang, R., & Chan, T. M. (2021). Using Observation to Determine Teachable Moments Within a Serious Game: A GridlockED as Medical Education (GAME) Study. AEM Education and Training, 5(2), e10456. <https://doi.org/10.1002/aet2.10456>

Buajeeb, W., Chokpipatkun, J., Achalanan, N., Kriwattanawong, N., & Sipiyaruk, K. (2023). The development of an online serious game for oral diagnosis and treatment planning: Evaluation of knowledge acquisition and retention. BMC Medical Education, 23(1), 830. <https://doi.org/10.1186/s12909-023-04789-x>

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). *Sparks of Artificial General Intelligence: Early experiments with GPT-4* (No. arXiv:2303.12712). arXiv.<http://arxiv.org/abs/2303.12712>

Bunogerane, G. J., Taylor, K., Lin, Y., & Costas-Chavarri, A. (2018). Using Touch Surgery to Improve Surgical Education in Low- and Middle-Income Settings: A Randomized Control Trial. Journal of surgical education, 75(1), 231–237. <https://doi.org/10.1016/j.jsurg.2017.06.016>

Cai, Z., Mao, P., Wang, D., He, J., Chen, X., & Fan, X. (2022). Effects of Scaffolding in Digital Game-Based Learning on Student’s Achievement: A Three-Level Meta-analysis. Educational Psychology Review, 34(2), 537–574. <https://doi.org/10.1007/s10648-021-09655-0>

Cardona, M. A., Rodríguez, R. J., & Ishmael, K. (2023). Artificial Intelligence and the Future of Teaching and Learning. U.S. Department of Education, Office of Educational Technology. <https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf>

Caroux, L. (2023). Presence in video games: A systematic review and meta-analysis of the effects of game design choices. Applied Ergonomics, 107, 103936. <https://doi.org/10.1016/j.apergo.2022.103936>

Caserman, P., Hoffmann, K., Müller, P., Schaub, M., Straßburg, K., Wiemeyer, J., Bruder, R., & Göbel, S. (2020). Quality Criteria for Serious Games: Serious Part, Game Part, and Balance. *JMIR Serious Games*, *8*(3), e19037.<https://doi.org/10.2196/19037>

Chen, C.-H., Shih, C.-C., & Law, V. (2020). The effects of competition in digital game-based learning (DGBL): A meta-analysis. Educational Technology Research and Development, 68(4), 1855–1873. <https://doi.org/10.1007/s11423-020-09794-1>

Chon, S.-H., Timmermann, F., Dratsch, T., Schuelper, N., Plum, P., Berlth, F., Datta, R. R., Schramm, C., Haneder, S., Späth, M. R., Dübbers, M., Kleinert, J., Raupach, T., Bruns, C., & Kleinert, R. (2019). Serious Games in Surgical Medical Education: A Virtual Emergency Department as a Tool for Teaching Clinical Reasoning to Medical Students. *JMIR Serious Games*, *7*(1), e13028.<https://doi.org/10.2196/13028>

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. Review of Educational Research, 86(1), 79–122. [https://doi.org/10.3102/00346 54315582065](https://doi.org/10.3102/00346)

Cook, A. (2022, January 14). Revealed: Sharp rise in claims ACC accepts for cancer misdiagnosis. Stuff. <https://www.stuff.co.nz/nz-news/350484394/number-of-claims-acc-accepts-for-cancer-misdiagnosis-doubles-in-five-years>

Cresswell, K. M., Panesar, S. S., Salvilla, S. A., Carson-Stevens, A., Larizgoitia, I., Donaldson, L. J., Bates, D., Sheikh, A., & World Health Organization's (WHO) Safer Primary Care Expert Working Group (2013). Global research priorities to better understand the burden of iatrogenic harm in primary care: an international Delphi exercise. PLoS medicine, 10(11), e1001554. <https://doi.org/10.1371/journal.pmed.1001554>

Csikszentmihalyi, M. (2014). Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi. Springer Netherlands. <https://doi.org/10.1007/978-94-017-9088-8>

Demirkol, O (2024). League of Legends player count. (n.d.). Retrieved July 25, 2024, from<https://www.thespike.gg/league-of-legends/beginner-guides/undefined/league-of-legends/beginner-guides/league-of-legends-player-count>

Diehl, L. A., Gordan, P. A., Esteves, R. Z., & Coelho, I. C. M. M. (2015). Effectiveness of a serious game for medical education on insulin therapy: A pilot study. Archives of Endocrinology and Metabolism, 59(5), 470–473. <https://doi.org/10.1590/2359-3997000000118>

Do, M., Sanford, K., Roseff, S., Hovaguimian, A., Besche, H., & Fischer, K. (2023). Gamified versus non-gamified online educational modules for teaching clinical laboratory medicine to first-year medical students at a large allopathic medical school in the United States. BMC Medical Education, 23(1), 959 <https://doi.org/10.1186/s12909-023-04951-5>

Dörner, R., Göbel, S., Kickmeier-Rust, M., Masuch, M., & Zweig, K. (2016). Entertainment Computing and Serious Games: International GI-Dagstuhl Seminar 15283, Dagstuhl Castle, Germany, July 5-10, 2015, Revised Selected Papers. Springer International Publishing. DOI: [10.1007/978-3-319-46152-6](http://dx.doi.org/10.1007/978-3-319-46152-6)

Dori A. (2015). 7 Psychological Tactics Used in Games to Hook Users.<https://www.doriadar.com/2015/10/12/7-psychological-tactics-used-in-games-to-hook-users/>

[Duncan, C. F., Youngstein, T., Kirrane, M. D., & Lonsdale, D. O. (2021). Diagnostic Challenges in Sepsis. *Current Infectious Disease Reports*, *23*(12), 22.](https://www.zotero.org/google-docs/?JkcHEl) <https://doi.org/10.1007/s11908-021-00765-y>

Evans, K. H., Daines, W., Tsui, J., Strehlow, M., Maggio, P., & Shieh, L. (2015). Septris: A Novel, Mobile, Online, Simulation Game That Improves Sepsis Recognition and Management. Academic Medicine, 90(2), 180–184. <https://doi.org/10.1097/ACM.0000000000000611>

Fadda, D., Pellegrini, M., Vivanet, G., & Zandonella Callegher, C. (2022). Effects of digital games on student motivation in mathematics: A meta‐analysis in K‐12. *Journal of Computer Assisted Learning*, *38*(1), 304–325.<https://doi.org/10.1111/jcal.12618>

Faustinella, F., & Jacobs, R. J. (2018). The decline of clinical skills: a challenge for medical schools. International journal of medical education, 9, 195–197.<https://doi.org/10.5116/ijme.5b3f.9fb3>

Fullerton, T. (2024). *Game Design Workshop. A Playcentric Approach to Creating Innovative Games* (5th ed.). Elsevier Morgan Kaufmann.<https://doi.org/10.1201/b22309>

Gerard, J. M., Scalzo, A. J., Borgman, M. A., Watson, C. M., Byrnes, C. E., Chang, T. P., Auerbach, M., Kessler, D. O., Feldman, B. L., Payne, B. S., Nibras, S., Chokshi, R. K., & Lopreiato, J. O. (2018). Validity Evidence for a Serious Game to Assess Performance on Critical Pediatric Emergency Medicine Scenarios. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *13*(3), 168–180.<https://doi.org/10.1097/SIH.0000000000000283>

Graber, M. L., Holmboe, E., Stanley, J., Danielson, J., Schoenbaum, S., & Olson, A. P. J. (2022). A call to action: next steps to advance diagnosis education in the health professions. Diagnosis (Berlin, Germany), 9(2), 166–175. <https://doi.org/10.1515/dx-2021-0103>

Gershgorn, D. (2017). The data that transformed AI research—and possibly the world. Quartz, July, 26(2013-2017), 52. <https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world>

Graber, M. L., Rencic, J., Rusz, D., Papa, F., Croskerry, P., Zierler, B., Harkless, G., Giuliano, M., Schoenbaum, S., Colford, C., Cahill, M., & Olson, A. P. J. (2018). Improving diagnosis by improving education: A policy brief on education in healthcare professions. *Diagnosis*, *5*(3), 107–118.<https://doi.org/10.1515/dx-2018-0033>

Gui, Y., Cai, Z., Yang, Y., Kong, L., Fan, X., & Tai, R. H. (2023). Effectiveness of digital educational game and game design in STEM learning: A meta-analytic review. International Journal of STEM Education, 10(1), 36. <https://doi.org/10.1186/s40594-023-00424-9>

Gorbanev, I., Agudelo-Londoño, S., González, R. A., Cortes, A., Pomares, A., Delgadillo, V., … Muñoz, Ó. (2018). A systematic review of serious games in medical education: quality of evidence and pedagogical strategy. Medical Education Online, 23(1). <https://doi.org/10.1080/10872981.2018.1438718>

Graafland, M., Schraagen, J. M., & Schijven, M. P. (2012). Systematic review of serious games for medical education and surgical skills training. *British Journal of Surgery*, *99*(10), 1322–1330.<https://doi.org/10.1002/bjs.8819>

Hage, R., Fourré, A., Ramonfosse, L., Leteneur, S., Jones, M., & Dierick, F. (2023). Description and rules of a new card game to learn clinical reasoning in musculoskeletal physiotherapy. *Journal of Manual & Manipulative Therapy*, *31*(4), 287–296.<https://doi.org/10.1080/10669817.2022.2132346>

Haleem, A., Javaid, M., & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 2(4), 100089. <https://doi.org/10.1016/j.tbench.2023.100089>

Hamari J., Shernoff D. J., Rowe E., Coller B., Asbell-Clarke J. & Edwards T. (2016). Challenging games help students learn: An empirical study on engagement, flow, and immersion in game-based learning. Computers in Human Behavior 54 (2016) 170e179 <https://doi.org/10.1016/j.chb.2015.07.045>

Hawkins, N., Younan, H. C., Fyfe, M., Parekh, R., & McKeown, A. (2021). Exploring why medical students still feel underprepared for clinical practice: a qualitative analysis of an authentic on‐call simulation. BMC Medical Education, 21, 1-11.<https://doi.org/10.1186/s12909-021-02605-y>

Hewitt, A., & Stubbs, M. (2017). Supporting law students’ skills development online – a strategy to improve skills and reduce student stress? *Research in Learning Technology*, *25*(0).<https://doi.org/10.25304/rlt.v25.1786>

Honey P., Mumford A. (1992). The manual of learning styles. Berkshire, England: Honey, Ardingly House.

Crimson Education NZ (2023). ​​How To Become a Doctor: Us vs UK. Retrieved February 24, 2025, from <https://www.crimsoneducation.org/nz/blog/how-to-become-a-doctor>

Goh, E., Gallo, R., Hom, J., Strong, E., Weng, Y., Kerman, H., Cool, J. A., Kanjee, Z., Parsons, A. S., Ahuja, N., Horvitz, E., Yang, D., Milstein, A., Olson, A. P. J., Rodman, A., & Chen, J. H. (2024). Large Language Model Influence on Diagnostic Reasoning: A Randomized Clinical Trial. JAMA Network Open, 7(10), e2440969. <https://doi.org/10.1001/jamanetworkopen.2024.40969>

Hu, Y., Wouters, P., Van Der Schaaf, M., & Kester, L. (2024). Timing of information presentation matters: Effects on secondary school students’ cognition, motivation and emotion in game‐based learning. *British Journal of Educational Technology*, bjet.13510.<https://doi.org/10.1111/bjet.13510>

Ilin, V. (2022). The role of user preferences in engagement with online learning. *E-Learning and Digital Media*, *19*(2), 189–208.<https://doi.org/10.1177/20427530211035514>

Ilin, V. (2020). The Good, the Bad and the Ugly. A Broad look at the Adaptation of Technology in Education. *EDU REVIEW. International Education and Learning Review / Revista Internacional de Educación y Aprendizaje*, *8*(2), 103–117.<https://doi.org/10.37467/gka-revedu.v8.2662>

Jackson, L. C., O’Mara, J., Moss, J., & Jackson, A. C. (2018). A Critical Review of the Effectiveness of Narrative-Driven Digital Educational Games: *International Journal of Game-Based Learning*, *8*(4), Article 4.<https://doi.org/10.4018/IJGBL.2018100103>

Jeuring, J., Grosfeld, F., Heeren, B., Hulsbergen, M., IJntema, R., Jonker, V., ... & van Zeijts, H. (2015). Communicate!—a serious game for communication skills—. In Design for Teaching and Learning in a Networked World: 10th European Conference on Technology Enhanced Learning, EC-TEL 2015, Toledo, Spain, September 15-18, 2015, Proceedings 10 (pp. 513-517). Springer International Publishing. <https://doi.org/10.1007/978-3-319-24258-3_49>

Karpinskyj, S., Zambetta, F., & Cavedon, L. (2014). Video game personalisation techniques: A comprehensive survey. *Entertainment Computing*, *5*(4), 211–218.<https://doi.org/10.1016/j.entcom.2014.09.002>

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274.<https://doi.org/10.1016/j.lindif.2023.102274>

Krishnamurthy, K., Selvaraj, N., Gupta, P., Cyriac, B., Dhurairaj, P., Abdullah, A., Krishnapillai, A., Lugova, H., Haque, M., Xie, S., & Ang, E.-T. (2022). Benefits of gamification in medical education. *Clinical Anatomy*, 35(6), 795–807. <https://doi.org/10.1002/ca.23916>

Koelewijn, G., Hennus, M. P., Kort, H. S. M., Frenkel, J., & Van Houwelingen, T. (2024). Games to support teaching clinical reasoning in health professions education: A scoping review. *Medical Education Online*, *29*(1), 2316971.<https://doi.org/10.1080/10872981.2024.2316971>

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, *2*(2), e0000198.<https://doi.org/10.1371/journal.pdig.0000198>

Landers, R. N. (2014). Developing a Theory of Gamified Learning: Linking Serious Games and Gamification of Learning. *Simulation & Gaming*, *45*(6), 752–768.<https://doi.org/10.1177/1046878114563660>

Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., & Kim, H. (2024). Few-shot is enough: Exploring ChatGPT prompt engineering method for automatic question generation in english education. Education and Information Technologies, 29(9), 11483–11515. <https://doi.org/10.1007/s10639-023-12249-8>

Lei, H., Wang, C., Chiu, M. M., & Chen, S. (2022). Do educational games affect students’ achievement emotions? Evidence from a meta‐analysis. *Journal of Computer Assisted Learning*, *38*(4), 946–959.<https://doi.org/10.1111/jcal.12664>

Li, Y.-C., Yang, K.-H., & Chang, C.-H. (2023). Development and implementation of a game-based learning system with real-time facial emotion recognition technology. *IIAI Letters on Informatics and Interdisciplinary Research*, *4*, 1.<https://doi.org/10.52731/liir.v004.156>

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American Psychologist, 57(9), 705–717. <https://doi.org/10.1037/0003-066X.57.9.705>

Majuri, J., Koivisto, J., & Hamari, J. (2018). Gamification of Education and Learning: A Review of Empirical Literature. In J. Koivisto, & J. Hamari (Eds.), Proceedings of the 2nd International GamiFIN Conference, GamiFIN 2018 (Vol. 2186, pp. 11-19). (CEUR Workshop Proceedings; Vol. 2186). CEUR-WS. <http://urn.fi/urn:nbn:de:0074-2186-5>

Marik, P. E. (2014). Don’t miss the diagnosis of sepsis! Critical Care, 18(5), 529. <https://doi.org/10.1186/s13054-014-0529-6>

Mayer, R. E. (2014). Computer games for learning: An evidence-based approach. MIT press. <https://doi.org/10.7551/mitpress/9427.001.0001>

Makary, M. A., & Daniel, M. (2016). Medical error-the third leading cause of death in the US. BMJ (Clinical research ed.), 353, i2139. <https://doi.org/10.1136/bmj.i2139>

Millwood, S. (2014). Developing a platform for learning from mistakes: changing the culture of patient safety amongst junior doctors. BMJ Open Quality, 3(1), u203658-w2114. <https://doi.org/10.1136/bmjquality.u203658.w2114>

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence‐centered design. ETS Research Report Series, 2003(1), i-29. <https://doi.org/10.1002/j.2333-8504.2003.tb01908.x>

Mollick, E. R., & Mollick, L. (2023). Assigning AI: Seven Approaches for Students, with Prompts. *SSRN Electronic Journal*.<https://doi.org/10.2139/ssrn.4475995>

Monrouxe, L. V., Bullock, A., Gormley, G., Kaufhold, K., Kelly, N., Roberts, C. E., Mattick, K., & Rees, C. (2018). New graduate doctors’ preparedness for practice: A multistakeholder, multicentre narrative study. BMJ Open, 8(8), e023146. <https://doi.org/10.1136/bmjopen-2018-023146>

Moore, M. (2016). *Basics of Game Design* (1st ed.). A. K. Peters. [ISBN: 9781439867761](https://www.oreilly.com/library/view/basics-of-game/9781439867761/)

Morgan, D. J., Scherer, L., Pineles, L., Baghdadi, J., Magder, L., Thom, K., Koch, C., Wilkins, N., LeGrand, M., Stevens, D., Walker, R., Shirrell, B., Harris, A. D., & Korenstein, D. (2024). Game-based learning to improve diagnostic accuracy: A pilot randomized-controlled trial. Diagnosis, 11(2), 136–141. <https://doi.org/10.1515/dx-2023-0133>

National Academies of Sciences, Engineering, and Medicine (2015). Improving Diagnosis in Health Care. Washington, DC: The National Academies Press. DOI: [10.17226/21794](https://nap.nationalacademies.org/catalog/21794/improving-diagnosis-in-health-care)

Naul, E., & Liu, M. (2020). Why Story Matters: A Review of Narrative in Serious Games. Journal of Educational Computing Research, 58(3), Article 3. <https://doi.org/10.1177/0735633119859904>

Nehme, J., & Chow, A. (2016). Touch Surgery Supporting the Plastic Surgical Community. *Plastic and Reconstructive Surgery–Global Open*, *4*(6), e779. [*DOI:* 10.1097/GOX.0000000000000778](https://journals.lww.com/prsgo/fulltext/2016/06002/touch_surgery_supporting_the_plastic_surgical.8.aspx)

NHS Resolution (2024). Annual report and accounts 2023/24. ISBN 978-1-5286-5053-3, HC 73. Retrieved from: <https://www.gov.uk/government/publications/nhs-resolution-annual-report-and-accounts-2023-to-2024>

Olgers, T. J., Bij De Weg, A. A., & Ter Maaten, J. C. (2021). Serious Games for Improving Technical Skills in Medicine: Scoping Review. *JMIR Serious Games*, *9*(1), e24093.<https://doi.org/10.2196/24093>

Olson, A. P. J., & Graber, M. L. (2020). Improving Diagnosis Through Education. Academic medicine: journal of the Association of American Medical Colleges, 95(8), 1162–1165. <https://doi.org/10.1097/ACM.0000000000003172>

Öz, G. Ö., & Ordu, Y. (2021). The effects of web based education and Kahoot usage in evaluation of the knowledge and skills regarding intramuscular injection among nursing students. Nurse education today, 103, 104910. <https://doi.org/10.1111/medu.15417>

Pauker, S. G., & Kassirer, J. P. (1980). The threshold approach to clinical decision making. The New England journal of medicine, 302(20), 1109–1117. <https://doi.org/10.1056/NEJM198005153022003>

Picton, A., Greenfield, S., & Parry, J. (2022). Why do students struggle in their first year of medical school? A qualitative study of student voices. *BMC medical education*, *22*(1), 100. <https://doi.org/10.1186/s12909-022-03158-4>

Plass, J., Homer, B., Kinzer, C., Frye, J., & Perlin, K. (2011). *Learning Mechanics and Assessment Mechanics for Games for Learning*.<https://doi.org/10.13140/2.1.3127.1201>

Reilly, J. B., Ogdie, A. R., Von Feldt, J. M., & Myers, J. S. (2013). Teaching about how doctors think: a longitudinal curriculum in cognitive bias and diagnostic error for residents. BMJ quality & safety, 22(12), 1044–1050. <https://doi.org/10.1136/bmjqs-2013-001987>

Rencic, J., Trowbridge, R. L., Fagan, M., Szauter, K., & Durning, S. (2017). Clinical reasoning education at

US medical schools: results from a national survey of internal medicine clerkship directors. Journal of

general internal medicine, 32, 1242-1246. <https://doi.org/10.1007/s11606-017-4159-y>

Reynolds, M. (1997). Learning Styles: A Critique. Management Learning, 28(2), 115-133. <https://doi.org/10.1177/1350507697282002>

Rose, D. H., & Meyer, A. (2007). Teaching every student in the digital age: Universal Design for Learning. Educational Technology Research and Development 55(5):521-525 [DOI:10.1007/s11423-007-9056-3](https://link.springer.com/article/10.1007/s11423-007-9056-3)

Ruedinger, E., Olson, M., Yee, J., Borman-Shoap, E., & Olson, A. P. J. (2017). Education for the Next Frontier in Patient Safety: A Longitudinal Resident Cerroriagnostic Error. American journal of medical quality : the official journal of the American College of Medical Quality, 32(6), 625–631. <https://doi.org/10.1177/1062860616681626>

Ryan R. M. & Deci L. E. (2010). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. American Psychological Association, Inc. 0003-066X/00/$5.00 Vol. 55, No. 1, 68-78 [DOI: 10.1037110003-066X.55.1.68](https://doi.org/10.1037//0003-066x.55.1.68)

Sabzalieva, E., & Valentini, A. (2023). *ChatGPT and artificial intelligence in higher education: Quick start guide*. UNESCO.<https://unesdoc.unesco.org/ark:/48223/pf0000385146>

Salehi, A. M., Mohammadi, H. A., Jenabi, E., Khanlarzadeh, E., & Ashtari, K. (2023). Quality of Evidence and Pedagogical Strategy in Using Gamification in Medical Education Literature: A Systematic Review. Simulation & Gaming, 54(6), 598-620. <https://doi.org/10.1177/10468781231195903>

Salen,﻿ K.,﻿ & ﻿Zimmerman,﻿ E.﻿ (2004).﻿ Rules of play: Game design fundamentals.﻿ MIT ﻿Press. ISBN: [9780262240451](https://mitpress.mit.edu/9780262240451/rules-of-play/)

Salloum, S. A., Al-Emran, M., Shaalan, K., & Tarhini, A. (2019). Factors affecting the E-learning acceptance: A case study from UAE. *Education and Information Technologies*, *24*(1), 509–530.<https://doi.org/10.1007/s10639-018-9786-3>

Schunk, D. H. (2019). Learning theories an educational perspective. Pearson Education, Inc. ISBN 10: [0134893751](https://www.abebooks.com/9780134893754/Learning-Theories-Educational-Perspective-Schunk-0134893751/plp) / ISBN 13: [9780134893754](https://www.abebooks.com/9780134893754/Learning-Theories-Educational-Perspective-Schunk-0134893751/plp)

Schuytema, P. (2007). *Game design: A practical approach* (1st ed.). ISBN 10: [1584504714](https://www.abebooks.com/9781584504719/Game-Design-Practical-Approach-Schuytema-1584504714/plp) / ISBN 13: [9781584504719](https://www.abebooks.com/9781584504719/Game-Design-Practical-Approach-Schuytema-1584504714/plp)

Scott I. A. (2009). Errors in clinical reasoning: causes and remedial strategies. BMJ (Clinical research ed.), 338, b1860. <https://doi.org/10.1136/bmj.b1860>

Seymour, C. W., Liu, V. X., Iwashyna, T. J., Brunkhorst, F. M., Rea, T. D., Scherag, A., Rubenfeld, G., Kahn, J. M., Shankar-Hari, M., Singer, M., Deutschman, C. S., Escobar, G. J., & Angus, D. C. (2016). Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA, 315(8), 762. <https://doi.org/10.1001/jama.2016.0288>

Shao, Y., Li, L., Dai, J., & Qiu, X. (2023). *Character-LLM: A Trainable Agent for Role-Playing* (No. arXiv:2310.10158). arXiv.<http://arxiv.org/abs/2310.10158>

Sharifzadeh, N., Kharrazi, H., Nazari, E., Tabesh, H., Edalati Khodabandeh, M., Heidari, S., & Tara, M. (2020). Health Education Serious Games Targeting Health Care Providers, Patients, and Public Health Users: Scoping Review. *JMIR Serious Games*, *8*(1), e13459.<https://doi.org/10.2196/13459>

Singh, H., Schiff, G. D., Graber, M. L., Onakpoya, I., & Thompson, M. J. (2017). The global burden of diagnostic errors in primary care. *BMJ quality & safety*, *26*(6), 484-494. DOI: [10.1136/bmjqs-2016-005401](http://dx.doi.org/10.1136/bmjqs-2016-005401)

Slussareff, M., Braad, E., Wilkinson, P., & Strååt, B. (2016). Games for Learning. In R. Dörner, S. Göbel, M. Kickmeier-Rust, M. Masuch, & K. Zweig (Eds.), *Entertainment Computing and Serious Games* (Vol. 9970, pp. 189–211). Springer International Publishing.<https://doi.org/10.1007/978-3-319-46152-6_9>

Smith, F. G. (2012). Analyzing a College Course that Adheres to the Universal Design for Learning (UDL) Framework. Journal of the Scholarship of Teaching and Learning, 12(3), 31–61. Retrieved from <https://scholarworks.iu.edu/journals/index.php/josotl/article/view/215>

Suspected sepsis: recognition, diagnosis and early management (2024). London: National Institute for Health and Care Excellence (NICE); 2024 Mar 19. (NICE Guideline, No. 51.) Available from: <https://www.ncbi.nlm.nih.gov/books/NBK553314/>

Swamy, M., Bloomfield, T. C., Thomas, R. H., Singh, H., & Searle, R. F. (2013). Role of SimMan in teaching clinical skills to preclinical medical students. *BMC medical education*, *13*, 1-6. <https://doi.org/10.1186/1472-6920-13-20>

Sweeney, T. E., Liesenfeld, O., & May, L. (2019). Diagnosis of bacterial sepsis: Why are tests for bacteremia not sufficient? Expert Review of Molecular Diagnostics, 19(11), 959–962. <https://doi.org/10.1080/14737159.2019.1660644>

Tao, M., Liang, X., Shi, T., Yu, L., & Xie, Y. (2024). *RoleCraft-GLM: Advancing Personalized Role-Playing in Large Language Models* (No. arXiv:2401.09432). arXiv.<http://arxiv.org/abs/2401.09432>

Ten Cate, O., & Scheele, F. (2007). Viewpoint: Competency-Based Postgraduate Training: Can We Bridge the Gap between Theory and Clinical Practice?: Academic Medicine, 82(6), 542–547. <https://doi.org/10.1097/ACM.0b013e31805559c7>

Thompson, C. G., & Von Gillern, S. (2020). Video-game based instruction for vocabulary acquisition with English language learners: A Bayesian meta-analysis. Educational Research Review, 30, 100332. <https://doi.org/10.1016/j.edurev.2020.100332>

Tolks, D., Schmidt, J. J., & Kuhn, S. (2024). The Role of AI in Serious Games and Gamification for Health: Scoping Review. JMIR Serious Games, 12, e48258. <https://doi.org/10.2196/48258>

Tsai, Y., & Tsai, C. (2020). A meta‐analysis of research on digital game‐based science learning. Journal of Computer Assisted Learning, 36(3), 280–294. <https://doi.org/10.1111/jcal.12430>

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science (New York, N.Y.), 185 (4157), 1124–1131. [DOI: 10.1126/science.185.4157.1124](https://doi.org/10.1126/science.185.4157.1124)

Urmeneta, A., & Romero, M. (Eds.). (2024). *Creative Applications of Artificial Intelligence in Education*. Springer Nature Switzerland.<https://doi.org/10.1007/978-3-031-55272-4>

Van Roy, R., & Zaman, B. (2018). Need-supporting gamification in education: An assessment of motivational effects over time. *Computers & Education*, *127*, 283–297.<https://doi.org/10.1016/j.compedu.2018.08.018>

Vygotsky, L. S., & Cole, M. (1978). Mind in society: the development of higher psychological processes. Cambridge: Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>

Wang, R., DeMaria, S., Goldberg, A., & Katz, D. (2016). *A Systematic Review of Serious Games in Training Health Care Professionals*. Simulation in healthcare: journal of the Society for Simulation in Healthcare, 11(1), 41–51. <https://doi.org/10.1097/SIH.0000000000000118>

Wang, L. H., Chen, B., Hwang, G.-J., Guan, J.-Q., & Wang, Y.-Q. (2022). Effects of digital game-based STEM education on students’ learning achievement: A meta-analysis. International Journal of STEM Education, 9(1), 1–13. <https://doi.org/10.1186/s40594-022-00344-0>

Martineau, L (2021). What is retrieval-augmented generation (RAG)? IBM Research. Retrieved on March 1, 2025. <https://research.ibm.com/blog/retrieval-augmented-generation-RAG>

Wimmers, P. F., Schmidt, H. G., & Splinter, T. A. (2006). Influence of clerkship experiences on clinical competence. Medical education, 40(5), 450–458. <https://doi.org/10.1111/j.1365-2929.2006.02447.x>

Xu, M., Luo, Y., Zhang, Y., Xia, R., Qian, H., & Zou, X. (2023). Game-based learning in medical education. Frontiers in public health, 11, 1113682. <https://doi.org/10.3389/fpubh.2023.1113682>

Yanagita, Y., Shikino, K., Ishizuka, K., Uchida, S., Li, Y., Yokokawa, D., ... & Ikusaka, M. (2023). Improving diagnostic accuracy using a clinical diagnostic support system for medical students during history-taking: a randomized clinical trial. *BMC Medical Education*, *23*(1), 383. DOI <https://doi.org/10.1186/s12909-023-04370-6>

Zairi, I., Ben Dhiab, M., Mzoughi, K., & Ben Mrad, I. (2022). The Effect of Serious Games on Medical Students’ Motivation, Flow and Learning. Simulation & Gaming, 53(6), 581-601. <https://doi.org/10.1177/10468781221123919>

Zhang, Z., Smischney, N. J., Zhang, H., Van Poucke, S., Tsirigotis, P., Rello, J., Honore, P. M., Kuan, W. S., Ray, J. J., Zhou, J., Shang, Y., Yu, Y., Jung, C., Robba, C., Taccone, F. S., Caironi, P., Grimaldi, D., Hofer, S., Dimopoulos, G., … Rocco, J. R. (2016). AME evidence series 001—The Society for Translational Medicine: Clinical practice guidelines for diagnosis and early identification of sepsis in the hospital. Journal of Thoracic Disease, 8(9), 2654–2665. <https://doi.org/10.21037/jtd.2016.08.03>

### Interviews

B. Peckler, personal communication, January 14, 2025

G. Bax, personal communication, February 4, 2025