# Game-Based Learning in Hybrid/Online Instruction

A Design Approach for Creating GBL College-Level Courses

A Continuing Edition

Joshua Parkinson

Continuing the Master's thesis submitted to the Information Design and Technology Program,

State University of New York Polytechnic Institute, December 2018

Updated with research and professional practice through 2026

Original thesis: December 2018

This edition: September 2026

**Abstract**

Educators working in hybrid and online college courses have a real opportunity to engage students through game-based learning (GBL) and game mechanics. The 2018 version of this thesis made that case and offered an example course built on tabletop role-playing game (RPG) mechanics. The 2018 thesis stands as submitted to SUNY Polytechnic Institute. This edition is not a resubmission. It continues that work to the present, drawing on eight more years of experience in instructional design, digital accessibility, and AI integration at a SUNY community college, and it tests the original design against a deeper reading of the research. The evidence supports games as learning environments. Across randomized and controlled studies, digital games improved learning compared with non-game instruction, and games with added scaffolding did better still (Clark et al., 2016). The same research shows that engagement and flow do not produce learning on their own. Students can be deeply absorbed in a game and still not learn the content (Admiraal et al., 2011), and learners without instructional support tend to learn the game instead of the subject (Ke, 2009). This edition therefore builds a design framework on socio-constructivist and situated views of learning. The framework has nine design principles organized around a repeating Brief, Play, Debrief cycle. It then applies the framework to a hybrid/online American history module aligned to the SUNY General Education learning outcomes, including the Information Literacy and Civic Discourse outcomes that take effect in Fall 2026. New sections address how role-play can coexist with historical accuracy, why collaboration should replace individual leaderboards, how to meet Universal Design for Learning and ADA Title II accessibility requirements, and how the design should be evaluated. It closes with limitations and a cautious look at generative AI as a support for the instructor who acts as Game Master.

*Keywords:* game-based learning, role-playing games, hybrid and online instruction, scaffolding, flow, universal design for learning, instructional design

**Acknowledgements**

I would like to express appreciation to all of my family, friends and colleagues who have encouraged me to continue my academic goals. Thank you to Gary Gygax and Dave Arneson for creating Dungeons and Dragons and my subsequent love for games in education. I would like to acknowledge my advisor, Dr. Steven Schneider, for his flexible pedagogy as well as his down-to-earth approach to information design. Thank you to C. T. Avis and K. Klump for your suggestions and the time you spent reading this project; several of the questions you asked in 2018 shaped this edition. Finally, I want to thank my wife for her persistent encouragement during all of my academic programs.

## Table of Contents

## About This Edition

I submitted this thesis to the Information Design and Technology program at SUNY Polytechnic Institute in December 2018, and it completed my Master of Science degree that month. That thesis stands as it was submitted. This document is not a resubmission, and it is not being submitted anywhere for credit. It is a continuation. It picks up the work where I left it in 2018 and carries it forward to today, with everything I have learned since.

A lot has happened in those years. In 2018 I was an academic advisor and Starfish manager at Herkimer County Community College, teaching as an adjunct on the side. In 2020 I became the college's Director of Instructional Design. Since then I have administered our learning management system and led the college's move from Blackboard to D2L Brightspace, reviewed courses each semester against SUNY's OSCQR course quality rubric, co-chaired the college's Accessibility Committee as the federal web accessibility rules for public colleges arrived, and served as the college's lead on AI integration, including writing its AI ethics and digital assessment policies. In 2024 I started building GamED, a game-based learning platform of my own. Every one of those experiences touched the questions this thesis asked.

When I went back and reread the 2018 thesis alongside the research I had collected for it, two things stood out. The first was that the core idea still holds up: a course can borrow the structure of a tabletop role-playing game to give students goals, roles, consequences, and a reason to work together. The second was that I had leaned on some of my sources for more than they actually said, and I had left a lot of good research sitting in my folders unused. Eight years of building real courses with real faculty also showed me where the 2018 design would have struggled in practice.

The course design at the center of this edition is the same one I proposed in 2018. What changed is how well it is supported and how carefully it is built. The major changes are:

- A learning-theory foundation. The 2018 thesis talked about "situated meaning" without ever naming the theory behind it. Chapter 2 now grounds the design in constructivist, sociocultural, and situated views of learning.
- A more honest treatment of flow. The original argued that flow leads to learning. My own sources say flow and engagement help, but do not produce learning without instructional support. The design now reflects that.
- A framework instead of a list. Chapter 3 organizes the design into nine principles and a repeating Brief, Play, Debrief cycle, each tied to evidence.
- An answer to the history question. My second reader asked whether letting students change history in a role-play would confuse the historical record. Chapter 4 now answers that directly.
- Rewards rebuilt around collaboration. Extra credit for reading LMS announcements is gone. Team progress and in-story recognition replace it.
- Accessibility. The original never mentioned it. After years of accessibility work at my college, I could not leave it out. The course now follows Universal Design for Learning and the ADA Title II web accessibility rule that applies to public colleges.
- Current outcomes and platforms. The example course is aligned to the SUNY General Education framework, and the platform section reflects what an LMS migration teaches about depending on any one tool, including the fact that Classcraft, which I recommended in 2018, shut down its original platform in 2024.

- An evaluation plan. The original described how to build the course but not how to know whether it worked. Chapter 5 adds that plan.
- Generative AI. It was not a classroom tool in 2018. Chapter 6 now considers what role it should and should not play.

Appendix A lists every change against the 2018 text, and Appendix B lists the reader comments from 2018 and where each one is addressed.

# Chapter 1: Introduction

The idea of building game mechanics into learning objectives is an easy one to consider, because most people have experienced a game in one way or another. Games teach players how to accomplish a goal, the different ways a goal can be reached, and what happens when a risk pays off or does not. Hunicke, LeBlanc, and Zubek (2004) describe games as designed artifacts rather than media, and argue that thinking about games this way "helps frame them as systems that build behavior via interaction" (p. 2). That framing is the starting point for this thesis. If a game is a system that builds behavior, then a course can be designed the same way: as a system of goals, rules, roles, and feedback that builds the behaviors a course is supposed to teach.

## Background and Problem

Hybrid and online courses are now a normal part of college. They give students flexibility, but they also make it easy for a student to become a name on a roster who reads, submits, and logs off. MacDonald and Thompson (2005), evaluating an online graduate course, found that learners were engaged in a "search for connection" and valued interactions that let them build content together (Abstract section). Students do not stop needing each other, or an instructor, just because the course moved online.

Games are good at exactly the things online courses struggle with. Good games give clear goals, fast feedback, a reason to try again after failing, and, in multiplayer games, a reason to rely on other people (Gee, 2005; Plass et al., 2015). Research on games as learning environments has grown a great deal. A large meta-analysis of randomized and controlled studies found that digital games improved learning compared with non-game instruction (Clark et al., 2016), and a longitudinal study of more than 1,400 adolescents found that playing strategic games, including

role-playing games, predicted higher self-reported problem-solving skills over four years (Adachi & Willoughby, 2013).

The same body of research carries a warning. Games can be engaging without being educational. Students in one well-designed history game showed clear signs of flow, yet their team flow was related to how well they played, not to what they learned (Admiraal et al., 2011). Reviews of instructional games keep reaching the same conclusion: without instructional support, learners learn to play the game rather than learn the content (Ke, 2009). The problem this thesis takes on is how to design a hybrid/online college course that gets the benefits of a game without falling into that trap.

**Purpose and Design Questions**

The purpose of this thesis is to present a research-grounded design approach for bringing tabletop role-playing game (RPG) mechanics into hybrid and online college courses, and to show that approach applied to a college-level American history module. It is design work. It does not report a completed experiment. It builds a framework from the literature, applies it to a real course structure, and lays out how the result should be evaluated. Three design questions guide the work:

1. What does research on learning, motivation, and games say should guide the design of game-based learning in hybrid and online college courses?
2. How can tabletop RPG mechanics be mapped to course outcomes in a way that protects content accuracy, supports every student, and keeps the focus on learning rather than on the game?
3. How should a course built this way be evaluated to find out whether it works, and for whom?

**Definitions**

The terms below are used throughout. Several of them were used loosely in the 2018 edition, so they are defined here.

- **Game.** A player-engaged system of goals, rules, and feedback. Hunicke et al. (2004) break a game into *mechanics* (the rules and components), *dynamics* (how those rules behave when people play), and *aesthetics* (the experience the player has).
- **Game-based learning (GBL).** Learning in which the game itself is the learning environment, so the content is taught through play (Plass et al., 2015). This thesis applies GBL to course delivery rather than to a stand-alone video game.
- **Gamification.** Adding game elements such as points, badges, or leaderboards to an activity that is not a game. Sailer and Homner (2020) define it, following Deterding and colleagues, as "the use of game design elements in non-game contexts" (p. 78). Plass et al. (2015) draw the same line: GBL designs the learning experience as play, while gamification adds game pieces to an existing task.
- **Flow.** A state of deep concentration in an activity where challenge and skill are balanced (Csikszentmihalyi, 2014). Chen (2006) summarizes eight components of flow, including clear goals, direct feedback, and a sense of control.
- **Scaffolding.** Temporary support that helps a learner do what they cannot yet do alone, such as a worked example, a conceptual briefing, a hint, or guided reflection (Barzilai & Blau, 2014; Clark et al., 2016).
- **Dynamic difficulty adjustment (DDA).** Changing the challenge of an activity to match the player. Chen (2006) separates *passive* DDA, where the system

changes difficulty for the player, from *active* DDA, where the player makes meaningful choices that set their own level of challenge.

- **Game Master (GM).** In a tabletop RPG, the person who runs the game, describes the world, applies the rules, and responds to the players. In this design, the instructor plays that role.
- **Player Character (PC).** The character a student plays within the course narrative.
- **Universal Design for Learning (UDL).** A framework for designing learning so that it works for the widest range of learners from the start, organized around multiple means of engagement, representation, and action and expression (CAST, 2024).

## Scope

This thesis focuses on the mechanics of traditional tabletop RPGs such as Dungeons and Dragons, delivered through tools a college already uses: a learning management system (LMS), synchronous video sessions, and a virtual tabletop or text-based story platform. It does not propose building a commercial-quality digital game. Digital game-based learning is well studied, and much of that research informs this design, but tabletop mechanics have advantages for a college course. They can be run by an instructor without a development team, and a human Game Master can adjust the experience in ways software still struggles to (Chen, 2006). The example course is an American history selective at a community college, delivered in both a fully online section and a hybrid section.

**Why This Topic, and Why Me**

I did not come to game-based learning from a background in game design. My undergraduate degree is in Community and Human Services. Before and during graduate school I worked as a domestic violence advocate and trained mediator, and then for several years as a substance abuse and problem gambling prevention specialist assigned to the Herkimer College campus. I have taught at the college as an adjunct since 2012, including Introduction to Technical Writing, First-Year Seminar, Wellness, and Chemical Dependency, and in 2015 I became an academic advisor and manager of the college's Starfish student success system. That was my job when I wrote the 2018 thesis. Three things from that earlier work show up in this thesis.

The first is role-play. In a Human Service Ethics course in 2011, my group worked through an end-of-life dilemma in which I played the human service worker. What I wrote afterward was that the role-play "taught me empathy," and that it was fun "having to get into the head of my client too" (Parkinson, 2011). I did not have a theory for it then, but I had experienced what the moral development research in Chapter 2 describes: taking on another person's role changes how you reason about a problem.

The second is facilitation. In group work training I learned that a group leader's job is not to have every answer but to "facilitate a medium that is productive" for the group (Parkinson, 2010). That is also a good description of a Game Master, and it is the role this design asks of the instructor.

The third is restraint. As a case worker I was taught that when writing case notes, less is best: include the important, include the relevant, and leave out the rest. In 2014, designing my first game-based pilot course, I applied that same rule to multimedia and called the extra material

"flare" that distracts learners (Parkinson, 2014d). I lost some of that restraint in the 2018 edition, which piled on points, boons, platforms, and extra credit. This edition puts it back.

I am also a gamer. When I started my graduate program I wrote that "if I was able to be consumed by games, imagine what it could do with students" (Parkinson, 2014b). That enthusiasm is still here. The difference now is that I want the design to earn it.

**What I Have Learned Since 2018**

In 2020 I became Director of Instructional Design at Herkimer County Community College. That role changed how I read my own thesis, in four ways.

The first is durability. Leading the college's move from Blackboard to D2L Brightspace, and supporting faculty through it, taught me that a course design has to survive the tools it is built on. Platforms get replaced, licenses end, and companies get acquired. A design that depends on one product is fragile, which is why Chapter 4 makes the LMS home base and treats every other tool as optional.

The second is quality. Reviewing courses each semester against SUNY's OSCQR course quality rubric has made me pay attention to the basics the 2018 design took for granted: clear navigation, clear expectations, aligned assessments, and materials students can actually use. A game layer does not make up for a course that misses those basics. It makes them more important, because students have more to keep track of.

The third is access. As co-chair of the college's Accessibility Committee, and as the college's designee for its accessibility compliance work, I have spent the last several years helping faculty and staff prepare course content for the web accessibility requirements that now apply to public colleges. That work made the absence of accessibility in the 2018 thesis

impossible to ignore. Chapters 3 and 4 now treat access as a design requirement from the start, not an accommodation added later.

The fourth is AI. Since generative AI tools became widely available, I have served as my college's lead on AI integration, written its AI ethics and digital assessment policies, and worked on an Introduction to AI course. That work, and the questions it raises about academic integrity, about what students should learn to do with these tools, and about what should stay human, shaped the cautious role Chapter 6 gives AI in this design.

I have also kept building. Since 2024 I have been developing GamED, a game-based learning platform grounded in the same theories this thesis draws on, including flow and self-determination theory. Building it has put the lessons of this edition in front of me in a very practical way: it is easy to make something engaging, and much harder to make sure the engagement leads to learning.

**Organization of the Thesis**

Chapter 2 reviews the research on how people learn, how games work as designed systems, motivation and flow, the evidence on game effectiveness, reflection, role-play and moral development, rewards, assessment, accessibility, and emerging AI tools. Chapter 3 turns that research into a design framework of nine principles and a repeating Brief, Play, Debrief cycle. Chapter 4 applies the framework to a hybrid/online American history module. Chapter 5 presents an evaluation plan. Chapter 6 discusses what changed in my thinking, the limitations of this work, and directions for future development. The appendices include a change log against the 2018 edition, the 2018 reader comments, sample course materials, a rubric, a debrief protocol, an accessibility checklist, sample evaluation items, and the history of the MyQuest prototype.

# Chapter 2: Literature Review

This chapter reviews the research that the design in Chapters 3 and 4 is built on. The 2018 edition moved quickly from game design frameworks to course mechanics. This edition starts one step earlier, with how people learn, because a game-based course is only as good as the learning theory under it. Plass, Homer, and Kinzer (2015) argue that games cannot be understood from a single perspective of learning, and that cognitive, motivational, affective, and sociocultural perspectives together are needed "to fully capture what games have to offer for learning" (p. 258). The chapter follows that advice. It moves from learning theory, to games as designed systems, to motivation and flow, to the evidence on what games do and do not accomplish, and then to the specific issues a role-playing course has to handle: reflection, moral reasoning and community, rewards, assessment, accessibility, and new AI tools.

## How People Learn: The Foundation Under the Design

### *Constructivist and Sociocultural Views*

When I first wrote about learning theory in graduate school, I compared Piaget and Vygotsky. Piaget's constructivism describes learners building knowledge by fitting new information into what they already know, or changing what they know to fit new experience. Vygotsky's sociocultural theory holds that social interaction comes first, and that a community plays a central role in making meaning. I wrote then that Piaget "has the foundation" while Vygotsky "has the structure of what it is to learn in today's society" (Parkinson, 2014a). I still think that is right, and it matters for this design because a role-playing course is social by nature. Students learn in parties, not alone.

Game researchers have built directly on Vygotsky. Barab, Thomas, Dodge, Carteaux, and Tuzun (2005) traced the theoretical framework of Quest Atlantis, a multiuser learning game, to

Vygotsky's work, drawing on human development, social constructivism, the social context of thought and activity, and the zone of proximal development. D'Aprile, Di Bitonto, De Asmundis, and Ulloa Severino (2015) went further and argued that classic needs and task analysis is not enough for designing game-based learning environments. In their view, socio-constructivist theory has to be part of the design, because learning in a game is tied up with the activities performed, the tools used, and the way knowledge is built through interaction. Their pilot of a nutrition game with 65 young learners reported on usability and engagement, and they concluded that game-based learning communities can combine online and face-to-face communication to improve learning through collaborative participation.

***Situated Cognition***

The term "situated meaning" appeared all through the 2018 edition without the theory behind it. That theory is situated cognition. Brown, Collins, and Duguid (1989) argued that much of schooling assumes knowledge can be pulled out of the situations where it is learned and used, and that this assumption limits learning. Knowledge, in their account, is partly a product of the activity, context, and culture in which it develops. They compared conceptual knowledge to a set of tools: a tool can only be fully understood through use, and it is possible to own a tool and not know how to use it, just as students can memorize definitions that sit inert. As an alternative, they proposed cognitive apprenticeship, in which learners take part in the authentic activities of a community, with support, the way an apprentice learns a trade.

I summarized a classic example of this in 2014: children who sold coconuts on the street could calculate prices in their heads while selling, but could not solve the same arithmetic when it was written as a school problem (Parkinson, 2014c). The math was the same. The situation was not. Oliver and Herrington (2001), in their guide to online teaching, apply the same idea to

course design. They recommend authentic tasks, meaning tasks similar to those faced in real life, which tend to be complex and ill-defined and rarely have a single right answer.

Shaffer, Squire, Halverson, and Gee (2004) connected situated learning to games directly. Communities of practice share ways of knowing, acting, and valuing, which they call an epistemic frame, and games can let players try on that frame by acting as a member of the community. Gee (2008) described the same idea as a "Situated Learning Matrix." A game like SWAT 4 does not start by teaching all of its content. It gives the player an identity, a SWAT team member with that role's goals and norms, and a tutorial that teaches just enough for the player to learn the rest by playing.

For this thesis, situated cognition does two things. It explains why role-play is worth the effort: acting as a historical agent who must use evidence is closer to the work of history than answering quiz questions about it. It also sets a standard. If the activity in the game is not the kind of thinking the course is supposed to teach, the situation is teaching the wrong thing.

## Games as Designed Systems

### *The MDA Framework*

Hunicke et al. (2004) broke games into mechanics, dynamics, and aesthetics. Mechanics are the rules and components. Dynamics are how those rules behave when players act on them over time. Aesthetics are the experiences the player has, such as challenge, fellowship, discovery, or narrative. A designer works from mechanics toward aesthetics, while a player experiences the game from aesthetics back toward mechanics. This is useful for course design because it separates what an instructor controls (the rules of an assignment, the structure of a session) from what an instructor hopes happens (curiosity, collaboration, persistence). An

instructor cannot assign fellowship. An instructor can design mechanics, such as interdependent team roles, that make fellowship likely.

### *Why Games Engage*

Prensky (2001) listed twelve reasons games engage players. Games are fun and are a form of play; they have rules, goals, interaction, and outcomes with feedback; they adapt to the player; they have win states, conflict or challenge, and problem solving; they involve other people; and they have representation and story. Gee (2005) described the learning principles that good games build in. Several of them are central to this design:

- **Identity.** Deep learning requires an extended commitment of self. Good games capture players by letting them inhabit or build a character, and learning a domain likewise means taking on the identity of people who do that work.
- **Customization.** Players can often adjust a game to fit their learning and playing styles.
- **Risk taking.** Good games lower the consequences of failure, so players explore and try new things, and failure becomes information.
- **Just in time and on demand.** Games give information right when players need it, or when players ask for it, rather than all at once up front.
- **Situated meanings.** Words and concepts are learned through the situations and actions they refer to, not through definitions alone.
- **Cross-functional teams.** In multiplayer games, each player masters a specialty but has to understand the others well enough to coordinate.

These principles are the reason tabletop RPGs are a good fit for college courses. A tabletop party is a cross-functional team by design. Characters are identities players build. A

Game Master gives information when the story calls for it. Failure is expected and survivable. Chapter 3 keeps these principles, but it treats them as design goals that need evidence behind them, not as proof that a course will work.

## Motivation, Flow, and Engagement

### *Self-Determination Theory*

Self-determination theory explains motivation through three basic psychological needs: autonomy (a sense of choice and volition), competence (a sense of effectiveness), and relatedness (a sense of connection with others). Ryan, Rigby, and Przybylski (2006) applied the theory to video games in four studies. Perceived in-game autonomy and competence were associated with enjoyment, game preference, and changes in well-being after play, and in a survey of an online multiplayer community, autonomy, competence, and relatedness each independently predicted enjoyment and intention to keep playing. Olson (2010), reviewing children's motivations for play, described related findings: when games provided opportunities for competence, mastery, and meaningful choices, violent content added little to how much players enjoyed or valued a game.

This matters because it explains why games motivate without depending on points or prizes. Players come back because the game lets them make choices that matter, get better at something, and do it with other people. That is also a good description of what a course should offer.

### *Flow*

Flow describes deep concentration in an activity where challenge and skill are in balance (Csikszentmihalyi, 2014). When the balance breaks, people drift into boredom when the challenge is too low, anxiety when it is too high, or apathy when both are low (Admiraal et al.,

2011). Chen (2006), drawing on Csikszentmihalyi, lists eight components of flow: a challenging activity that requires skill, the merging of action and awareness, clear goals, direct feedback, concentration on the task, a sense of control, loss of self-consciousness, and a changed sense of time. Not every component has to be present.

Chen's contribution for this thesis is his approach to difficulty. Players differ, so a single difficulty level cannot keep everyone in flow. *Passive* dynamic difficulty adjustment has the system watch the player and change the game, but games cannot directly read what a player is thinking, so these systems guess from incomplete data. *Active* adjustment gives players meaningful choices of challenge and embeds those choices into play, so players steer their own experience. Chen's recommendation was to widen the range of activities and difficulties, let players choose their pace, and build those choices into the core of the game (Chen, 2006). A human Game Master can do both kinds of adjustment. The instructor can notice a struggling student and adjust, and the course can offer students choices of challenge.

### *Engagement and Challenge as Predictors of Learning*

Hamari, Shernoff, Rowe, Coller, Asbell-Clarke, and Edwards (2016) surveyed 173 players of two physics learning games, one used in high school classrooms and one used across a semester of an undergraduate engineering course. Engagement in the game had a clear positive effect on perceived learning, but immersion did not. Challenge had a positive effect on learning both directly and through increased engagement, and it was an especially strong predictor of learning. Skill affected learning only by increasing engagement. The authors concluded that the challenge of an educational game should keep up with learners' growing abilities.

Fu, Su, and Yu (2009) built a way to measure this experience. Their EGameFlow scale measures learner enjoyment of e-learning games across eight dimensions: immersion, social

interaction, challenge, goal clarity, feedback, concentration, control, and knowledge improvement. They validated it with 166 students in a university online course and found its validity and reliability satisfactory. Chapter 5 adapts these dimensions for evaluating the course.

**What Flow Does Not Do**

The 2018 edition quoted Admiraal et al. (2011) to argue that flow "invokes a growth principle... leading to long-term learning effects." That sentence comes from the closing paragraph of their article, where the authors offer it as a promising direction for future research. Their own findings point somewhere more complicated, and the design has to account for them.

Admiraal and colleagues studied Frequency 1550, a one-day game about medieval Amsterdam that 216 secondary students played in teams, with one team moving through the city and another working from headquarters. Students generally experienced flow, although technology and navigation problems distracted them. The key finding was this: "Team flow appears to be related to the group performance in the game, but not with student learning outcome" (Admiraal et al., 2011, p. 1192). They noted that this matched other studies in which students were more engaged but did not show greater conceptual understanding. At the same time, some game activities did relate to learning. The less students were distracted by technical problems, and the more engaged they were in competition between groups, the more they learned. The authors also noted that if scaffolding is used to help students do what they cannot do alone, the challenge still needs to be within their reach.

I summarized a second study for my research methods course in 2017 that makes the same point from a different angle (Parkinson, 2017a). Barzilai and Blau (2014) compared three conditions in an online financial math game: play only, study and then play, and play and then study. The group that studied a conceptual scaffold before playing solved problems better than

the other two groups. Flow predicted students' perceived learning and enjoyment, but perceived learning and enjoyment did not predict problem-solving performance. Adding the scaffold also did not reduce the flow or enjoyment students reported. In other words, students who feel like they learned a lot are not necessarily the ones who learned the most, and a short briefing before play does not spoil the fun.

Erhel and Jamet (2013) looked at how a game is framed and what feedback it gives. In their first experiment, learners told to treat a medical learning game as a learning activity showed deeper comprehension than learners told it was entertainment, with no drop in motivation. In a second experiment, they added knowledge-of-correct-response feedback. With that feedback, learners in the entertainment condition did better on comprehension than those in the learning condition. Feedback that tells learners what they got right and wrong helped them process the content more deeply, even when they thought they were playing.

Taken together, these studies do not say flow is unimportant. Engagement and challenge are clearly linked to learning (Hamari et al., 2016). They say that flow is not enough. Learning comes from engagement combined with a briefing before play, feedback during play, and a clear connection between the game and the content. The 2018 design had pieces of this, including a tutorial, a story recap, and a final report, but it never explained that those pieces are what make the game educational. This edition does.

## The Evidence on Game-Based Learning

### *Systematic Reviews and Meta-Analyses*

Connolly, Boyle, MacArthur, Hainey, and Boyle (2012) reviewed 129 papers on the impacts of computer games and serious games for players 14 and older. Games were linked to a wide range of perceptual, cognitive, behavioral, affective, and motivational outcomes, and the

most frequent were knowledge acquisition and content understanding, along with affective and motivational outcomes. The studies were so varied in their questions, theories, and methods that the authors could not combine them statistically, and relatively few used randomized controlled designs.

Ke (2009) reviewed 89 instructional gaming studies. Earlier reviews had found no clear causal link between games and academic performance, and many studies described positive effects while few described negative ones. The finding most relevant here came from studies of instructional game design: instructional support features are a necessary part of instructional games, and learners without that support "learn to play the game rather than learn domain-specific knowledge embedded in the game" (Ke, 2009, p. 21). Yaman, Nerdel, and Bayrhuber (2008) found something similar in computer simulations. The type of instructional support mattered: learners gained factual knowledge either way, but gains in understanding depended on the kind of support they received.

Clark, Tanner-Smith, and Killingsworth (2016) conducted the most rigorous synthesis in this review. They included only randomized controlled trials and controlled quasi-experiments with K-16 students. Digital games improved learning compared with non-game conditions ($g = 0.33$), and augmented game designs outperformed standard versions of the same games ($g = 0.34$). The largest group of augmented designs was enhanced scaffolding, which outperformed the standard versions ($g = 0.41$). Games played over multiple sessions showed better learning than non-game conditions, while single-session games did not. When game characteristics were controlled, single-player games without competition and collaborative team competition both outperformed single-player competitive games. Their conclusion is the one this thesis is built around: games "provide new and powerful affordances, but it is the design within the medium to

leverage those affordances that will determine the efficacy of a learning environment" (Clark et al., 2016, p. 116).

Two more findings from Clark et al. (2016) are cautions for this design. Schematic and text-based games produced larger effects than realistic ones, and games with story lines that were unrelated to the learning content produced larger effects than games with related story lines. Both differences shrank to nonsignificance, or near it, when other game characteristics were controlled, so neither is a rule. They are still a warning that richer graphics and more elaborate stories do not automatically mean more learning.

### *Studies Close to This Design*

Several studies in my research folder are close to the context of this thesis. Papastergiou (2009) compared a computer memory game with an equivalent non-game website in high school computer science. The game was more effective for learning and more motivating, and it was equally effective for boys and girls. Barr (2017) randomly assigned undergraduate arts and humanities students to play commercial video games over eight weeks or to a control group. The game group showed large improvements on self-report measures of communication, adaptability, and resourcefulness. Adachi and Willoughby (2013) followed 1,492 adolescents over four years and found that strategic games, including role-playing and strategy games, predicted higher self-reported problem-solving skills, which in turn predicted higher grades.

### *What Teachers Need*

Perrotta, Featherstone, Aston, and Houghton (2013) reviewed the evidence for teachers and school leaders and offered practical guidance. Game-based learning can improve engagement and motivation, but they cautioned against relying on games alone to raise attainment. They recommended using games within a clear teaching process: keep the academic

content inside the game's fiction and rules rather than bolting it on, and plan the roles of teacher and learners carefully. Teachers "should play roles that allow them to mediate the experience for learners" (p. 27), which they describe as giving guidance when needed, making sure rules are followed, and keeping a respectful atmosphere. They also advised against separating badges, scores, and leaderboards from the game's fiction and rules, and they warned that teachers need time, preparation, and support to do this well. Van Eck (2006) made a related point years earlier: if advocates keep insisting only that games can be effective, they risk giving the impression that all games are good for all learners and all outcomes.

**Reflection and Metacognition**

If flow is not enough, reflection is one of the main things that turns play into learning. Kim, Park, and Baek (2009) had ninth graders play a commercial online role-playing game while using three metacognitive strategies: self-recording (writing down plans and activities), modeling (watching others), and thinking aloud (discussing game play with peers during breaks). Thinking aloud was the strongest influence on students' social problem-solving ability, followed by modeling, with individual self-recording the weakest. Social problem solving, in turn, strongly affected both learning achievement and game achievement. The more social the reflection was, the more it helped.

Plass et al. (2015) connect reflection to failure. They argue that games allow graceful failure, where failure is "by design an expected and sometimes even necessary step in the learning process" (p. 261). Lower stakes encourage risk taking, and play creates chances for self-regulated learning: setting a goal, monitoring progress, and judging whether a strategy worked. Reflection is where that self-regulation becomes visible, both to the student and to the instructor.

## Role-Play, Moral Reasoning, and Community

### *Moral Development*

Kohlberg and Hersh (1977) described moral reasoning as developing through three levels, each with two stages. At the preconventional level, right and wrong are judged by consequences such as punishment and reward. At the conventional level, meeting the expectations of family, group, or nation is valued for its own sake. At the postconventional level, people reason from principles that hold apart from any particular group's authority. What moves a person through these stages is social interaction. Relationships "demand that one take others' perspectives (role-taking)," and moral development comes from reworking those role-taking experiences into more adequate ideas of justice (Kohlberg & Hersh, 1977, p. 57). Encountering moral conflicts that one's current reasoning cannot resolve pushes that development forward. The authors recommended classroom moral discussion, but warned that discussion alone becomes pedantic and should be part of a broader involvement of students in real problems of their community.

Wright, Weissglass, and Casey (2017) tested whether imaginative role-play could support that development. Twelve college students played six sessions of about four hours each in a customized Dungeons and Dragons game with embedded moral dilemmas, such as whether to torture a prisoner for information, which they had to work through as a group. The players showed significant growth in moral development on the Defining Issues Test and the Self-Understanding Interview, and two control groups did not. The authors concluded that role-play games can function as an engaging "moral training ground" (p. 1). The sample was small, so the study is promising evidence rather than proof, but it matches Kohlberg's account closely. Players

took on roles, faced dilemmas their current reasoning could not settle, and worked them out with others.

Adams (2013) analyzed how Dungeons and Dragons players talk about their games and found shared fantasy themes related to democracy, friendship, extraordinary experiences, and ethics. My own experience fits here too. In the 2011 ethics role-play described in Chapter 1, taking on a role I disagreed with forced me to separate my personal values from my professional responsibility, and what I took away was empathy (Parkinson, 2011).

### *Community and Connection*

The 2018 draft had a section on community that was cut from the final version, although its sources stayed in the reference list. Community belongs in the design. MacDonald and Thompson (2005) found that quality online learning depended on structure, content, delivery, service, and outcomes working together, and that learners were searching for connection, valued interactions, and built content socially. Barab et al. (2005) described their approach to Quest Atlantis as socially responsive design, building structures that engage with and potentially transform learners and the contexts they participate in. They also described a balance that applies directly here: they gave neither entertainment nor education priority, because favoring education would lose the students and favoring entertainment would lose the teachers who decide whether the game is used at all.

Lenhart, Kahne, Middaugh, Macgill, Evans, and Vitak (2008) surveyed teens nationally and found that those with the most civic gaming experiences, meaning game play that parallels civic learning opportunities such as helping others and debating ethical issues, were more likely to report interest and engagement in civic and political activities. One detail is important for an online course. Playing games with others in person was related to civic and political outcomes,

but playing with others online was not. That does not mean online play cannot build community. It means an online course should not assume community will happen just because students are logged in at the same time. It has to be designed.

**Rewards, Points, and Gamification**

The 2018 design gave extra credit, called "luck points," for interacting with the LMS, and extra credit for character backstories. My earlier thesis ideas went further. My 2017 plan for a campus game proposed a leaderboard with prizes such as a TV or a game console for the student with the most points (Parkinson, 2017c). The research gives good reasons to rethink this.

Sailer and Homner (2020) conducted a meta-analysis of gamification in learning. Gamification had small positive effects on cognitive ($g = .49$), motivational ($g = .36$), and behavioral ($g = .25$) learning outcomes, and only the cognitive effect held up in the most rigorous studies. Two design choices made a difference. Including game fiction, meaning a story or narrative frame, and combining competition with collaboration were especially effective for behavioral outcomes, and competition combined with collaboration may also help motivational outcomes. Clark et al. (2016) found that single-player competitive games were the least effective arrangement for learning. Perrotta et al. (2013) advised against pulling badges, scores, and leaderboards away from the fiction and rules of the game.

Self-determination theory helps explain why. The needs that predicted enjoyment and continued play were autonomy, competence, and relatedness (Ryan et al., 2006). Points for opening an announcement do little for any of the three, while team goals, meaningful choices, and recognition of growing skill speak to all of them. In 2014 I wrote that intrinsic motivation is what brought me to graduate school and that a degree was "just a bonus" (Parkinson, 2014a). A course designed around that belief should not pay students in points to open announcements.

Recognition in this edition is built into the story and shared by teams, and it rewards the skills the course is teaching.

**Assessing Choices, Not Just Knowledge**

My second reader asked a hard question about the 2018 design. If students in a history role-play can make different choices than the real historical figures made, does that confuse the facts? And if they cannot, are they really role-playing? (C. T. Avis, personal communication, December 28, 2018). Assessment research offers a way through.

Schwartz and Arena (2009) argued that most assessments measure how much knowledge students have accumulated, while a basic goal of education is preparing students to make good choices on their own. They proposed choice as the construct assessment should be built around, and argued that interactive digital environments make it possible to see what students choose to learn and do. In a history role-play, the most important thing to assess is not whether students' characters matched history. It is whether students can explain, with evidence, why history went the way it did and what would have had to be different for it to go another way. That is a historian's reasoning, and it matches the SUNY General Education outcome that students "apply historical and contemporary evidence to draw, support, or verify conclusions" (State University of New York [SUNY], n.d.). Chapter 4 uses this to answer the reader's question.

**Designing for Every Student**

***Universal Design for Learning***

Accessibility did not appear anywhere in the 2018 edition. That was a gap. In 2014, while designing my first game-based pilot course, I wrote that learning works best when it is multimodal, with material presented in multiple forms and students given multiple ways to interact with it and show what they know (Parkinson, 2014d). The current version of that idea is

CAST's Universal Design for Learning Guidelines 3.0 (CAST, 2024). Several of the guidelines read almost like a description of good game design: optimize choice and autonomy (7.1), nurture joy and play (7.3), optimize challenge and support (8.2), foster collaboration and collective learning (8.3), foster belonging and community (8.4), offer action-oriented feedback (8.5), and promote individual and collective reflection (9.3). Others are warnings for a game-based course: support multiple ways to perceive information (1.2), vary and honor the methods for response, navigation, and movement (4.1), and optimize access to accessible materials and assistive technologies (4.2). A course that depends on a visual map, fast voice chat, and timed decisions will lock out some students unless it is designed not to.

### *Cognitive Load*

Plass et al. (2015) note a real tension in game design between reducing cognitive load and increasing visual appeal. Every added element competes for a learner's limited attention. Clark et al.'s (2016) finding that simpler, schematic games did at least as well as realistic ones points the same direction. My own rule from case work, include the important, include the relevant, and leave out the rest, is a fair summary of what this research recommends (Parkinson, 2014d).

### *The Legal Requirement*

For a public college, accessibility is also the law. In 2024 the U.S. Department of Justice adopted a rule under Title II of the Americans with Disabilities Act that sets WCAG 2.1 Level AA as the technical standard for the web content and mobile apps of state and local government entities, which include public colleges. An interim final rule in April 2026 extended the compliance dates: public entities with a total population of 50,000 or more must comply by April 26, 2027, and smaller entities and special district governments by April 26, 2028 (Extension of Compliance Dates, 2026). Course materials, story documents, recorded sessions, and third-party

game platforms used in a public college course all fall within this conversation. Any platform chosen for a game-based course has to be checked for accessibility before it is adopted, not after a student requests an accommodation.

**Emerging Tools: Generative AI**

Much has changed since 2018, and the most recent research in my folder is about generative AI. Narreddy, Joordens, and Prompiengchai (2025) reviewed large language models (LLMs) for formative assessment in higher education, including in game-based learning and storytelling. LLMs can support personalized feedback, adaptive goal setting, and instructor planning, but the authors identified challenges including cost, privacy and security, technical bias, the rapid pace of change, and the need for teacher training and for empirical evaluation. Marengo, Pagano, Lund, and Santamato (2025) reviewed studies combining AI and gamification in higher education and found potential for personalized learning and for assessing soft skills such as collaboration and leadership, but noted limited long-term evidence, concerns about data privacy and technological equity, and findings drawn mostly from well-resourced settings. Pang and Wei (2025) analyzed 67 case studies from a 2024 AI in education conference and reached similar conclusions about promise and about accuracy, bias, and ethical concerns.

Two studies speak to design. Sadik (2025) redesigned a software engineering final project so students used a generative AI tool to design educational games, positioning the tool as an assistant rather than a replacement for student thinking, with explicit attention to ethical and responsible use. Alharthi (2025) surveyed and interviewed game design professionals and found that generative AI speeds up ideation and prototyping but raises concerns about originality and creative dependence. SUNY now expects students to understand these issues. The revised Information Literacy outcome, effective Fall 2026, asks students to demonstrate an

understanding of the ethical dimensions of information use and creation, "whether from traditional sources or emerging technologies, such as artificial intelligence" (SUNY, n.d.). Chapter 6 considers a limited role for AI as an assistant to the Game Master, with a human in charge.

**Summary**

The research supports the core idea of the 2018 thesis. Games can be effective learning environments, role-play can support perspective-taking and moral reasoning, and the social structure of a tabletop party fits how people learn. The research also changes the design in five important ways.

1. Engagement and flow matter, but they do not produce learning on their own. Scaffolding before play, feedback during play, and reflection after play are what connect the game to the content (Admiraal et al., 2011; Barzilai & Blau, 2014; Clark et al., 2016; Erhel & Jamet, 2013; Ke, 2009; Kim et al., 2009).
2. The activity inside the game has to be the thinking the course is meant to teach, which is the central claim of situated cognition (Brown et al., 1989; Gee, 2008; Shaffer et al., 2004).
3. Challenge should be tuned to each learner, and learners should have meaningful choices of challenge (Chen, 2006; Hamari et al., 2016).
4. Collaboration and in-story recognition are better motivators than individual competition and detached points (Clark et al., 2016; Perrotta et al., 2013; Ryan et al., 2006; Sailer & Homner, 2020).

5. A game-based course has to be designed for every student and meet accessibility requirements, and its effects need to be measured rather than assumed (CAST, 2024; Fu et al., 2009; Van Eck, 2006).

Chapter 3 turns these findings into a design framework.

# Chapter 3: A Design Framework for Scaffolded Story Play

This chapter turns the research in Chapter 2 into a design framework. I call it Scaffolded Story Play. The name is meant to hold both halves of the argument: the course is built around a story students play, and that play is scaffolded so it leads to learning. The framework has two parts. The first is a repeating module cycle of Brief, Play, Debrief, Write History, and Level Up. The second is a set of nine design principles that decide what goes into each part of the cycle. The 2018 edition listed Gee's GBL principles and Lidwell, Holden, and Butler's (2010) design principles side by side. This framework keeps many of those ideas, but each principle here is tied to evidence and to a specific decision in the course.

## Where the Framework Comes From

Three findings from the literature carry most of the weight. First, design matters more than medium. Games can support learning, but it is the design inside the game that determines whether they do (Clark et al., 2016). Second, engagement is necessary but not sufficient. Students can be absorbed in play without learning the content, and they learn the content when play is paired with instructional support, feedback, and reflection (Admiraal et al., 2011; Barzilai & Blau, 2014; Erhel & Jamet, 2013; Ke, 2009). Third, people learn in situations and with others. The activity inside the game has to be the thinking the course is meant to teach, and the social structure of the game has to make students need each other (Brown et al., 1989; D'Aprile et al., 2015; Kim et al., 2009).

A tabletop RPG already has a structure that fits those findings, if it is used deliberately. Before a campaign, good tables hold what players call a session zero, where the group agrees on expectations, tone, and boundaries. Before each session, the Game Master sets the scene. During

play, players make choices and deal with consequences. After play, groups often talk through what happened. The framework names those moments and gives each one a job.

## The Module Cycle

Each learning module runs through the same cycle. Students learn the rhythm early, which reduces the load of figuring out what to do next and leaves more attention for the content. Table 1 summarizes the cycle, what happens in each phase, and the research behind it.

**Table 1**

*The Scaffolded Story Play Module Cycle*

| Phase | What happens | Why (evidence) |
|---|---|---|
| Session Zero (once, at course start) | Course tutorial; roles, expectations, and boundaries agreed; opt-in choice between Story Play and the standard path; accessibility and communication preferences collected | Tutorial teaches just enough to begin (Gee, 2008); game-based learning is not for every student (Admiraal et al., 2011); design for access from the start (CAST, 2024) |
| Brief | Short conceptual briefing and primary sources before play; mission objectives stated in plain terms | Studying a scaffold before play improved problem solving without reducing flow (Barzilai & Blau, 2014); learning framing deepens comprehension (Erhel & Jamet, 2013) |
| Play | Synchronous or asynchronous story session; teams make choices using evidence; Game Master mediates | Situated activity and cross-functional teams (Brown et al., 1989; Gee, 2005); teacher mediates play (Perrotta et al., 2013) |
| Debrief | Guided group discussion right after play; think-aloud about decisions; feedback on what was right and wrong | Thinking aloud with peers strengthened social problem solving and achievement (Kim et al., 2009); corrective feedback supports deep processing (Erhel & Jamet, 2013) |
| Write History | Team and individual reports that compare the team's choices with the historical record, using evidence | Assess reasoning about choices (Schwartz & Arena, 2009); SUNY GE evidence and critical thinking outcomes (SUNY, n.d.) |
| Level Up | Instructor reviews performance and pulse survey; adjusts challenge; students choose next mission tier | Challenge predicts learning (Hamari et al., 2016); active and passive difficulty adjustment (Chen, 2006) |

*Note.* Session Zero happens once. The other five phases repeat in every module.

## Nine Design Principles

### *Principle 1: Brief Before Play*

Every session starts with a short briefing that gives students the concepts and sources they will need, and a clear statement of what the mission is supposed to teach. Barzilai and Blau (2014) found that students who studied a conceptual scaffold before playing solved problems better than students who only played or who studied afterward, and the scaffold did not reduce their flow or enjoyment. Clark et al. (2016) found that games with enhanced scaffolding outperformed the same games without it. Ke (2009) found that learners without instructional support tend to learn the game instead of the content. The 2018 design had a tutorial at the start of the course but nothing like this before each session. The briefing is short on purpose, and it is framed as mission intelligence inside the story, but its job is instructional.

### *Principle 2: Make the Content the Mechanic*

The actions students take in the game should be the thinking the course is supposed to teach. In a history course, that means gathering evidence, weighing sources, placing events in context, and arguing from evidence. Perrotta et al. (2013) recommend making academic content integral to the game rather than an add-on, and embedding tasks in the game's fiction and rules. Brown et al. (1989) argue that knowledge learned apart from its use tends to sit inert. Gee (2005) calls this situated meaning. The test for any mechanic is simple: if a student could succeed at it without doing historical thinking, the mechanic needs to change. Clark et al.'s (2016) finding that a relevant story line did not automatically improve learning is a reminder that a history-themed story is not the same thing as history-shaped play.

***Principle 3: Tune Challenge, Do Not Chase Flow***

The goal is not to keep students in flow. The goal is to keep challenge at the edge of what each student can do, because challenge and engagement predict learning (Hamari et al., 2016) and because flow alone does not (Admiraal et al., 2011). The course does this in two ways that come from Chen (2006). Passive adjustment is the instructor noticing, from performance and a short pulse survey, that a team is bored or overwhelmed, and changing the next mission. Active adjustment is giving students meaningful choices of challenge, which in this course are mission tiers students choose each module. Chen's advice was to embed those choices into play so they feel like part of the game, not like a settings menu. UDL's guideline to optimize challenge and support (CAST, 2024, 8.2) points the same way.

***Principle 4: Fail Gracefully, With Feedback***

Good games lower the cost of failure so players will take risks (Gee, 2005), and Plass et al. (2015) describe failure in games as an expected and sometimes necessary step in learning. The 2018 design kept this idea with its Time Agents, who can reset time up to three times after a failed major event. This edition keeps the resets and adds the part that makes failure useful: feedback. Erhel and Jamet (2013) found that feedback on correct and incorrect responses helped learners process content more deeply. A failed mission in this course always comes with specific feedback about which evidence was missing or misread, and a side quest aimed at that gap before the team tries again. The UDL guideline to offer action-oriented feedback (CAST, 2024, 8.5) describes this well.

***Principle 5: Debrief Every Session***

Play is followed by a structured debrief. Kim et al. (2009) found that talking through game play with peers had the strongest effect on social problem solving of the three reflection

strategies they tested, and that social problem solving strongly influenced learning. Kohlberg and Hersh (1977) describe perspective-taking and working through moral conflict with others as the engine of moral development, and classroom discussion as one of the ways schools can support it. The debrief is where students connect what their characters did to what actually happened, and where the instructor corrects misconceptions before they settle. It is required, it is short, and it happens as close to play as possible. Appendix D provides the protocol.

***Principle 6: Build a Table, Not a Leaderboard***

Students succeed as teams, and recognition is shared and built into the story. The course has no public individual leaderboard and no points for routine LMS activity. Clark et al. (2016) found that single-player competitive arrangements were the least effective for learning. Sailer and Homner (2020) found that combining competition with collaboration and including game fiction made gamification more effective. Perrotta et al. (2013) advise against pulling scores and badges away from the fiction and rules of the game. Self-determination theory adds that the needs driving enjoyment and persistence are autonomy, competence, and relatedness (Ryan et al., 2006). Teams in this course can still compete in friendly ways, such as which cell files the most complete report, but the competition is between teams that depend on their own members, which is the arrangement the research supports.

***Principle 7: The Instructor Is the Game Master***

The instructor's role is to run the game in service of learning. My second reader asked me to define that role in 2018, and the research defines it well. Perrotta et al. (2013) say teachers should take roles that let them mediate the experience: giving guidance when needed, making sure rules are followed, and keeping a respectful atmosphere. Brown et al.'s (1989) cognitive apprenticeship describes an expert who models, coaches, and gradually fades support. My own

training in group work described a leader's job as facilitating a productive space rather than having every answer (Parkinson, 2010). In this course the Game Master sets scenes, voices non-player characters, adjudicates choices against the historical record, adjusts challenge, protects the respectful atmosphere of the table, and leads the debrief. The Game Master does not play against the students.

***Principle 8: Design for Every Player***

Every part of the game has an accessible path, and nothing essential depends on a single sense, a fast reaction, or one piece of technology. Students can choose the standard path instead of Story Play at any time without penalty, because not every student benefits from game-based delivery (Admiraal et al., 2011). The course follows the UDL Guidelines 3.0 (CAST, 2024), and all digital materials and platforms are checked against WCAG 2.1 Level AA before adoption, the standard public colleges must meet under the ADA Title II rule by April 26, 2027, for larger entities (Extension of Compliance Dates, 2026). The design also limits extra elements to reduce cognitive load. Plass et al. (2015) describe the tension between cognitive load and visual appeal, and Clark et al. (2016) found that simpler, schematic games did at least as well as realistic ones. When a feature is interesting but not necessary, it comes out.

***Principle 9: Assess Choices and the Reasoning Behind Them***

Assessment focuses on how students use evidence to make and explain choices, not only on what they can recall. Schwartz and Arena (2009) argue that choice, rather than accumulated knowledge, should be the organizing idea of assessment, since preparing students to choose well is a core goal of education. For a history course, this means grading the reasoning in the Write History reports: what the team chose, what actually happened, what evidence explains the difference, and what would have had to change for history to turn out differently. This approach

also answers the historical accuracy question raised in 2018, which Chapter 4 takes up in detail. Recall still matters and is still assessed with quizzes, but it is not the only thing that counts.

### The Framework at a Glance

Table 2 pulls the nine principles together and shows how each one changes the 2018 design.

**Table 2**

*Design Principles, Evidence, and Changes From the 2018 Edition*

| Principle | Core evidence | Change from 2018 |
|---|---|---|
| 1. Brief before play | Barzilai & Blau (2014); Clark et al. (2016); Ke (2009) | Adds a scaffold before every session, not only a course tutorial |
| 2. Make the content the mechanic | Brown et al. (1989); Gee (2005); Perrotta et al. (2013) | Every mission action requires historical thinking; story alone is not treated as enough |
| 3. Tune challenge, do not chase flow | Admiraal et al. (2011); Chen (2006); Hamari et al. (2016) | Flow is no longer claimed to cause learning; mission tiers give active difficulty choices |
| 4. Fail gracefully, with feedback | Erhel & Jamet (2013); Gee (2005); Plass et al. (2015) | Time resets kept; every failure now comes with targeted feedback and a side quest |
| 5. Debrief every session | Kim et al. (2009); Kohlberg & Hersh (1977) | Story recap becomes a required, structured debrief with a protocol |
| 6. Build a table, not a leaderboard | Clark et al. (2016); Ryan et al. (2006); Sailer & Homner (2020) | Luck points for LMS activity removed; team and in-story recognition replace them |
| 7. The instructor is the Game Master | Brown et al. (1989); Perrotta et al. (2013) | GM role defined, answering the 2018 reader comment |
| 8. Design for every player | CAST (2024); Extension of Compliance Dates (2026); Plass et al. (2015) | New accessibility requirements; opt-out at any time; fewer extra elements |
| 9. Assess choices and reasoning | Schwartz & Arena (2009); SUNY GE (n.d.) | Reports graded on evidence-based reasoning about divergence from history |

### What the Framework Does Not Promise

The framework is a set of design commitments backed by evidence, not a guarantee. Most of the studies behind it were conducted with digital games, often with younger students, and often over short periods. Wright et al. (2017) studied twelve students. Barzilai and Blau

(2014) studied an online math game, not a history role-play. The principles are reasonable inferences from that research to a college course, and Chapter 5 describes how to test whether they hold up in this setting.

## Chapter 4: Applying the Framework to a Hybrid/Online American History Course

This chapter applies Scaffolded Story Play to the example course from the 2018 edition: a three-credit American history course at a SUNY community college, offered in a fully online section and a hybrid section. The narrative is also carried forward. Students are temporary agents of The Historical Company, an organization that sends Time Agents into the past to verify the record of major events. The first module covers the opening of the American Revolution, the fighting at Lexington and Concord in April 1775.

The chapter moves from the course outcomes, to the two ways students can take the course, to the story and how it protects historical accuracy, to the roles, the module cycle, difficulty, recognition, platforms, accessibility, and the work this asks of an instructor.

### Course Outcomes

The 2018 edition used objectives from a local course outline. Looking back, one of its student learning outcomes, "identify, analyze and evaluate arguments as they occur in their own or other's work and will develop well-reasoned arguments of their own," closely matches the Critical Thinking and Reasoning outcome in SUNY's general education framework. That course outline is no longer available to cite, so this edition aligns the course directly to the SUNY General Education (SUNY GE) framework. SUNY requires approved general education courses to meet its learning outcomes, while allowing campuses to adapt the local wording as long as the intent stays intact (SUNY, n.d.).

An American history course would be approved in the US History and Civic Engagement category. All SUNY undergraduate degree programs must also address the core competencies of Critical Thinking and Reasoning, Information Literacy, and, beginning Fall 2026, Civic Discourse. The Information Literacy outcome was revised for Fall 2026 to include the ethical use

of information from emerging technologies such as artificial intelligence (SUNY, n.d.). A history course is a natural place to address several of these. Table 3 shows how the first module aligns with them.

**Table 3**

*Alignment of Module 1 With SUNY General Education Learning Outcomes*

| SUNY GE outcome (SUNY, n.d.) | Where students practice it in Module 1 | How it is assessed |
|---|---|---|
| US History and Civic Engagement: demonstrate understanding of United States society and/or history, including the diversity of individuals and communities that make up the nation | Agent perspectives include colonists of different social classes, free and enslaved Black New Englanders, Loyalists, British regulars, and women on the home front | Rift Report section on whose perspectives shaped events; quiz |
| US History and Civic Engagement: understand the role of individual participation in US communities and government | Missions to town meetings, militia musters, and committees of correspondence | Debrief prompt; individual reflection |
| US History and Civic Engagement: apply historical and contemporary evidence to draw, support, or verify conclusions | Every mission decision must cite a source from the briefing packet or the Company Archive | Rift Report rubric, Evidence criterion |
| Critical Thinking and Reasoning: clearly articulate an issue or problem; identify, analyze, and evaluate ideas, data, and arguments; acknowledge limitations such as perspective and bias; develop well-reasoned arguments | Comparing conflicting eyewitness accounts of who fired first at Lexington | Rift Report rubric, Reasoning and Perspective criteria |
| Information Literacy (revised Fall 2026): locate information effectively; evaluate sources for authority, validity, bias, and origin; understand the ethical dimensions of information use, including AI | Veteran-tier missions locate additional sources; all tiers evaluate source origin; Company policy on AI-generated "records" | Source evaluation log; AI use statement in the Rift Report |
| Civic Discourse (new, Fall 2026): deliberate through reasoned inquiry that considers multiple viewpoints; practice advocacy, dissent, and dialogue that attends constructively to conflict | Team deliberation before each major decision; structured debrief | Debrief participation rubric; peer feedback |

*Note.* Outcome wording is condensed from the SUNY GE framework. Campuses may adapt local wording while keeping the intent of each outcome (SUNY, n.d.).

### Two Paths, One Set of Outcomes

Students choose between two paths at Session Zero: Story Play or the Standard Path. The 2018 edition offered the same choice, citing Admiraal et al. (2011) to argue that a game should not be used for every student. That still holds, and self-determination theory adds a second reason. A choice about how to take a course supports autonomy, while forcing students into a game they do not want works against it (Ryan et al., 2006).

The two paths share everything that affects grades. Both use the same briefings, the same primary sources, the same quizzes, and the same rubric for the module's main written work. The Standard Path completes a traditional version of each assignment: a discussion, a source analysis, and an essay answering the same questions the Rift Report asks. Students can switch paths at the start of any module without penalty. Because the outcomes and rubrics match, the two paths also make it possible to compare them fairly, which Chapter 5 takes advantage of.

### The Story: The Historical Company

The Historical Company is the fiction that holds the course together. Its stated mission is to keep the most accurate record of major historical events, verified by multiple reports. Students apply for temporary assignment as Time Agents. They work in small cells, travel to key moments, gather evidence, and report back. The first two pieces of in-story correspondence from the 2018 edition are kept, revised, and included in Appendix C.

The story follows Perrotta et al.'s (2013) advice to place learning activities inside the fiction and rules of the game rather than bolting them on. An agent's job is historical research, so the tasks that move the story forward are the tasks the course is meant to teach: finding evidence, weighing accounts, placing events in context, and arguing from evidence.

## Protecting the Historical Record

In 2018 my second reader asked a question I did not fully answer in the thesis. If students can make different choices than the historical figures did, does that confuse the actual facts of history? And if they cannot, are they really role-playing? In my reply I proposed the time agent idea, where students go back in time to establish major events, and if events are altered, identify which point history depended on (C. T. Avis, personal communication, December 28, 2018). The design below builds that reply into the course.

### *The Record and the Rift*

The course separates two things and keeps them clearly labeled.

- **The Record** is what happened. It lives in the Company Archive, a set of briefing documents and primary sources in the LMS. The Game Master never presents invented events as part of the Record, and fictional characters created for the story are marked as fictional in the Agent Handbook.
- **The Rift** is what a cell's choices did inside the story. Agents can make choices the historical figures did not make. When they do, the story opens a Rift: a divergence between the timeline the cell is in and the Record.

Every module ends with a Rift Report. The report asks four questions. What did your cell choose, and why? What does the Record show actually happened? What evidence explains why events unfolded the way they did? What would have had to be different for events to go the way your cell's choices pushed them? The fourth question is where the learning shows. To answer it, a student has to understand the causes, constraints, and people behind an event well enough to explain why one path was taken and not another. That is historical reasoning, and it matches the SUNY GE outcome of applying evidence to draw, support, or verify conclusions (SUNY, n.d.).

It is also the kind of choice-centered assessment Schwartz and Arena (2009) recommend. The rubric is in Appendix E.

This answers both halves of the 2018 question. Students really are role-playing, because their choices matter in the story. The historical facts are protected, because the course never lets the story replace the Record, and because every module ends with students writing history as it happened and explaining why.

### *Time Resets as Graceful Failure*

The 2018 edition gave Time Agents three time resets. If a cell failed a major event because its report missed critical evidence, the cell could reset time and try again, but only three times before the reset device stopped working. The resets stay. They are the course's version of what Gee (2005) calls lowering the consequences of failure and what Plass et al. (2015) call graceful failure. What changes is what happens before a retry. Each failure produces specific feedback on what evidence was missing or misread, and the cell completes a short side quest aimed at that gap before resetting (Erhel & Jamet, 2013). A reset is not a do-over. It is a second attempt with better information.

## Roles

### *The Instructor as Game Master*

The instructor runs the game. Following Perrotta et al. (2013), the Game Master mediates the experience by giving guidance when needed, applying the rules consistently, and keeping the table respectful. In practice the Game Master:

- sets each scene and voices non-player characters, drawing only on the Record for historical figures;
- decides the outcomes of cell choices using the evidence cells present;

- watches for teams that are bored or overwhelmed and adjusts the next mission;
- protects the respectful atmosphere of the table, which matters more in a course that asks students to voice perspectives they may not share;
- leads the debrief and corrects misconceptions; and
- fades support over the term, moving from modeling how to reason about sources early on to coaching and then stepping back, as in cognitive apprenticeship (Brown et al., 1989).

The Game Master is not the players' opponent. My training in group work described the leader's job as creating a productive space rather than having every answer (Parkinson, 2010), and that is the stance this role needs.

***Cells and Rotating Roles***

Each cell has three to five agents. The 2018 edition used Primary Characters and Secondary Characters. This edition keeps that structure, gives the roles in-story names, and makes their interdependence explicit, following Gee's (2005) cross-functional teams principle.

- **Field Leads** (the 2018 Primary Characters) represent the cell during play, coordinate the cell's decisions, and are responsible for pulling the team's evidence into the Rift Report.
- **Archivists** (the 2018 Secondary Characters) research characters, places, and sources, and review the Field Lead's draft report for errors of fact or interpretation before it is filed.

Roles rotate each module, so every student takes both roles over the term. Each role holds information the other needs. Field Leads cannot make well-supported decisions without the

Archivists' research, and Archivists cannot see how their research plays out without the Field Leads. That interdependence supports relatedness (Ryan et al., 2006) and reflects the finding that collaborative team arrangements did better for learning than competitive individual ones (Clark et al., 2016).

### *Perspectives*

Each cell's agents are assigned perspectives within the event, such as a Lexington militia member, a Loyalist merchant, a British regular, a woman managing a farm while men muster, or a free Black member of the militia. Taking on perspectives supports the SUNY GE outcome on the diversity of individuals and communities, and it is also the role-taking Kohlberg and Hersh (1977) describe as the basis of moral development. The Game Master sets clear boundaries at Session Zero. Students explore perspectives through evidence and in-character reasoning, and they do not perform stereotypes. Any student can take a different perspective without explanation.

## The Module Cycle in Practice

Table 4 shows the first module, covering the events leading to the fighting at Lexington and Concord, on a two-week schedule. The hybrid section holds the synchronous session on campus. The online section holds it by video conference. An asynchronous play-by-post version runs in the LMS discussion area for students who cannot attend live.

**Table 4**

*Module 1 Schedule: "The Shot Heard Round the World"*

| Days | Phase | Story Play path | Standard Path |
|---|---|---|---|
| 1 to 2 | Brief | Company briefing on the 1774 crisis in Massachusetts; primary source packet; five-question readiness check with feedback | Same briefing, packet, and readiness check |

| Days | Phase | Story Play path | Standard Path |
|---|---|---|---|
| 3 to 5 | Play | 60 to 75 minute synchronous mission, or play-by-post mission over three days; cells decide how to act on the night of April 18 and morning of April 19, 1775, citing evidence for each decision | Guided source analysis of accounts of April 19, 1775 |
| 5 | Debrief | 20-minute structured debrief immediately after the live session; asynchronous debrief thread for play-by-post cells | Discussion post responding to the same debrief questions |
| 6 to 10 | Write History | Cell Rift Report plus a short individual reflection | Essay answering the four Rift Report questions |
| 11 to 12 | Level Up | Pulse survey; module quiz; cell chooses next mission tier; Game Master adjusts | Module quiz; choice of next assignment option |

*Note.* Both paths use the same readiness check, quiz, and written-work rubric.

### *Brief*

The briefing is short, framed in the story, and aimed at what the mission requires. For Module 1, it covers the tension in Massachusetts after the Coercive Acts, the role of committees and town militias, and the question the mission turns on: who fired first at Lexington, and why do accounts disagree? The readiness check gives immediate right-and-wrong feedback, following Erhel and Jamet (2013). This is the scaffold-before-play step that Barzilai and Blau (2014) found improved problem solving without reducing enjoyment.

### *Play*

During play, cells face decisions. Should the militia muster on Lexington Green or hold back? Whose account of the morning should the Company trust? Each decision requires the cell to cite evidence from the briefing packet. The Game Master responds to the choice and its evidence, not to how clever the idea sounds. A well-supported decision that differs from history opens a Rift. A poorly supported decision fails, which triggers feedback, a side quest, and possibly a time reset.

### *Debrief*

The debrief follows the protocol in Appendix D. It asks cells to think aloud about how they decided, what evidence mattered most, where their story departed from the Record, and what they would do differently. Kim et al. (2009) found that this kind of peer discussion of play had the strongest effect on social problem solving. The debrief is also where the Game Master names and corrects any historical misconceptions that came up in play, so they do not carry into the written report.

### *Write History*

The cell files a Rift Report, and each agent writes a short individual reflection. The report answers the four Rift questions and includes a source evaluation log. Starting Fall 2026, the report also includes a brief statement on whether and how the cell used AI tools, reflecting SUNY's revised Information Literacy outcome. In the story, the Company does not accept unverified records, and an AI-generated summary is treated as an unverified record until it is checked against sources.

### *Level Up*

At the end of the module the instructor reviews the quiz, the reports, and a short pulse survey. The survey asks a few questions drawn from the EGameFlow dimensions of challenge, goal clarity, and feedback (Fu et al., 2009). The instructor uses this information for passive adjustment of the next module, and cells choose their next mission tier for active adjustment (Chen, 2006).

## Difficulty: Mission Tiers

Each mission is offered at three tiers. Cells choose a tier at the Level Up step and can change tiers each module. The tiers change the amount of support, not the outcomes or the grade

ceiling. Every tier is assessed against the same rubric. This matters for equity: students who choose more support should not be capped at a lower grade for doing so.

**Table 5**

*Mission Tiers*

| Tier | Support provided | What agents do |
|---|---|---|
| Scout | Curated sources with guiding questions for each; decision points narrowed to two or three options | Use provided evidence to justify decisions; compare two accounts |
| Agent | Curated sources without guiding questions; open decision points | Select relevant evidence; reconcile conflicting accounts |
| Veteran | Starting sources only | Locate and evaluate at least two additional sources; argue from conflicting evidence |

*Note.* All tiers are graded with the same Rift Report rubric (Appendix E).

This follows Chen's (2006) recommendation to widen the range of challenge and let players choose their own level, and it follows Hamari et al.'s (2016) finding that challenge that keeps up with growing ability predicts learning. It also applies the UDL guideline to optimize challenge and support (CAST, 2024, 8.2).

**Recognition Without Leaderboards**

The 2018 edition awarded "luck points" as extra credit for interacting with LMS announcements and discussions, extra credit for rich character backstories, and avatar customization for strong performance. This edition removes points for routine LMS activity entirely. Rewarding students for opening announcements does nothing for autonomy, competence, or relatedness (Ryan et al., 2006), and it separates the reward from the fiction and rules of the game, which Perrotta et al. (2013) advise against.

Recognition in this edition works in three ways.

1. **Clearance levels.** Cells earn Company clearance by showing growth in four skills tied to the course outcomes: using evidence, placing events in context, weighing perspectives, and building arguments. Clearance unlocks access to new

areas of the Company Archive, which contain sources for later missions. The reward is in the story and it feeds back into learning.

2. **Team recognition.** Recognition goes to cells, not individuals, and is shared in the debrief. Friendly competition between cells is fine. Sailer and Homner (2020) found that competition combined with collaboration was particularly effective, and Clark et al. (2016) found that collaborative team competition outperformed individual competition.

3. **Agent dossiers.** Character backstories remain, but a dossier counts toward participation only when it is grounded in historical evidence about the kind of person the agent is portraying. The creative work stays, and it becomes part of the content rather than a separate reward.

**Platforms in 2026**

The 2018 edition recommended building the course on a learning management system, a virtual tabletop such as Roll20, and a gamification platform such as Classcraft. Since then I have led my college through a full LMS migration, from Blackboard to D2L Brightspace, and that experience taught me to design courses that can move from one system to another. Two other things have changed. First, the original Classcraft platform shut down on June 30, 2024, after its acquisition by Houghton Mifflin Harcourt, which relaunched the name as a K-8 product (Houghton Mifflin Harcourt, 2024). A course built on it would have lost its foundation. Second, accessibility requirements for public colleges are now specific and have deadlines. The platform plan below is built to survive both kinds of change.

- **The LMS is home base.** Every graded item, the Company Archive, briefings, the Agent Handbook, and the rubric live in the college's LMS. Story

documents are built as accessible pages with real headings, alternative text, and readable structure, not as image-based handouts. Branching correspondence like the 2018 "Yes, continue" links can be built with the LMS's release conditions.

- **Synchronous sessions** use the college's supported video conferencing tool with live captions, a text chat channel that is a full alternative to speaking, and recording only with students' informed consent.
- **A virtual tabletop is optional.** A map on a virtual tabletop can add a lot to a session, but nothing essential lives only on the map. Every map location and event is also described in text in the Archive. Any virtual tabletop has to pass the college's accessibility review before use.
- **The asynchronous path is play-by-post.** Tabletop players have long run games through written posts, and the LMS discussion area supports that directly. It also serves students who cannot attend live sessions and students who process better in writing.
- **No single vendor is required.** The design should be able to lose any one third-party tool without losing the course.

**Accessibility Plan**

Co-chairing my college's Accessibility Committee has taught me that accessibility is far easier to build in than to repair later. A game-based course raises the stakes because it adds maps, handouts, and live sessions that all have to work for every student. The course follows the UDL Guidelines 3.0 (CAST, 2024) and checks all digital content and platforms against WCAG 2.1 Level AA, the standard adopted under the ADA Title II web rule. Public entities with a total population of 50,000 or more, which includes many public colleges, must comply by April 26,

2027 (Extension of Compliance Dates, 2026). Table 6 shows how specific UDL guidelines appear in the course.

**Table 6**

*UDL Guidelines 3.0 in the Course Design*

| UDL guideline (CAST, 2024) | Course feature |
|---|---|
| 7.1 Optimize choice and autonomy | Choice of path, perspective, and mission tier; switch at any module |
| 7.3 Nurture joy and play | Story Play path; in-story recognition |
| 7.4 Address biases, threats, and distractions | Session Zero boundaries; no stereotyped portrayals; opt out of any perspective without explanation |
| 8.2 Optimize challenge and support | Scout, Agent, and Veteran tiers with the same grade ceiling |
| 8.3 Foster collaboration, interdependence, and collective learning | Field Lead and Archivist roles that depend on each other |
| 8.5 Offer action-oriented feedback | Readiness check feedback; targeted feedback before any time reset |
| 9.3 Promote individual and collective reflection | Structured debrief and individual reflection every module |
| 1.2 Support multiple ways to perceive information | Every map, image, and spoken scene also available as text; captions on all video |
| 4.1 Vary and honor the methods for response, navigation, and movement | Participate by voice, chat, or play-by-post; no timed reactions required |
| 4.2 Optimize access to accessible materials and assistive technologies | LMS-based accessible pages; platform accessibility review before adoption |
| 5.1 Use multiple media for communication | Rift Report may include an annotated timeline, audio narration, or a written report |

Appendix F provides a practical accessibility checklist for running a live session.

## What This Asks of the Instructor

This design is more work to build than a standard online course, and it would be unfair to pretend otherwise. The 2018 edition noted that development takes considerable time but maintenance becomes lighter once modules exist. That is true, but the research adds two cautions. Perrotta et al. (2013) found that teachers need time, preparation, technical support, and professional development to make game-based learning work, and that they should not be left to

do it alone. Barab et al. (2005) found that a game has to work for teachers and staff as well as students, because they decide whether it gets used.

My own job for the last several years has been helping faculty build and improve courses, and that experience agrees with the research. In practice, it means three things. A course like this should be built in partnership with instructional design staff, not by one faculty member working nights. The first offering should be treated as a pilot with a smaller number of modules in Story Play form. And the materials, including the Archive, correspondence, rubrics, and debrief protocol, should be built as reusable templates so other courses can adapt them. The 2018 edition suggested that an introductory case studies course might use simulated encounters, and that an algebra course might use escape-room puzzles. Those are still good ideas, and the framework in Chapter 3 is meant to travel to them, but each course has to work through the principles for its own content.

## Chapter 5: Evaluation Plan

The 2018 edition ended its methodology with a description of how to build the course. It did not say how anyone would know whether the course worked. My first thesis topic statement in 2017, written for an earlier version of the project, asked, "How do we evaluate student engagement from this particular project?" (Parkinson, 2017b). The final thesis never answered that question. This chapter does. It describes how the first offering of the course should be evaluated, what should be measured, how the data should be analyzed, and what the evaluation can and cannot show.

My work since 2018 shapes this plan as much as the research does. I designed structured success plans for students on negative academic standing and tracked how those students did over time, and as a member of my college's Middle States working committee on the design and delivery of the student learning experience, I have seen how much an institution needs evidence, not impressions, to decide whether something is working. The research in Chapter 2 shapes the plan in two ways. First, engagement and learning have to be measured separately, because they do not always move together (Admiraal et al., 2011; Barzilai & Blau, 2014). A course that students enjoy is not proof that it teaches. Second, the field has struggled with inconsistent methods and measures, which made even large reviews hard to combine (Connolly et al., 2012; Ke, 2009). A good evaluation should use comparisons, validated or clearly adapted instruments, and transparent analysis, so the results mean something to someone outside the classroom.

### Evaluation Questions

1. **Learning.** Do students on the Story Play path meet the module learning outcomes at rates equal to or higher than students on the Standard Path, as measured by the shared quiz and the shared written-work rubric?

2. **Engagement and motivation.** How do students on each path describe their engagement, challenge, goal clarity, feedback, and sense of autonomy, competence, and relatedness?

3. **Reasoning.** How do Story Play students reason about their choices and about the historical record in Rift Reports and debriefs, and does that reasoning change across modules?

4. **Access.** Can every student participate fully? What barriers appear, and for whom?

5. **Implementation.** What does the design require of the instructor, and which parts of the framework were delivered as intended?

**Design**

The evaluation uses a mixed-methods design that takes advantage of the two paths. Because the Story Play path and the Standard Path share outcomes, briefings, quizzes, and rubrics, students on both paths can be compared on the same measures. This is a quasi-experimental comparison, not a randomized trial. Students choose their path, and that choice is itself a variable, since students who choose Story Play may differ from those who do not. Randomly assigning students to a path would violate the choice at the center of the design (Principle 8), so the evaluation addresses the selection problem in other ways, described below.

The first offering should be treated as a pilot, followed by at least one more offering before strong conclusions are drawn. Clark et al. (2016) found that games played over multiple sessions produced better learning than single sessions, so the evaluation looks at learning across the full term rather than after a single module.

## Measures

### *Learning Outcomes*

Learning is measured with the assessments both paths already share. The module readiness checks and quizzes measure factual and conceptual knowledge. The written-work rubric (Appendix E) measures evidence use, reasoning, perspective, and argument. To reduce bias, a sample of Rift Reports and Standard Path essays should be scored by a second rater who does not know which path produced the work, with rubric scores compared for agreement. A short pretest at the start of the course on historical knowledge and source analysis gives a baseline for comparison.

### *Engagement and Experience*

Fu et al.'s (2009) EGameFlow scale measures eight dimensions of learner enjoyment in e-learning games: immersion, social interaction, challenge, goal clarity, feedback, concentration, control, and knowledge improvement. The course uses an adapted short form built on those dimensions. Appendix G gives sample items written for this course. Because the items are adapted, the evaluation should report their internal consistency rather than assuming the original scale's reliability carries over. The survey is given to Story Play students at mid-term and at the end of the term. A version with wording that fits a non-game course is given to Standard Path students, so dimensions such as goal clarity, feedback, and challenge can be compared.

A smaller set of items addresses the three needs from self-determination theory: autonomy, competence, and relatedness (Ryan et al., 2006). The end-of-module pulse survey uses three items from the challenge, goal clarity, and feedback dimensions. It is part of teaching, used for the Level Up step, and it doubles as a running engagement measure.

### *Reasoning*

Rift Reports, individual reflections, and debrief notes are the richest data the course produces. They show how students connect choices to evidence and how their reasoning develops. These are analyzed qualitatively. Saldaña (2009) describes structural coding, which labels segments of data according to the question they respond to, and notes that it suits data collected with a consistent set of questions. That fits the four Rift Report questions well. A first cycle of structural coding organizes the data by question, and a second cycle groups codes into patterns, such as how students treat conflicting accounts or how they explain causation. Rambaree (2007) describes using computer-assisted qualitative data analysis software to make coding more rigorous and transparent. Keeping the codebook, coding decisions, and memos in such a tool creates an audit trail that another reader can follow.

### *Access*

Accessibility is measured, not assumed. Data include accommodation requests, student reports of barriers through an anonymous form available all term, the results of accessibility checks on course materials and platforms against WCAG 2.1 Level AA, and participation patterns across the live, chat, and play-by-post options. The goal is to find out who could not fully participate and why, so the design can be fixed.

### *Implementation*

The instructor keeps a brief log after each session recording time spent, adjustments made, problems, and whether each phase of the module cycle happened as planned. A fidelity checklist based on the nine principles records, for each module, whether the briefing, debrief, feedback before resets, and tier choice actually occurred. Without this, a disappointing result could not be separated from a design that was never fully delivered.

**Analysis**

Quiz and rubric scores are compared across paths, adjusting for pretest scores and other available baseline measures to account for differences between the students who chose each path. Results are reported with effect sizes and confidence intervals, not just significance tests, so they can be compared with findings like those in Clark et al. (2016) and Sailer and Homner (2020). Given the small numbers in a pilot, these comparisons are descriptive and are used to plan the next offering, not to make broad claims.

Survey data are summarized by dimension and path, with reliability reported for the adapted items. Following Hamari et al. (2016) and Admiraal et al. (2011), the analysis looks at whether engagement, challenge, and learning scores are related, rather than assuming that more engagement means more learning.

Qualitative findings are analyzed with the coding approach above, and then integrated with the quantitative results. For example, if Story Play students score higher on the Reasoning criterion, the coded Rift Reports should show what that reasoning looks like. If they do not, the coded data may show where the game drew attention away from the content.

**Ethics and Privacy**

Course evaluation that will be shared beyond the classroom requires review by the college's institutional review board. Participation in the research is voluntary and separate from participation in the course. Students can take either path and complete all course work without agreeing to have their data used, and the instructor should not know who consented until grades are submitted. Session recordings are made only with informed consent, stored according to college policy, and never required for a grade. Student records are protected under federal privacy law and college policy, and data are de-identified before analysis.

If AI tools are used in teaching or in analysis, the concerns raised in the recent literature apply directly. Narreddy et al. (2025) and Marengo et al. (2025) both identify privacy and security as major concerns. Student work should not be entered into any AI tool that is not approved by the college for student data, and any AI-assisted analysis should be disclosed and checked by a human reader.

**Limitations of the Plan**

The evaluation cannot establish cause the way a randomized trial could, because students choose their path. Adjusting for baseline differences helps but does not remove the problem. A single instructor teaching both paths may also affect results, in either direction. Novelty is another risk: students may be more engaged by a new format simply because it is new, and that effect may fade. Finally, a pilot at one college with modest enrollment will produce results that are useful for improving the course but not strong enough to generalize. These are reasons to evaluate carefully and repeat the evaluation, not reasons to skip it.

# Chapter 6: Discussion and Conclusion

## What Changed in My Thinking

The biggest change is in how I think about fun. My first game-based course module in 2014 leaned almost entirely on the idea that games are engaging, and in 2018 I argued that flow leads to long-term learning. I was not wrong that engagement matters. Hamari et al. (2016) found that engagement and challenge predict learning, and Clark et al. (2016) found that games can outperform non-game instruction. But I was treating engagement as the finish line. The research I had already collected said otherwise. Students in Frequency 1550 were in flow and did not learn more because of it (Admiraal et al., 2011). Students who felt they learned the most were not necessarily the ones who solved problems best (Barzilai & Blau, 2014). Learners without support learned the game (Ke, 2009). Engagement is the condition that makes learning possible in a game-based course. The briefing, the feedback, and the debrief are what make it happen.

The second change is about rewards. My campus game idea in 2017 had a leaderboard with a TV as the top prize, and the 2018 thesis gave extra credit for reading announcements. Both were built on the idea that points make people participate. The evidence on gamification is more modest than that. Effects are real but small, and they depend on design choices such as game fiction and pairing competition with collaboration (Sailer & Homner, 2020). Self-determination theory says people keep playing because a game gives them choice, growing skill, and connection (Ryan et al., 2006). I believed that about myself in 2014, when I wrote that intrinsic motivation brought me to graduate school. I did not build my own course around it until this edition.

The third change is about who the design is for. The 2018 edition never mentioned accessibility. That was a real gap in a design that relied on maps, voice sessions, and a

gamification platform. Co-chairing my college's Accessibility Committee and serving as its designee for accessibility compliance since then has made clear to me that access is not a feature to add at the end. It is part of whether a design works at all, and for a public college it is now a legal requirement with a date attached (Extension of Compliance Dates, 2026). Universal Design for Learning turned out to fit this design naturally. Several of its guidelines, such as optimizing choice, nurturing play, optimizing challenge, fostering belonging, and promoting reflection, read like a description of a good game (CAST, 2024).

The fourth change is about AI. In 2018 generative AI was not part of the conversation. Since then I have become my college's lead on AI integration and written its AI ethics and digital assessment policies. The questions that work raises, about academic integrity, about what students should learn to do with these tools, and about what has to stay human, are the same questions a Game Master faces if AI is brought to the table. Chapter 6's section on AI reflects those conversations.

The last change is about my second reader's question. In 2018 I had an answer to whether role-play would confuse the historical record, but I only half built it into the thesis. Separating the Record from the Rift, and making every module end with students explaining why history went the way it did, turns what looked like a weakness of role-play into its strongest assessment.

**Implications for Instructional Designers**

The framework in Chapter 3 is written for a history course, but most of it is not about history. For designers considering game-based learning in any hybrid or online course, the research reviewed here points to a few practical conclusions.

- Start with the thinking the course is supposed to teach, then find a game structure that requires it. Do not start with a game and look for content to put in it (Brown et al., 1989; Perrotta et al., 2013).
- Put a briefing before play and a debrief after it. These are the cheapest parts of the design and some of the best supported (Barzilai & Blau, 2014; Kim et al., 2009).
- Give feedback on failure before a retry (Erhel & Jamet, 2013).
- Prefer team structures and in-story recognition over individual points and leaderboards (Clark et al., 2016; Sailer & Homner, 2020).
- Offer a non-game path and let students switch (Admiraal et al., 2011).
- Check every platform for accessibility before adopting it, and make sure nothing depends on a single tool or vendor (CAST, 2024; Houghton Mifflin Harcourt, 2024).
- Plan to evaluate from the first offering (Fu et al., 2009; Van Eck, 2006).

**Limitations**

This work has limitations that should be stated plainly. It is design work. The framework and course are grounded in research, but they have not yet been tested with students, and Chapter 5 is a plan, not results. Most of the empirical studies behind the framework involved digital games rather than tabletop-style role-play, many involved younger students, and several involved short interventions or small samples. Wright et al. (2017) studied twelve students, Kim et al. (2009) studied ninth graders, and Barzilai and Blau (2014) studied school-age children. The principles are reasonable applications of that research to college courses, but they are applications, and some may not transfer.

The literature review is also limited by its origin. It was built primarily from the research I collected during my graduate programs, supplemented for this edition with a small number of widely cited open-access sources. It is not a systematic review. Some important studies, particularly newer work on tabletop role-play in higher education, are likely missing.

Finally, the design depends heavily on the instructor. A skilled Game Master can adjust difficulty, protect the respectful atmosphere of the table, and turn a debrief into real learning. A less experienced or overloaded instructor may struggle to do all of that, which is why Chapter 4 recommends building courses like this with instructional design support and treating the first offering as a pilot (Perrotta et al., 2013).

## Future Directions

### *From MyQuest to HeroMaker*

Before this thesis settled on course design, my project was a game. MyQuest was a choice-based text adventure, built in TiddlyWiki, meant to teach new players how to create a Dungeons and Dragons character by making story choices that fill in a real character sheet (Parkinson, 2018b). By December 2018 the design document had grown into HeroMaker, in which players choose one of four adventurers, learn the rules through a narrative quest, and export their choices as a backstory they can take to a real table (Parkinson, 2018a). Appendix H describes that history. HeroMaker is a direct application of the principles in this thesis to a different learning goal: it teaches a system of rules through situated play with a built-in tutorial, the approach Gee (2008) describes. Finishing and evaluating it is a natural next project.

### *Generative AI as the Game Master's Assistant*

The most significant change in the tools available since 2018 is generative AI. A course like this one asks a lot of an instructor, and some of that work, such as drafting non-player

character dialogue, writing variations of a mission for each tier, or giving first-pass feedback on a draft report, is the kind of work AI tools are now being used for. The research is promising but cautious. Large language models can support personalized feedback and instructor planning, but reviews raise concerns about accuracy, bias, cost, privacy, and the need for teacher training and evaluation (Narreddy et al., 2025; Pang & Wei, 2025). Reviews of AI combined with gamification find potential for personalized learning and assessing collaboration, with limited long-term evidence and real concerns about data privacy and equity (Marengo et al., 2025). Game designers report that AI speeds up ideation but can create dependence and weaken originality (Alharthi, 2025).

If AI is used in this course, the framework sets the limits. The Game Master stays in charge. AI can help draft materials, but nothing enters the Record without being checked against sources, because an unverified AI summary is exactly the kind of record the Historical Company does not accept. Student work does not go into tools the college has not approved for student data. And students are taught to treat AI the same way the course treats any source, which is also what SUNY's revised Information Literacy outcome now asks of them (SUNY, n.d.). Sadik's (2025) course design, which positioned AI as an assistant students learned to use responsibly rather than a shortcut, is a useful model. Studying whether AI support makes a design like this more sustainable for instructors, without weakening the human judgment the Game Master role depends on, is a worthwhile direction for future research.

### *GamED*

Since 2024 I have been building GamED, a game-based learning platform grounded in the same foundations this edition draws on, including flow and self-determination theory. This edition gives that work a sharper set of design tests. The framework in Chapter 3 suggests that

the platform should put scaffolding before play, give feedback before retries, favor team structures over individual leaderboards, and be evaluated for learning rather than engagement alone. Applying the nine principles to GamED, and studying whether they hold outside a single course, is the most direct way to continue this work.

### *Other Courses*

The 2018 edition suggested that an introductory case studies course could use simulated encounters and that an algebra course could use escape-room puzzles. My own background suggests another: a human services ethics course built around role-play, where students take the roles of clients, families, and professionals in a dilemma, followed by a structured debrief. That is close to what I experienced as an undergraduate, and the moral development research in Chapter 2 suggests why it worked (Kohlberg & Hersh, 1977; Wright et al., 2017). Each of these would need to work through the nine principles for its own content and outcomes.

## Conclusion

Educators working in hybrid and online courses have a tremendous opportunity to engage students through game-based learning and game mechanics. I believed that in 2018, and I still do. What this edition adds, with eight more years of practice behind it, is a clearer sense of what turns that opportunity into learning. Games give students goals, roles, consequences, and a reason to work together. Research on learning explains why those things matter: people learn by doing meaningful work in real situations, with other people. Research on games adds a caution: engagement opens the door, but briefing, feedback, and reflection are what carry students through it.

Scaffolded Story Play is my attempt to build a course around both of those ideas. It keeps the Time Agents, the cells, the time resets, and the two paths from the 2018 edition. It adds a

briefing before play, a debrief after it, feedback before every retry, challenge students can choose, recognition shared by teams, an honest line between the story and the historical record, a design that works for every student, and a plan to find out whether it works. As research has stated and this thesis has tried to show, it is the design within the medium that determines how well students learn (Clark et al., 2016). The game is the medium. The design is the job.

## References

Adachi, P. J. C., & Willoughby, T. (2013). More than just fun and games: The longitudinal relationships between strategic video games, self-reported problem solving skills, and academic grades. *Journal of Youth and Adolescence, 42*(7), 1041–1052. https://doi.org/10.1007/s10964-013-9913-9

Adams, A. S. (2013). Needs met through role-playing games: A fantasy theme analysis of Dungeons & Dragons. *Kaleidoscope: A Graduate Journal of Qualitative Communication Research, 12*, Article 6. http://opensiuc.lib.siu.edu/kaleidoscope

Admiraal, W., Huizenga, J., Akkerman, S., & ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185–1194. https://doi.org/10.1016/j.chb.2010.12.013

Alharthi, S. A. (2025). Generative AI in game design: Enhancing creativity or constraining innovation? *Journal of Intelligence, 13*(6), Article 60. https://doi.org/10.3390/jintelligence13060060

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development, 53*(1), 86–107.

Barr, M. (2017). Video games can develop graduate skills in higher education students: A randomised trial. *Computers & Education, 113*, 86–97. https://doi.org/10.1016/j.compedu.2017.05.016

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education, 70*, 65–79.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42. https://doi.org/10.3102/0013189X018001032

CAST. (2024). *Universal design for learning guidelines version 3.0* [Graphic organizer]. https://udlguidelines.cast.org

Chen, J. (2006). *Flow in games* [Master's thesis, University of Southern California].

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/10.3102/0034654315582065

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Csikszentmihalyi, M. (2014). Toward a psychology of optimal experience. In *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi* (pp. 209–226). Springer.

D'Aprile, G., Di Bitonto, P., De Asmundis, R., & Ulloa Severino, A. (2015). Social, constructivist and informal learning processes: Together on the edge for designing digital game-based learning environments. *Journal of e-Learning and Knowledge Society, 11*(3), 23–39.

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education, 67*, 156–167.

Extension of Compliance Dates for Nondiscrimination on the Basis of Disability; Accessibility of Web Information and Services of State and Local Government Entities, 91 Fed. Reg. 20902 (April 20, 2026) (to be codified at 28 C.F.R. pt. 35).

Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education, 52*(1), 101–112. https://doi.org/10.1016/j.compedu.2008.07.004

Gee, J. P. (2005). Good video games and good learning. *Phi Kappa Phi Forum, 85*(2), 33–37.

Gee, J. P. (2008). Learning and games. In K. Salen (Ed.), *The ecology of games: Connecting youth, games, and learning* (pp. 21–40). MIT Press.

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*, 170–179. https://doi.org/10.1016/j.chb.2015.07.045

Houghton Mifflin Harcourt. (2024, February 21). *HMH launches immersive and dynamic learning experience with HMH Classcraft* [Press release]. https://www.hmhco.com/about-us/press-releases/hmh-launches-immersive-and-dynamic-learning-experience-with-hmh-classcraft

Hunicke, R., LeBlanc, M., & Zubek, R. (2004). MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI* (Vol. 4, No. 1, p. 1722).

Ke, F. (2009). A qualitative meta-analysis of computer games as learning tools. In R. E. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (Vol. 1, pp. 1–32). Information Science Reference.

Kim, B., Park, H., & Baek, Y. (2009). Not just fun, but serious strategies: Using meta-cognitive strategies in game-based learning. *Computers & Education, 52*(4), 800–810. https://doi.org/10.1016/j.compedu.2008.12.004

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory Into Practice, 16*(2), 53–59.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, A. R., Evans, C., & Vitak, J. (2008). *Teens, video games, and civics*. Pew Internet & American Life Project. http://www.pewinternet.org/files/old-media//Files/Reports/2008/PIP_Teens_Games_and_Civics_Report_FINAL.pdf.pdf

Lidwell, W., Holden, K., & Butler, J. (2010). *Universal principles of design* (Rev. ed.). Rockport.

MacDonald, C. J., & Thompson, T. L. (2005). Structure, content, delivery, service, and outcomes: Quality e-learning in higher education. *International Review of Research in Open and Distance Learning, 6*(2).

Marengo, A., Pagano, A., Lund, B., & Santamato, V. (2025). Research AI: Integrating AI and gamification in higher education for e-learning optimization and soft skills assessment through a cross-study synthesis. *Frontiers in Computer Science, 7*, Article 1587040. https://doi.org/10.3389/fcomp.2025.1587040

Narreddy, C., Joordens, S., & Prompiengchai, S. (2025). Harnessing large language models for scalable and effective formative assessment in higher education: A review. *Trends in Higher Education, 4*(4), Article 65. https://doi.org/10.3390/higheredu4040065

Oliver, R., & Herrington, J. (2001). *Teaching and learning online: A beginner's guide to e-learning and e-teaching in higher education*. Centre for Research in Information Technology and Communications, Edith Cowan University.

Olson, C. K. (2010). Children's motivations for video game play in the context of normal development. *Review of General Psychology, 14*(2), 180–187. https://doi.org/10.1037/a0018984

Pang, W., & Wei, Z. (2025). Shaping the future of higher education: A technology usage study on generative AI innovations. *Information, 16*(2), Article 95. https://doi.org/10.3390/info16020095

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*(1), 1–12. https://doi.org/10.1016/j.compedu.2008.06.004

Parkinson, J. (2010). *Transition phase group work* [Unpublished course paper]. Group Work Practices, SUNY Empire State College.

Parkinson, J. (2011). *Human service ethics role play summary* [Unpublished course paper]. Human Service Ethics, SUNY Empire State College.

Parkinson, J. (2014a). *Emerging technology and the basic needs of today* [Unpublished course paper]. Learning in Emerging Technology: Theory and Practice, SUNY Empire State College.

Parkinson, J. (2014b). *MALET journal blog* [Unpublished journal]. SUNY Empire State College.

Parkinson, J. (2014c). *Research and share 2: Situated cognition, immersive environments, and backward design* [Unpublished course paper]. Learning in Emerging Technology: Theory and Practice, SUNY Empire State College.

Parkinson, J. (2014d). *Universal instructional design and the flexibility in use, participation and presentation principle* [Unpublished blog post]. Designing Online Learning Environments, SUNY Empire State College.

Parkinson, J. (2017a). *Literature review* [Unpublished course paper]. Research Methods, SUNY Polytechnic Institute.

Parkinson, J. (2017b). *Preliminary topic statement: Creating a digital campus environment using video game mechanics, dynamics and aesthetics* [Unpublished manuscript]. IDT 599, SUNY Polytechnic Institute.

Parkinson, J. (2017c). *Thesis game* [Unpublished planning document]. SUNY Polytechnic Institute.

Parkinson, J. (2018a). *Game design document: HeroMaker* [Unpublished design document]. SUNY Polytechnic Institute.

Parkinson, J. (2018b). *MyQuest* [Unpublished project proposal]. SUNY Polytechnic Institute.

Perrotta, C., Featherstone, G., Aston, H., & Houghton, E. (2013). *Game-based learning: Latest evidence and future directions* (NFER Research Programme: Innovation in Education). National Foundation for Educational Research.

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist, 50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Prensky, M. (2001). Fun, play and games: What makes games engaging. In *Digital game-based learning* (Chapter 5). McGraw-Hill.

Rambaree, K. (2007). Bringing rigour in qualitative social research: The use of a CAQDAS. *University of Mauritius Research Journal, 13A*.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*(4), 347–363. https://doi.org/10.1007/s11031-006-9051-8

Sadik, O. (2025). Using generative AI to design and develop educational games in engineering education. *International Journal of Designs for Learning, 16*(2), 182–193. https://doi.org/10.14434/ijdl.v16i2.42059

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*, 77–112. https://doi.org/10.1007/s10648-019-09498-w

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage.

Schwartz, D. L., & Arena, D. (2009). *Choice-based assessments for the digital age* [White paper]. Stanford University School of Education.

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2004). *Video games and the future of learning* [Working paper]. University of Wisconsin–Madison and Academic Advanced Distributed Learning Co-Laboratory.

State University of New York. (n.d.). *SUNY general education framework (SUNY GE)*. Retrieved September 13, 2026, from https://system.suny.edu/academic-affairs/academic-policies/general-education/suny-ge/

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review, 41*(2), 16–30.

Wright, J. C., Weissglass, D. E., & Casey, V. (2017). Imaginative role-playing as a medium for moral development: Dungeons & Dragons provides moral training. *Journal of Humanistic Psychology*. Advance online publication. https://doi.org/10.1177/0022167816686263

Yaman, M., Nerdel, C., & Bayrhuber, H. (2008). The effects of instructional support and learner interests when learning using computer simulations. *Computers & Education, 51*(4), 1784–1794. https://doi.org/10.1016/j.compedu.2008.05.009

## Appendix A: Changes From the 2018 Edition

This appendix lists the substantive changes between the thesis submitted in December 2018 and this continuing edition. Minor edits for grammar, formatting, and APA style are not listed individually.

**Table 7**

*Change Log: 2018 Edition to 2026 Revised Edition*

| 2018 edition | Revised edition | Reason |
|---|---|---|
| No learning-theory section; "situated meaning" used without explanation | Chapter 2 opens with constructivist, sociocultural, and situated cognition foundations | The design needed a stated theory of learning (Brown et al., 1989; D'Aprile et al., 2015; Plass et al., 2015) |
| Flow presented as producing long-term learning, quoting Admiraal et al. (2011) | Flow treated as necessary but not sufficient; Admiraal's actual finding reported | Team flow was related to game performance, not learning outcomes (Admiraal et al., 2011); see also Barzilai & Blau (2014) |
| Gee and Lidwell principles listed side by side | Nine design principles and a Brief, Play, Debrief cycle, each tied to evidence (Chapter 3) | A framework is easier to apply and test than a list |
| Tutorial at course start only | Briefing before every session | Scaffold-before-play evidence (Barzilai & Blau, 2014; Clark et al., 2016) |
| "Story recap" feedback session | Required structured debrief with protocol (Appendix D) | Peer think-aloud evidence (Kim et al., 2009) |
| Time resets without guidance | Time resets kept; targeted feedback and side quest required before each reset | Feedback supports deeper processing (Erhel & Jamet, 2013) |
| Historical accuracy question unanswered in the text | The Record and the Rift; Rift Report assesses reasoning about divergence | Reader comment (C. T. Avis, personal communication, December 28, 2018); choice-based assessment (Schwartz & Arena, 2009) |
| Primary and Secondary Characters | Field Leads and Archivists with explicit interdependence and rotation | Cross-functional teams (Gee, 2005); collaborative structures (Clark et al., 2016) |
| Game Master role not defined | Game Master role defined | Reader comment (C. T. Avis, personal communication, December 28, 2018); Perrotta et al. (2013) |
| "Luck points" for LMS activity; extra credit for backstories; avatar rewards | Removed; clearance levels, team recognition, and evidence-based agent dossiers | Gamification evidence (Sailer & Homner, 2020); self-determination theory (Ryan et al., 2006) |
| Difficulty adjustment described as a grading system | Mission tiers (active adjustment) and instructor adjustment (passive), with equal grade ceiling | Chen (2006); Hamari et al. (2016) |

| 2018 edition | Revised edition | Reason |
|---|---|---|
| Objectives from a 2017 course outline no longer available | Alignment to SUNY GE outcomes, including Fall 2026 Information Literacy and Civic Discourse | SUNY (n.d.) |
| Classcraft and Roll20 recommended | LMS as home base; optional accessible virtual tabletop; play-by-post path; no required vendor | Original Classcraft platform shut down in 2024 (Houghton Mifflin Harcourt, 2024) |
| No accessibility content | UDL Guidelines 3.0 plan; WCAG 2.1 AA review; ADA Title II compliance date | CAST (2024); Extension of Compliance Dates (2026) |
| No evaluation plan | Mixed-methods evaluation plan (Chapter 5) | Engagement and learning must be measured separately |
| Section on "Sense of Community" drafted but cut; its references left in the reference list | Community restored in Chapter 2 with sources actually cited; orphaned references removed or used | Reference list accuracy |
| Hamari et al., Papastergiou, and the course outline cited in text but missing from references | All in-text citations matched to the reference list | APA accuracy |
| Trybus (2010), Moreno et al. (2007), Prensky (2003), Gee and Hayes (2012), Crawford et al. (2014) cited | Replaced with sources held in the research archive that support the same points | Every claim in the revision is traceable to a source on file |
| Claim that Wright et al. (2017) found role-play fosters a sense of community (in December 2018 drafts) | Removed; only Wright et al.'s measured findings on moral reasoning are reported | The claim does not appear in the source |

## Appendix B: 2018 Reader Comments and Where They Are Addressed

My second reader, C. T. Avis, left comments on the December 2018 draft. Most concerned APA formatting and wording, and most were fixed before submission. The substantive comments below were only partly addressed in 2018. Each is quoted from the draft comments (C. T. Avis, personal communication, December 28, 2018).

**Table 8**

*Substantive Reader Comments and Responses*

| Comment | Where addressed in this edition |
|---|---|
| "This is fairly specific to D&D players knowledge. Does the general reader need some explanation of these terms?" | Definitions section in Chapter 1; Game Master and Player Character defined; D&D-specific terms reduced throughout |
| "Spell out or refer to them as game masters. Define their role." | Chapter 3, Principle 7; Chapter 4, The Instructor as Game Master |
| "This is an interesting idea, but how does it work in education? Consider a follow-up." | Dynamic difficulty adjustment is now explained through mission tiers and instructor adjustment (Chapter 3, Principle 3; Chapter 4, Mission Tiers) |
| A request for an example of dynamic difficulty in practice | Worked examples added throughout Chapter 4, including the Module 1 schedule and sample mission decisions |
| "Do PCs have a choice to do things differently that [sic] the historical figures in the example history course? If so, could this lead to confusion as the actual facts of history? If not, are they really role playing?" | Chapter 4, Protecting the Historical Record: the Record and the Rift, and the Rift Report |
| "This is fairly formal document. Do you want the informal 'you' in it?" | This edition uses first person where I describe my own experience and reasoning, and a formal voice elsewhere; "you" is limited to in-story correspondence |

## Appendix C: Sample Module Materials

### Correspondence 1: Application Accepted

*The following is revised from the 2018 edition. It is delivered as an accessible LMS page at the start of Module 1.*

Good day, Agent Candidate.

Your application for temporary assignment with The Historical Company has been accepted. We are eager to begin the next step of the hiring process.

The Historical Company is dedicated to preserving the record of major historical events. Every event in our Archive has been verified by multiple reports. Our agents do not guess, and they do not make things up. They find evidence.

As part of your application, you were asked to read the Company briefing on the crisis in Massachusetts in 1774 and the primary source packet. To prove you are ready for field work, file a short report on the events of April 19, 1775, famously remembered as "the shot heard round the world." Include key people, places, organizations, and governments, and cite the documents from your packet that support each point.

A few rules for your first report:

1. There is no length requirement. There is an evidence requirement.

2. Some information must be obtained before you can continue.

3. Working with other candidates is encouraged.

4. Hints may be broad. Read them carefully anyway.

If your report does not include the evidence our Archivists need, you will receive a note explaining what is missing and a short field assignment to find it before you try again.

**Correspondence 2: Assignment to a Cell**

Excellent work. Your report has been accepted, and your clearance for time travel is approved.

Your next assignment will be with other agents. You will be placed in a cell. In each mission, some agents serve as Field Leads and some serve as Archivists, and those roles rotate. Field Leads represent the cell and bring its evidence together. Archivists research the people, places, and sources the cell needs and check the final report before it is filed. Neither role can succeed without the other.

Remember the first rule of the Company: the Record is what happened. In the field, your choices may send events in another direction. When that happens, your job is not to pretend history changed. Your job is to explain, with evidence, why it went the way it did.

Are you ready for field assignment?

[Yes, continue to your cell assignment] [No, take the Standard Path for this module]

*Selecting "No" opens the Standard Path version of Module 1 without penalty. Students can choose again at the start of each module.*

**Sample Mission Decision: Lexington Green**

*Agent-tier version. Scout-tier versions add guiding questions to each source and narrow the options; Veteran-tier versions provide only the starting sources and require agents to locate two more.*

It is before dawn on April 19, 1775. Word has come that British regulars are marching toward Lexington. Your cell is with the town's militia on the Green. Two accounts in your packet describe what happened next, and they disagree about who fired first.

Decide: What do you advise the militia to do, and which account do you trust more? Cite at least one source for each part of your decision. Your Game Master will respond based on your evidence.

## Appendix D: Debrief Protocol

The debrief follows every session. It takes about 20 minutes live, or runs as a structured discussion thread for play-by-post cells. The Game Master facilitates and does not grade the debrief on correctness; the goal is honest reasoning. The prompts draw on the think-aloud and modeling strategies studied by Kim et al. (2009) and on the role of moral discussion described by Kohlberg and Hersh (1977).

1. **What happened?** Each cell summarizes its key decision in two or three sentences.
2. **How did you decide?** Cells think aloud about the evidence that mattered most and any disagreements within the cell.
3. **Where did the story leave the Record?** Cells identify any Rift and describe what actually happened.
4. **Why did history go the way it did?** Cells name the causes, constraints, or people that made the actual outcome more likely.
5. **Whose perspective was missing?** Cells consider who in the event their decision affected but did not account for.
6. **Game Master's corrections.** The Game Master names any historical misconceptions that came up and points to the source that corrects them.
7. **What would you do differently?** Each agent writes one sentence to carry into the Rift Report.

## Appendix E: Rift Report Rubric

The same rubric is used for the Story Play Rift Report and the Standard Path essay, and for all mission tiers.

**Table 9**

*Rift Report and Standard Path Essay Rubric*

| Criterion | Proficient (meets outcome) | Developing | Beginning |
|---|---|---|---|
| Evidence | Cites relevant sources accurately for each claim; evaluates source origin and reliability | Cites sources for most claims; limited evaluation of reliability | Few or inaccurate citations; sources not evaluated |
| The Record | Accurately describes what happened, including key people, places, and institutions | Mostly accurate with minor errors or gaps | Significant factual errors or confuses the story with the Record |
| Reasoning | Explains why events unfolded as they did and what would have had to differ, using evidence | Offers causes but connects them to evidence inconsistently | Describes events without explaining causes |
| Perspective | Identifies multiple perspectives in the event and acknowledges bias and limitations in accounts | Identifies more than one perspective with limited analysis | Presents a single perspective |
| Argument | Clear, well-reasoned argument with a logical structure | Argument present but uneven | No clear argument |
| AI and source ethics (Fall 2026) | States whether and how AI tools were used; verifies any AI-assisted content against sources | Statement present but verification unclear | No statement, or unverified AI content presented as fact |

*Note.* Criteria are aligned with the SUNY GE US History and Civic Engagement, Critical Thinking and Reasoning, and Information Literacy outcomes (SUNY, n.d.). Proficient corresponds to the course's mastery threshold.

## Appendix F: Accessibility Checklist for a Live Session

This checklist is used before each synchronous session. It reflects the UDL Guidelines 3.0 (CAST, 2024) and the WCAG 2.1 Level AA standard adopted under the ADA Title II web rule (Extension of Compliance Dates, 2026). It is a practical checklist for instructors, not a substitute for the college's formal accessibility review.

- Live captions are turned on, and the recording (if students consented) will have edited captions.
- A text chat channel is open and the Game Master reads it as closely as spoken input.
- Every map, image, and handout used in the session also exists as described text in the Company Archive.
- No decision requires a timed reaction; cells have time to deliberate and can post decisions in writing.
- Visual information is never conveyed by color alone.
- Any virtual tabletop or third-party tool in use has passed the college's accessibility review, and nothing essential lives only in that tool.
- Documents shared during the session are accessible pages or tagged files with real headings and alternative text.
- Students who cannot attend live have a play-by-post version of the same mission with the same deadline structure.
- Session Zero boundaries on perspectives and portrayals are restated, and any student can change perspectives without explanation.
- Accommodation plans on file are reflected in the session setup.

## Appendix G: Sample Evaluation Items

The items below were written for this course. They are adapted to the eight dimensions of the EGameFlow scale described by Fu et al. (2009) and to the three needs in self-determination theory (Ryan et al., 2006). They are not the original validated items, so their reliability must be checked before results are reported. Students respond on a five-point scale from *strongly disagree* to *strongly agree*.

**Table 10**

*Sample Adapted Items by Dimension*

| Dimension | Sample item (Story Play version) |
|---|---|
| Concentration | During missions, I was able to stay focused on the task. |
| Goal clarity | I understood what each mission was supposed to help me learn. |
| Feedback | I got feedback that showed me what I did well and what I needed to fix. |
| Challenge | The missions were challenging without being overwhelming. |
| Control | I felt my choices affected what happened in the mission. |
| Immersion | I became absorbed in the story during missions. |
| Social interaction | Working with my cell helped me understand the history better. |
| Knowledge improvement | I can explain the events of this module better than before it started. |
| Autonomy | I had meaningful choices about how to take part in this course. |
| Competence | I felt I was getting better at using historical evidence. |
| Relatedness | I felt connected to other students in this course. |

*Note.* Standard Path versions replace mission and story wording with assignment wording. The end-of-module pulse survey uses the challenge, goal clarity, and feedback items.

## Appendix H: From MyQuest to HeroMaker

Before this thesis became a course design, it was a game. This appendix records how the project developed, because the prototype is the most direct test of the ideas in this thesis and a natural next step.

### Campus Minimap (Fall 2017)

My first topic statement proposed a game that would help new students explore college life, using video game mechanics, dynamics, and aesthetics as described by Hunicke et al. (2004). Students would explore campus offices, make appointments, join clubs, and earn points and badges for behaviors associated with staying in college (Parkinson, 2017b). A companion planning document proposed a leaderboard with prizes (Parkinson, 2017c). The idea was sound in its goal, but it relied heavily on points and prizes, which Chapter 2 of this edition now questions.

### MyQuest (Spring and Summer 2018)

By April 2018 the project had become MyQuest, a choice-based text adventure for players 14 and older that teaches the mechanics of Dungeons and Dragons through play. Players create a character by making story choices, and those choices fill in a real fifth-edition character sheet, with an introduction to dice checks, skill checks, and alignment along the way (Parkinson, 2018b). I built working prototypes in TiddlyWiki over the summer of 2018. Game mechanic information such as races, classes, and statistics came from freely available rules resources, used for educational purposes only.

### HeroMaker (December 2018)

The final design document renamed the project HeroMaker. Players choose one of four adventurers, each with a place in the party, and move through a narrative investigation with side

quests in a city. The interface is framed as a Hero's Journal. Each choice teaches a piece of the rules, and the choices are saved and exported as a backstory the player can take to a real table (Parkinson, 2018a).

**How HeroMaker Connects to This Thesis**

HeroMaker applies several of the principles in Chapter 3 to a different learning goal. It teaches a rule system through situated play with a tutorial that teaches just enough to begin (Gee, 2008). Its content is its mechanic, since choosing a character's abilities is how players learn what the abilities mean. It lets players choose from characters with different roles in a party, which is the cross-functional team principle (Gee, 2005). What it would need, based on this edition, is a clearer briefing and debrief structure, feedback on choices, an accessible text-first interface, and an evaluation plan. Completing and testing HeroMaker would test whether the framework in this thesis works outside a college course.