TU WIEN Informatics

Shwizard: Ein Lernspiel zu Shader-Konzepten

Senkung der Einstiegshürde in die Computergrafik durch spielbasiertes Lernen

BACHELORARBEIT

zur Erlangung des akademischen Grades

Bachelor of Science

im Rahmen des Studiums

Medieninformatik und Visual Computing

eingereicht von

Jonathan Seyr
Matrikelnummer e12119032

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Dipl.-Ing. Dipl.-Ing. Dr.techn. Michael Wimmer
Mitwirkung: Annalena Ulschmid, MSc BSc

Wien, 22. Juli 2026

Jonathan Seyr   Michael Wimmer

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

TU WIEN Informatics

Shwizard: An Educational Game for Teaching Shader Concepts

Lowering the Barrier to Computer Graphics Education via Game-Based Learning

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Media Informatics and Visual Computing

by

Jonathan Seyr
Registration Number e12119032

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Dipl.-Ing. Dipl.-Ing. Dr.techn. Michael Wimmer
Assistance: Annalena Ulschmid, MSc BSc

Vienna, July 22, 2026

Jonathan Seyr	Michael Wimmer

Erklärung zur Verfassung der Arbeit

Jonathan Seyr

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.
Ich erkläre weiters, dass ich mich generativer KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Im Anhang „Übersicht verwendeter Hilfsmittel" habe ich alle generativen KI-Tools gelistet, die verwendet wurden, und angegeben, wo und wie sie verwendet wurden. Für Textpassagen, die ohne substantielle Änderungen übernommen wurden, habe ich jeweils die von mir formulierten Eingaben (Prompts) und die verwendete IT-Anwendung mit ihrem Produktnamen und Versionsnummer/Datum angegeben.

Wien, 22. Juli 2026

Jonathan Seyr

Declaration of Authorship

Jonathan Seyr

I hereby declare that I have written this thesis independently, that I have completely specified the utilized sources and resources and that I have definitely marked all parts of the work – including tables, maps and figures – which belong to other works or to the internet, literally or extracted, by referencing the source as borrowed.
I further declare that I have used generative AI tools only as an aid, and that my own intellectual and creative efforts predominate in this work. In the appendix "Overview of Generative AI Tools Used" I have listed all generative AI tools that were used in the creation of this work, and indicated where in the work they were used. If whole passages of text were used without substantial changes, I have indicated the input (prompts) I formulated and the IT application used with its product name and version number/date.

Vienna, July 22, 2026

Jonathan Seyr

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Mein besonderer Dank gilt Annalena Ulschmid für die großartige Betreuung, die wertvollen Ideen und die kontinuierliche Unterstützung während der gesamten Arbeit.

Weiters bedanke ich mich herzlich bei allen Teilnehmerinnen und Teilnehmern der Studie, die sich die Zeit genommen haben, *Shwizard* zu spielen und die Fragebögen auszufüllen – ohne sie wäre diese Arbeit nicht möglich gewesen.

Schließlich danke ich meinem Freund Bozhidar Dejkoski, der mir während der gesamten Zeit den Rücken gestärkt und mir geholfen hat, durchzuhalten.

Acknowledgements

I would like to thank everyone who contributed to the completion of this thesis.

My special thanks go to Annalena Ulschmid for her excellent guidance, valuable ideas, and continuous support throughout this work.

I would also like to sincerely thank all study participants who took the time to play *Shwizard* and fill out the questionnaires – without them, this work would not have been possible.

Finally, I thank my boyfriend Bozhidar Dejkoski, who supported me throughout and helped me persevere.

Kurzfassung

Diese Arbeit untersucht, ob ein Serious Game grundlegendes Wissen über Shader-Programmierung und Computergrafik an Personen ohne Vorerfahrung auf diesem Gebiet vermitteln kann. Das für diese Studie entwickelte Spiel *Shwizard* ist ein Unity-WebGL-Open-World-Spiel, in dem Spielerinnen und Spieler Shader-Konzepte in einer spielinternen Enzyklopädie erlernen, Multiple-Choice-Quizfragen beantworten, um magische Zaubersprüche freizuschalten, und diese Zaubersprüche anwenden, um Quests abzuschließen. Die Lerninhalte umfassen zehn Shader-Themen, von allgemeiner GPU-Architektur bis hin zu spezifischen Effekten wie Wasser-Shadern, Beleuchtungsmodellen und Outline-Shadern.

Zur Evaluierung des Spiels wurde eine Studie mit Fragebögen vor und nach dem Spielen durchgeführt. Die Wissenswerte bei Fragen zu Shader-Konzepten stiegen nach dem Spielen signifikant an, mit einer mittleren Effektstärke. Die Teilnehmenden bewerteten das Spiel als ansprechend und die präsentierten Informationen als verständlich; eine Mehrheit beurteilte das Spielformat als effektiver als das Lernen aus einem Lehrbuch. Das wahrgenommene Lernen korrelierte substanziell mit dem gemessenen Wissenszuwachs. Die Ergebnisse stützen die Eignung von Serious Games als Medium für einführende Computergrafik-Ausbildung, zeigen jedoch auch Deckeneffekte in einer technisch erfahrenen Stichprobe sowie Verbesserungspotenziale im Spieldesign auf.

Abstract

This thesis investigates whether a serious game can effectively convey foundational knowledge about shader programming and computer graphics to players with no prior experience in the field. The game developed for this study, *Shwizard*, is a Unity WebGL open-world game in which players learn about shader concepts through an in-game encyclopedia, answer multiple-choice quiz questions to unlock magical spells, and apply those spells to complete quests. The educational content covers ten shader topics ranging from general GPU architecture to specific effects such as water shaders, lighting models, and outline shaders.

A pre- and post-game questionnaire study was conducted to evaluate the game. Knowledge scores on a set of shader concept questions increased significantly after playing, with a medium effect size. Participants rated the game as engaging and the presented information as clear, and a majority judged the game format to be more effective than learning from a textbook. Perceived learning correlated substantially with measured knowledge gain. The findings support the viability of serious games as a medium for introductory computer graphics education, while also highlighting ceiling effects in a technically experienced sample and areas for design improvement.

Contents

CHAPTER 1

Introduction

This thesis presents *Shwizard*, a serious game developed in Unity WebGL that teaches foundational shader concepts to players with no prior programming or computer graphics experience, and evaluates it in a pre–post user study. Playing the game led to a statistically significant increase in participants' knowledge of shader concepts with a medium effect size, participants rated the game as engaging and its content as clear, and a majority judged the game format to be more effective than learning the same material from a textbook.

The problem addressed by this work is how to convey shader knowledge to complete beginners. A *shader* is a small program that runs on the graphics card and most commonly determines how 3D content appears on screen—for example, the color of each pixel or the position of each point of a 3D model. The term *serious game* refers to a game designed primarily for a purpose other than pure entertainment, such as education, training, or awareness raising [1]. Combining the two, this thesis addresses the following research questions:

- **RQ1:** Does playing *Shwizard* lead to a measurable increase in players' factual knowledge of shader concepts?
- **RQ2:** Do players find the game-based learning format engaging and clear?
- **RQ3:** Do players perceive the game as more effective than a traditional learning method such as reading a textbook?

Shaders are central to virtually all real-time graphics, from games to visualization software, yet shader programming is rarely taught outside of university-level computer graphics courses. Dedicated learning resources exist—interactive guides such as The Book of Shaders [2] and community platforms such as Shadertoy [3]—but they are code-first:

they require the learner to read and write shader code from the outset and therefore presuppose programming experience. Serious games, in turn, have demonstrated positive learning effects across many domains [4, 5] and have been applied to university-level programming education [6] and to computer graphics coursework [7], but no prior serious game has targeted shader concepts for an audience without programming background. This leaves a gap: The group for whom shaders are most inaccessible is precisely the group existing materials serve worst. The main obstacles are that shader concepts are abstract and invisible—they describe computations, not objects—and that code, the usual medium for making them concrete, is itself the barrier. A solution that makes shader effects tangible without code would lower the entry barrier to a field that is both practically relevant and widely perceived as intimidating.

Shwizard overcomes these obstacles through three design decisions. First, shaders are reframed as magical spells: castable abilities that visibly transform objects in an open game world, making the invisible effects of shader programs directly observable without any code. Second, the concepts are explained in an in-game encyclopedia using natural language and diagrams instead of source code, written for readers without technical background. Third, learning is verified and rewarded within the game loop itself: a short quiz gates each spell, so that understanding a topic is what grants the ability to use it. To evaluate this design, a two-part questionnaire study was conducted in which participants were tested on their shader knowledge before and after playing, with game telemetry linking both measurements anonymously.

The evaluation with 64 participants (41 complete cases) showed a significant knowledge gain with improvement on every question tested, engagement and clarity ratings in the upper range of the scales, and a majority preference for the game over a textbook. Beyond the headline result, the analysis surfaced two less obvious findings: participants' self-perceived learning closely tracked their actually measured learning, and the perceived clarity of the learning content—rather than the knowledge gain itself—was the variable most strongly associated with the overall quality of the experience. Obtaining these results was not trivial: the sample turned out to be more technically experienced than the intended target audience, producing a strong ceiling effect that had to be identified and separated from the treatment effect, and the anonymous pre–post linking had to work reliably without collecting any identifying data.

The main contributions of this thesis are:

- The design and implementation of *Shwizard*, to the author's knowledge the first serious game targeting shader concepts for an audience without programming background
- A pre–post user study with 64 participants (41 complete cases) demonstrating a statistically significant knowledge gain (Cohen's $d = 0.55$) with improvement on every question tested

- evidence that self-reported learning tracks objectively measured knowledge gain (Spearman's $\rho = 0.58$), supporting the validity of self-report instruments in this setting
- The finding that content clarity is a central driver of the overall experience of an educational game, with a strong negative association to frustration ($\rho = -0.76$)
- A majority judgment (56 %) that the game format is more effective than a textbook for this material
- An explicit analysis of the threats to the validity of these results—most notably a ceiling effect in a technically experienced, self-selected sample and the absence of a control condition—which bounds the conclusions and defines the agenda for follow-up studies.

The remainder of this thesis is organized as follows. Chapter 2 situates this work in the broader literature on serious games and shader education; future work will focus on a controlled comparison against a non-game version of the same content and on evaluating the game with genuine beginners. Chapter 3 describes the methodology: the study design, the development and implementation of the game, and features that were considered but not implemented. Chapter 4 presents the results, including the statistical evaluation, a summary of participant responses, a discussion of the findings, and the threats to their validity. Chapter 5 concludes the thesis with a summary of the findings and directions for future work.

CHAPTER 2

Related Work

2.1 Serious Games and Their Effectiveness

The empirical evidence for the effectiveness of serious games, as defined in Chapter 1, is substantial. In a widely cited meta-analysis covering 77 learning comparisons, Wouters et al. [4] found that serious games were more effective than conventional instruction in terms of both learning (Cohen's $d = 0.29$) and retention ($d = 0.36$), though—contrary to common assumption—not significantly more motivating. Importantly for the present work, their moderator analyses showed that learning effects were strongest when games were supplemented with other instructional methods and played in multiple sessions.

A more recent decade-spanning review by Zhonggen [5] similarly concluded that serious game-based learning is significantly more effective than non-game-based learning and that learners remain engaged longer, while also cataloguing disadvantages such as high development cost and highly variable design quality. Backlund and Hendrix [8], surveying empirical studies across formalized school contexts from pre-school to higher education, found a generally positive picture but noted considerable methodological variation between studies. Taken together, these findings suggest that serious games can work, but that their success depends on the alignment between game mechanics and learning objectives rather than on the medium itself.

2.2 Serious Games in University and School Education

Serious games are used at both school and university level, although the emphasis differs. At primary and secondary school level, games are widely used to introduce computational thinking and basic programming concepts, typically through block-based programming and puzzle mechanics; systematic reviews report a large and growing body of such interventions [8]. At university level, serious games have been applied to harder technical subjects. A prominent example is *Prog&Play* by Muratet et al. [6], a real-time strategy game developed to strengthen programming skills of first-year computer science students, motivated by high dropout rates in introductory programming courses. Experimental deployment across several French universities showed positive effects on student motivation and practice [9].

Closest in subject matter to the present work is *Immersivio* by Moradi et al. [7], a problem-based serious game for learning 3D computer graphics, evaluated in a quasi-experimental study with undergraduate students in Malaysia (n = 50). The game is structured as a series of three levels, each presenting problem-solving tasks built around geometric concepts of 3D computer graphics, with the game design following the problem-based learning model of Hmelo-Silver—emphasizing higher-order thinking, collaborative learning, and playing for learning. The experimental group using the game achieved significantly higher post-test scores in a 3D computer graphics course than the control group. Immersivio targets enrolled computer science students and course-level content; to the author's knowledge, no prior serious game has targeted shader concepts specifically for an audience without any programming background, which is the research gap *Shwizard* aims to fill.

2.3 Similar Games and Learning Tools for Shaders

Dedicated learning resources for shader programming exist, but none in the form of a game. The *Book of Shaders* by Gonzalez Vivo and Lowe [2] is an interactive, step-by-step online guide to fragment shaders with live-editable code examples. Shadertoy [3] is a browser-based shader editor and community platform that has become the de facto center of the shader programming community, but it deliberately provides no instruction and assumes existing knowledge. Tools such as *Shader School* [10] provide workshop-style exercises for OpenGL Shading Language (GLSL). All of these resources are code-first: they require the learner to read and write shader code from the outset. *Shwizard* differs fundamentally by embedding shader concepts in game mechanics and natural-language explanations, removing code entirely, and targeting complete beginners.

CHAPTER 3

Methodology

3.1 Development of the Game

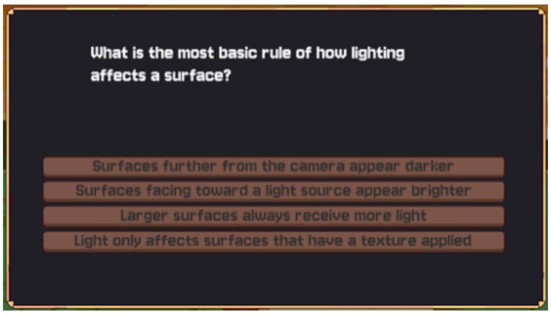
Development began with a game design document that specified the core concept, the spell metaphor, the list of shader topics, the quest structure, and the intended target audience. This document served as the reference point for all subsequent design and implementation decisions; the design described in the following reflects its final iteration.

The central design challenge of *Shwizard* was to make shader concepts feel meaningful and tangible to players without a programming background. Rather than presenting shaders as code, the game reframes them as magical spells: abilities that visibly transform the appearance of the world. This metaphor preserves the essential truth of what shaders do—they control how things look—while removing the intimidating association with programming. The wizard theme also provides a natural narrative justification for the learning loop: spells must be studied before they can be cast.

The core loop consists of three repeating stages. In the *study* stage, the player opens the in-game encyclopedia and reads a short text explaining a shader concept, accompanied by illustrative diagrams (Figure 3.1a). In the *quiz* stage, the player answers three multiple-choice questions (Figure 3.1b); all three must be answered correctly to proceed, with incorrect answers requiring a retry. In the *unlock* stage, the player receives a new spell corresponding to the concept and can cast it on objects in the world, visibly applying the shader effect. Unlocked spells are selected through a radial spell wheel that displays the name of the selected spell and the topic that unlocked it (Figure 3.3). The game is completed after solving four quests, each requiring specific spells.

The educational content covers ten topics in a progression from foundations to applications: shaders as programs on the GPU in general, the fragment shader, the vertex shader, outline shaders, general lighting, ambient lights, spot lights, point lights, noise functions, and water shaders. Each topic is presented as a text of roughly 800 to 1000 characters

(a) Encyclopedia entry for the water shader topic.



(b) Quiz interface with four answer options.

Figure 3.1: The two learning-focused interfaces of *Shwizard*: The in-game encyclopedia (a) and the multiple-choice quiz (b).

written for readers with no technical background, using real-world analogies and avoiding technical jargon, with one or two diagrams per topic. The three quiz questions per topic were designed to be answerable directly from the accompanying text, so that the quiz measures reading comprehension rather than prior familiarity.

3.2 Implementation

Shwizard was developed with the Unity engine [11] and deployed as a WebGL application on a custom domain, making it playable in a standard desktop web browser without installation. The open world (Figure 3.2) was built using Unity's terrain and lighting tools, with custom shader materials used both as scenery and as demonstrations of the concepts being taught. The visual assets of the game were partially sourced from third-party providers, including asset packs from Synty Studios and itch.io. The corresponding licenses and attributions are listed in the game's credits.

The encyclopedia is an in-game UI panel accessible at any time from the Heads-Up Display (HUD). The quiz state is tracked per topic and a spell is unlocked only when all three questions have been answered correctly. Before entering the open world, the player must complete a fixed set of introductory topics—*General*, *Fragment Shader*, *Vertex Shader*, and *Outline Shader*—which establish the foundational vocabulary and unlock the interaction mechanics the rest of the game relies on. The remaining topics can then be studied in any order while exploring. Spell casting combines Unity's particle system with material replacement, so that outline, lighting, and water effects are applied visibly to world objects in real time. Figure 3.6 shows an example of a transparency spell applied to a tree. The outline shader taught in the game also serves a functional purpose in the interface: whenever the player looks at an object that can be targeted by a spell, that object is highlighted with an outline, signalling interactivity. In this way one of the taught concepts is itself used as a core interaction cue.

Figure 3.2: The stylized low-poly open world of *Shwizard*. A compass at the top of the screen marks points of interest with question marks.

Figure 3.3: The radial spell wheel used to select unlocked spells, here with the *Greenify* spell selected. The wheel displays the spell name and the shader topic that unlocked it (here: the fragment shader).

Since spells can be applied to most objects in the game world, the player is also given a dedicated reset spell that reverts the targeted object to its original state. Nearly the entire environment can be manipulated this way; the only exceptions are world boundaries and other objects whose alteration could break the game or allow the player to leave the intended play area.

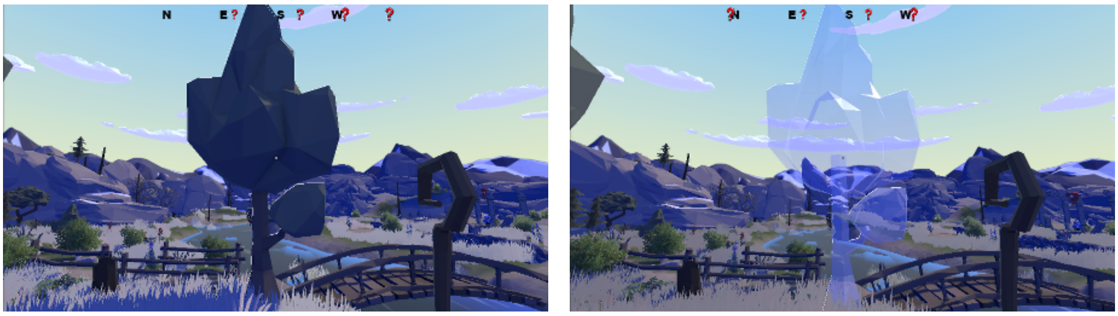
When a quest's objective is fulfilled, a floating green checkmark appears above the corresponding quest giver to confirm completion (Figure 3.5).

Figure 3.4: An interactable object highlighted with a white outline when the player looks at it. The outline shader, taught as an encyclopedia topic, doubles as the interaction cue for spell targeting.

Figure 3.5: Visual feedback on quest completion: a floating green checkmark appears above the quest giver once the objective has been fulfilled.

Initially, the plan was to cache the entire original object whenever a shader was applied to it. This idea was discarded for performance reasons; instead, only the original material of each modified object is cached.

Telemetry was implemented by encoding session variables as URL parameters when the player finishes the game and is redirected to the post-game questionnaire. This avoided the need for a dedicated backend or user accounts, kept the study fully anonymous, and made the game session and questionnaire data trivially linkable.



(a) Before casting. (b) After casting.

Figure 3.6: A spell applying an alpha-transparency ("fade") shader to a tree. The immediate, visible effect of casting a spell was the most praised element in participant feedback (see Chapter 4).

3.2.1 Implemented Spells

Table 3.1 lists all spells implemented in *Shwizard*, together with the shader topic that unlocks each spell and the effect it applies. Each spell becomes available once the player has passed the quiz of the corresponding encyclopedia topic; the *Fade* and *Reset* spells are available from the start.

Table 3.1: All implemented spells, their unlocking topics, and their effects.

Spell	**Unlocked by**	**Effect**
Greenify	Fragment Shader	Recolors the target object green
Fade	— (default)	Makes the target object transparent
Reset	— (default)	Reverts the target to its original state
Heat	Ambient Light	Shifts the ambient light to a warm orange tone
Glow	Spot Light	Conjures a light-emitting sphere
Liquify	Water Shader	Turns the target translucent blue and animates it with a wave-like motion

3.2.2 Quests

The game contains four quests, each requiring the player to apply specific spells to objects in the world. The quests thereby function as applied exercises: they can only be solved with spells whose underlying shader topics have been studied and passed.

Quest 1 – Reviving a tree. The quest giver asks the player to bring a dead tree back to life. The quest is solved by casting the *Greenify* spell on a specific tree next to the quest giver, recoloring its bare crown green.

Quest 2 – Warming the region. The quest giver complains that the region has grown too cold. The player solves this by casting the *Heat* spell, shifting the ambient lighting from a cold blue to a warm orange tone and thereby changing the perceived temperature of the entire area.

Quest 3 – Repopulating the pond. The player is asked to help bring fish back to a small pond, which initially holds still, grey water. Casting the *Liquify* spell transforms the surface into lively, animated blue water, making the pond hospitable again and completing the quest.

Quest 4 – Light in the cave. For the final quest, the player must help the quest giver find a lost item inside a dark cave. Casting the *Glow* spell inside the cave illuminates it, allowing the quest giver to locate the item.

3.3 Considered Features

Several features were considered during development but ultimately not implemented, primarily to keep the scope of a bachelor project manageable and the play session short enough for a study protocol.

- **A shader code view:** an optional panel showing simplified GLSL code alongside each spell was considered, to bridge toward real shader programming. It was cut to keep the experience strictly code-free for the target audience.
- **Additional topics:** more advanced concepts such as physically based rendering, shadow mapping, and post-processing were drafted but excluded to keep the expected playtime short.
- **Adaptive difficulty and spaced repetition** of quiz questions, which would have required per-player state persistence beyond a single session.
- **A quest progress indicator** in the HUD. This was de-prioritized during development—and, as the results in Chapter 4 show, it became a heavily requested improvement in participant feedback.
- **Separate area for quests and sandbox to experiment with the spells:** This was initially considered, but it would have exceeded the available development time.

3.4 Study Design and Procedure

The study was designed to address the three research questions defined in Chapter 1. It followed a pre-post design consisting of three phases: a *pre-game questionnaire*, a *game session*, and a *post-game questionnaire*, completed in this order by each participant. The two questionnaires were linked by a self-generated anonymous participant code.

- **Pre-game questionnaire:** Collected demographics, five knowledge questions about shader concepts, and attitude ratings on interest in the topic and game-based learning. It was distributed before participants received the game link, so the game could not influence baseline responses.
- **Game session:** Participants played *Shwizard* in the browser. The intended session ended with the completion of all four quests, but participants were free to return to the study earlier without finishing every quest. The game recorded lightweight telemetry—total play time, completed quests, completed encyclopedia entries, and incorrectly answered quiz questions—which was appended to the post-game questionnaire URL.
- **Post-game questionnaire:** Repeated the five knowledge questions and collected experience ratings, NASA Task Load Index (NASA-TLX) workload items, a game-versus-textbook comparison, a recommendation rating in the form of the Net Promoter Score (NPS) [12], and open-ended feedback.

Before starting the pre-game questionnaire, participants were shown a study information and consent form (see: Appendix C) describing the purpose of the study, the data collected, its handling under the General Data Protection Regulation (GDPR), and their rights; participation required agreement with this form.

To link the two questionnaires without collecting identifying information, each participant generated the same anonymous code in both questionnaires from information only they would reliably reproduce: the first two letters of their mother's first name, the two-digit day of their own birth, and the last two letters of their place of birth. For example, a participant whose mother is named Maria, who was born on the 2nd of a month in Vienna, would generate the code `ma02na`. This scheme creates an identifier for matching pre- and post-game responses while remaining anonymous, as the code cannot be traced back to the participant by the investigators.

The five knowledge questions included an explicit "I don't know" option to discourage guessing, and each was answerable from the game's encyclopedia content. The questionnaire items deliberately did not replicate the in-game quiz questions word for word: while covering the same concepts, their phrasing and answer options were slightly varied. This was done so that the post-game measurement captures understanding of the underlying concept rather than recognition of a memorized question–answer pair from the game—had the items been identical, a participant could improve their score simply by recalling which option they had clicked to pass the quiz, without having understood the material. All experience rating scales used a forced four-point format without a neutral midpoint, with each point anchored by a descriptive phrase.

Recruitment took place through the investigators' personal networks and university channels. This convenience-sampling approach introduces a self-selection bias that should be kept in mind when interpreting the results: participants were not drawn randomly from the intended target population of complete beginners, but from a pool likely to over-represent people with a technical or academic background, an interest in games, and a pre-existing willingness to help with a student study. As reported in Chapter 4, the resulting sample was indeed predominantly male and considerably more experienced with shaders than the target audience, with a large share already achieving the maximum knowledge score before playing. The sample is therefore not representative of the population the game is ultimately designed for, and the implications of this are discussed alongside the other threats to validity in Section 4.3.5.

Knowledge gain was tested with a paired t-test and a Wilcoxon signed-rank test, with Cohen's d for paired samples as effect size; the significance threshold was set to $\alpha = 0.05$. The two tests were used in combination because the paired t-test assumes normally distributed difference scores, an assumption that cannot be taken for granted for a bounded 0–5 knowledge score in a sample of this size. The Wilcoxon signed-rank test, as its non-parametric counterpart, makes no such distributional assumption and therefore serves as a robustness check: agreement between both tests indicates that the conclusion does not depend on the normality assumption. Continuous variables are reported as mean and standard deviation; ordinal rating data are reported as response counts with median and Interquartile Range (IQR). Associations between variables were explored using Spearman rank correlations, as the involved variables are largely ordinal. For details on the statistical methods used, we refer the reader to standard works on statistical analysis [13, 14].

CHAPTER 4

Results

4.1 Evaluation

4.1.1 Participants

The 41 participants who completed the full protocol had a mean age of 26.6 years (SD = 5.0, range 19–44). The sample comprised 31 men, seven women, and 3 participants who preferred not to disclose their gender. 17 were students, eleven employed, nine both studying and employed, and four reported other occupational statuses.

Prior familiarity with the subject was mixed but skewed towards technical: 21 participants had experimented with shaders a little and 6 used shaders regularly, while 11 had heard of shaders without using them and only 3 had never heard of them. Baseline interest in the topic was high (median 3 of 4, IQR $[2, 4]$; 18 very interested, 11 moderately, 10 slightly, 2 not at all). Attitudes toward game-based learning were favorable (median 3 of 4, IQR $[2, 3]$): 20 participants somewhat and 10 strongly preferred learning through games, 11 somewhat preferred traditional methods, and none strongly preferred traditional methods.

Telemetry showed a mean play time of 17.8 minutes (SD = 10.7, range 2.8–45.2). Participants answered on average 4.2 quiz questions incorrectly during play (SD = 5.1, median 2). NASA-TLX workload ratings on 0–10 scales were low across the demand-related dimensions: mental demand M = 3.2 (SD = 2.1), temporal demand M = 2.3 (SD = 2.2), effort M = 3.0 (SD = 2.4), and frustration M = 3.0 (SD = 2.6). The performance item was rated at M = 4.8 (SD = 3.0, median 5), the only dimension centered around the middle of the scale. Figure 4.1 visualizes these distributions; the demand-related ratings cluster in the lower third of the scale, with only isolated participants reporting high effort or frustration.

An analysis of which quiz questions were answered incorrectly showed a broad distribution rather than isolated problem questions: 29 of the 30 quiz questions were missed by at least one participant, and no single question was missed by more than 11 of the 41 participants. Since the telemetry records each question at most once per participant, these figures count distinct questions failed at least once rather than repeated attempts. At topic level, however, a moderate pattern emerges when the number of participants who completed each topic is taken into account: the noise functions, outline shader, and general lighting topics produced the highest error rates (0.78–0.92 recorded errors per completing participant), while the introductory general topic produced by far the lowest (0.16)—its first question being the only one in the game that no participant ever answered incorrectly. The three most frequently missed individual questions were the first fragment shader question and the second and third outline shader questions (11 participants each). This ordering is consistent with conceptual difficulty rather than flawed item wording: the topics with the highest error rates are precisely those introducing the most novel and abstract mechanisms.

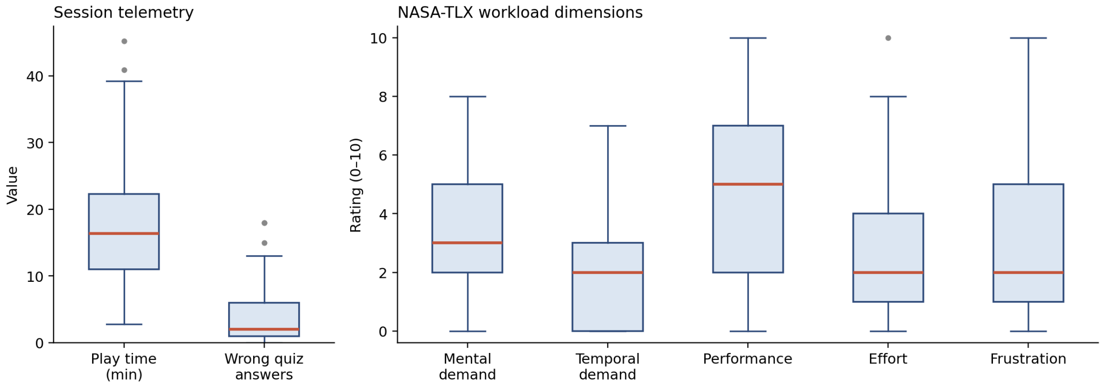


Figure 4.1: Distributions of session telemetry (left) and NASA-TLX workload ratings (right) as boxplots (n = 41). Boxes span the interquartile range, the line marks the median, and dots mark outliers. For all workload dimensions, higher values indicate greater perceived load and are therefore less favorable, except for the performance item, where higher values indicate greater self-assessed success.

4.1.2 Knowledge Gain

Knowledge scores increased from a pre-game mean of 3.71 out of 5 (SD = 1.71, median 5) to a post-game mean of 4.59 (SD = 0.92, median 5). The mean gain of 0.88 points (SD = 1.60) was statistically significant (Wilcoxon signed-rank test: W = 23.0, p = 0.001), with a rank-biserial correlation value of r = 0.51 indicating a medium to large effect.

Figure 4.2 shows how the distribution shifted. The number of participants answering all five questions correctly rose from 21 to 32, while every lower score level decreased. Nineteen participants improved their score, twenty remained unchanged, and two scored lower after playing (by one and by four points, respectively). Notably, 19 of the 20 unchanged participants had already achieved the maximum score before playing, so no improvement was possible for them; among participants with room to improve, the large majority did improve. The participant whose score dropped by four points played for under four minutes and switched four previously correct answers to “I don’t know” in the post-test, a pattern indicative of disengaged responding rather than knowledge loss.

Because more than half of the sample achieved the maximum score before playing, the aggregate effect necessarily understates what the game achieved for those who had something to learn. A subgroup analysis restricted to the 20 participants with a pre-game score below the maximum makes this visible: their scores rose from M = 2.35 (SD = 1.53) to M = 4.40 (SD = 0.94), a mean gain of 2.05 points (SD = 1.32). The improvement was near-universal in this group—19 of the 20 improved, one remained unchanged, and none declined—and 13 of them answered all five questions correctly after playing. Filtering out the participants who already scored perfectly before playing more than accounts for the ceiling effect: the rank-biserial effect size rises from $r = 0.51$ in the full sample to $r = 0.87$ in this subgroup (Wilcoxon signed-rank: $W = 0.0$, $p < 0.001$), roughly a threefold increase. Among the ten participants with the lowest prior knowledge (pre-game score of two or below), the mean score rose from 1.0 to 4.0.

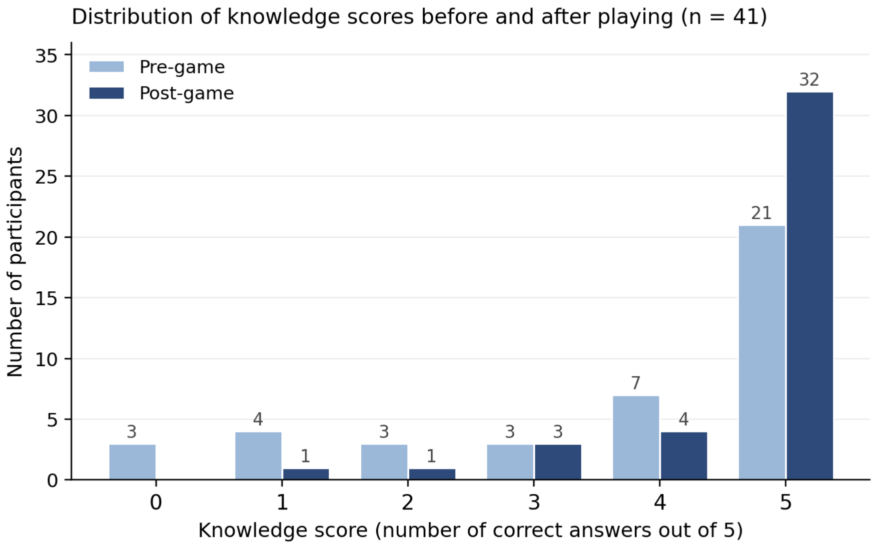


Figure 4.2: Distribution of knowledge scores before and after playing (n = 41). The number of participants answering all five questions correctly rose from 21 to 32, while every lower score level decreased.

Improvement was observed on every one of the five knowledge questions (Table 4.1). The largest gains occurred for the fragment shader, distance falloff, and noise functions (+22.0 percentage points each), while the smallest gain occurred for ambient light (+9.8), which had the highest baseline. After playing, every question was answered correctly by at least 85 % of participants.

Table 4.1: Per-question correct answer rates before and after playing (n = 41).

Question	**Pre correct**	**Post correct**	**Change (pp)**
Fragment shader	63.4 %	85.4 %	+22.0
Vertex	80.5 %	92.7 %	+12.2
Distance falloff	73.2 %	95.1 %	+22.0
Ambient light	82.9 %	92.7 %	+9.8
Noise function	70.7 %	92.7 %	+22.0

4.1.3 Experience Ratings

All experience items are summarized as response counts with median and IQR on the underlying 1–4 scale (1–5 for the textbook comparison). Figure 4.3 visualizes the response distributions of the four-point items, Figure 4.4 those of the game-versus-textbook comparison.

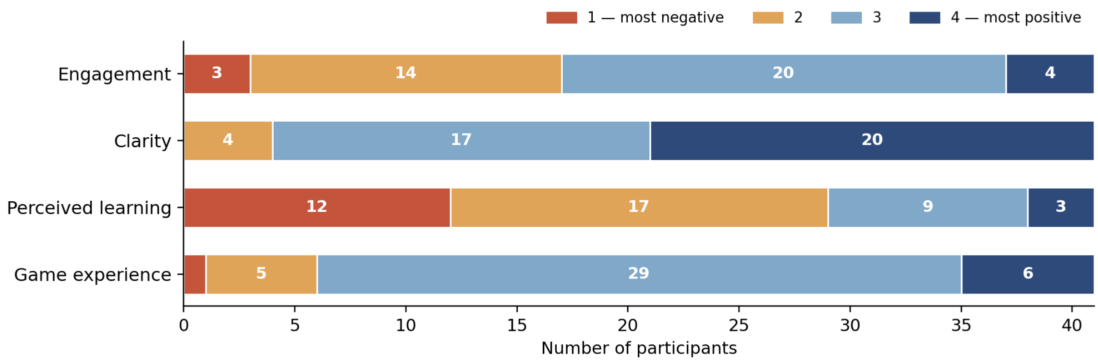


Figure 4.3: Response distributions of the four-point post-game experience ratings (n = 41).

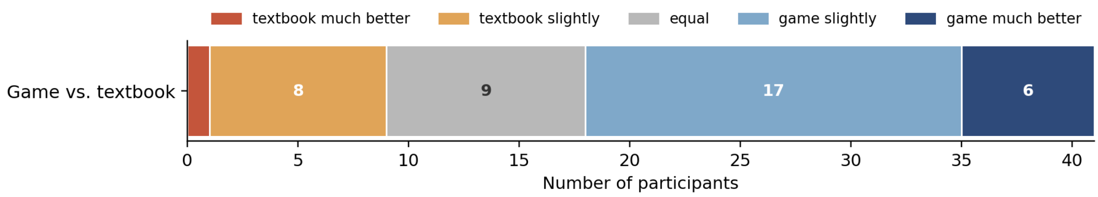


Figure 4.4: Response distribution of the five-point game-versus-textbook comparison (n = 41).

Perceived learning was more moderate, with a median of 2 (IQR $[1, 3]$); it was the only item whose distribution leaned toward the negative end of the scale, with a majority of participants reporting that they had learned little or nothing new. The game-versus-textbook comparison leaned in the opposite direction (median 4 of 5, IQR $[3, 4]$): a majority judged the game to be more effective than a textbook, while only a small group favored the textbook and just one participant selected the strongest pro-textbook option. Willingness to recommend the game was measured with the NPS [12], which classifies ratings of 9–10 as promoters, 7–8 as passives, and 0–6 as detractors, and subtracts the percentage of detractors from that of promoters. With a mean rating of 6.6 (SD = 2.5, median 7), promoters were outnumbered by detractors, yielding an NPS of −24.

4.1.4 Correlation Analysis

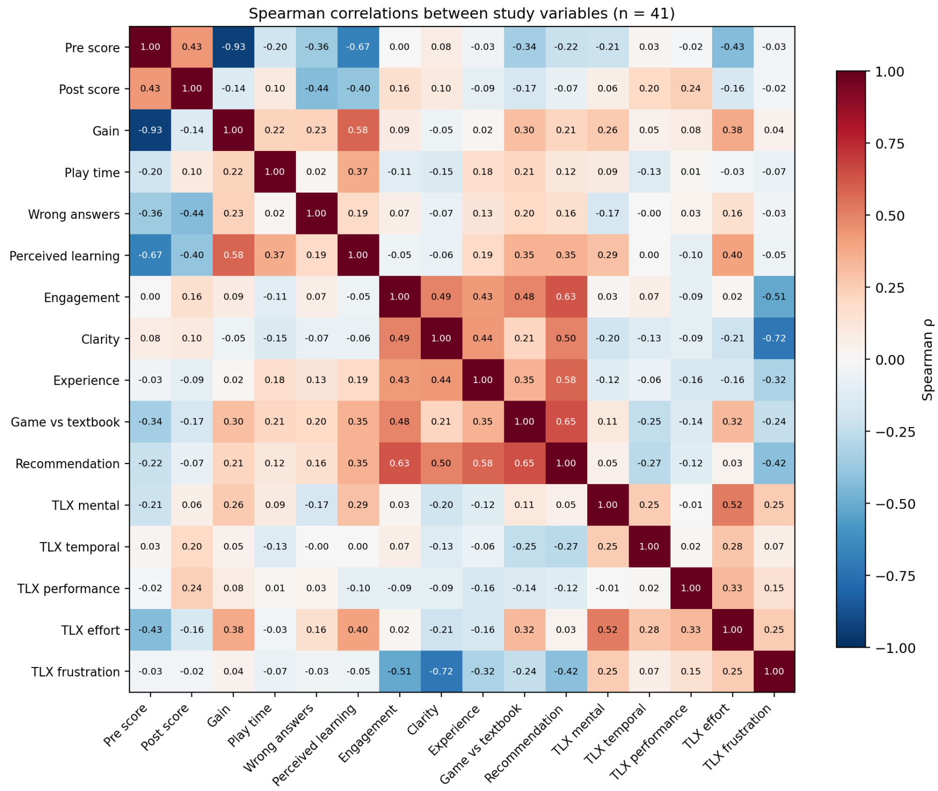


Figure 4.5: Spearman rank correlations between study variables (n = 41).

Figure 4.5 shows Spearman correlations between the main study variables, including the individual NASA-TLX workload dimensions. Three patterns stand out. Firstly, knowledge gain was almost perfectly negatively correlated with pre-game knowledge score ($\rho = -0.93$), the statistical signature of the ceiling effect: participants who started at the maximum could not improve. Second, perceived learning tracked actual learning: participants' self-reported learning correlated substantially with their measured gain ($\rho = 0.58$) and negatively with their prior knowledge ($\rho = -0.67$), indicating that self-reports were grounded in reality rather than mere enthusiasm. Third, the experience measures formed a coherent cluster: engagement correlated with clarity ($\rho = 0.49$), overall experience ($\rho = 0.43$), the game-versus-textbook judgment ($\rho = 0.48$), and the recommendation rating ($\rho = 0.63$).

Among the workload dimensions, frustration was the only one substantially associated with this cluster: it correlated strongly negatively with clarity ($\rho = -0.76$) and negatively with engagement ($\rho = -0.51$), overall experience ($\rho = -0.44$), and recommendation ($\rho = -0.52$), whereas mental demand, temporal demand, and effort showed no meaningful association with how positively participants evaluated the game (all $|\rho| \leq 0.29$). The dimensions were therefore analyzed individually rather than aggregated into a Raw NASA-TLX score [15], since a composite would have masked this distinction. Effort was the only workload dimension related to learning, correlating moderately with knowledge gain ($\rho = 0.35$). Taken together, this suggests that perceived clarity of the learning content—and the frustration that arises in its absence—rather than cognitive workload or knowledge gain itself, was a primary driver of how positively participants evaluated the game. As noted in the methodology, all correlations are rank-based; a comparison with Pearson correlations revealed no additional strong associations, so the reported patterns are not an artifact of the chosen coefficient.

4.2 Summary of Responses

Open-ended responses about what participants enjoyed most overwhelmingly centered on the spell mechanic: applying shaders to objects and seeing the effect immediately and in real time was the single most frequently praised element, followed by the relaxed pacing and the visual presentation of the world. Several participants explicitly connected the mechanic to their understanding, describing that seeing a shader visibly change a surface made the written explanation concrete.

Suggestions for improvement most commonly concerned orientation and progress feedback—above all a visible indication of how many quests remain—followed by requests for more content and more varied spell effects, and minor usability refinements in the encyclopedia and quiz interface. No participant reported technical problems severe enough to prevent completion. The qualitative shift in the open-ended “What is a shader?” question mirrored the quantitative results: pre-game answers were frequently vague or tentative, while post-game answers more often referenced concrete mechanisms such as per-pixel color calculation.

4.3 Discussion

For reference, the research questions as stated in 1 are listed again below:

- **RQ1:** Does playing *Shwizard* lead to a measurable increase in players' factual knowledge of shader concepts?
- **RQ2:** Do players find the game-based learning format engaging and clear?
- **RQ3:** Do players perceive the game as more effective than a traditional learning method such as reading a textbook?

4.3.1 Knowledge Gain (RQ1)

The results provide clear evidence for RQ1: Playing *Shwizard* was associated with a statistically significant and practically meaningful increase in factual knowledge of shader concepts, with improvement on every question measured. The effect sizes reported in the results are rank-based (r), as is appropriate for the ordinal knowledge scores. To additionally place the effect in the context of the serious games literature—which conventionally reports Cohen's d—we also computed d for the full sample. At $d = 0.55$, the effect exceeds the average meta-analytic effect of serious games over conventional instruction reported by Wouters et al. ($d = 0.29$) [4]. This comparison should be treated cautiously, since the present study measured gain against a no-instruction baseline rather than against an alternative teaching method, and since—as the subgroup analysis in Section 4.1.2 showed—the full-sample figure is itself deflated by the ceiling effect.

The magnitude of the gain must be interpreted in light of a substantial ceiling effect, visible directly in the near-perfect negative correlation between prior knowledge and gain ($\rho = -0.93$). The sample was more technically experienced than the intended target audience, so the aggregate figure likely understates the true effect for the intended audience of complete beginners. At the same time, the pre-post design cannot isolate the specific contribution of the game mechanics: the gain may stem from the encyclopedia texts, the testing effect of the in-game quizzes, the interactive spell mechanics, or their combination. A control condition reading the same texts without the game would be required to disentangle these factors.

4.3.2 Engagement and Clarity (RQ2)

The experience data support a positive answer to RQ2, though with an important qualification. Clarity was rated highly—90 % of participants found the information mostly or very clear—and perceived workload and frustration were low. Engagement, however, was more equivocal: while 59 % found the game quite or very engaging, this is only a modest majority, and a substantial minority found it merely slightly engaging or not engaging at all. One plausible explanation lies in the format of the learning content itself. Knowledge in *Shwizard* is conveyed primarily through written encyclopedia entries, which—despite being embedded in a game world—structurally resemble the textbook the game is meant to improve upon. The reading-heavy delivery may have limited how engaging the experience felt, even as it kept the content clear.

The correlation analysis adds a further nuance: clarity was among the variables most strongly connected to the rest of the experience cluster, with an especially strong negative association with frustration ($\rho = -0.76$). For an educational game, this suggests that the comprehensibility of the learning content is not merely a pedagogical property but a core driver of the overall user experience—participants who found the content clear enjoyed the game, and those who struggled with it became frustrated regardless of the game's other qualities.

4.3.3 Perceived Effectiveness (RQ3)

The answer to RQ3 leans positive: 56 % of participants judged the game more effective than a textbook and only 22 % favored the textbook, just one of them strongly. Perceived learning was modest in absolute terms—a median of "I learned a little"—but the correlation analysis shows this reflects honest self-assessment in a knowledgeable sample rather than a failure of the game: perceived learning tracked measured gain closely ($\rho = 0.58$), and the participants who reported learning nothing new were overwhelmingly those who already knew the material. The negative NPS (-24) alongside positive experience ratings indicates participants found the game good but not remarkable—a reasonable verdict on a research prototype with limited content, and one consistent with the recommendation rating being driven primarily by engagement ($\rho = 0.63$) and the perceived advantage over traditional formats ($\rho = 0.65$) rather than by learning outcomes.

4.3.4 Design Reflections

The qualitative feedback validates the central design decision: making shader effects tangible through castable spells was the most praised element, and several participants explicitly credited it with making the concepts concrete—one noting that being able to apply the water shader to a rock "demystifies the concept" for beginners. The most consistent criticism concerned discoverability of the controls: at least eleven participants reported that the key bindings were shown only once at the start and could not be recalled afterwards, leaving some "lost" or pressing keys at random until a spell fired. This is the same class of orientation problem as the frequently requested quest progress indicator, and both point to affordances that were deprioritized during development (Section 3.3); notably, one participant explicitly attributed their lower clarity rating to the inability to review the controls, directly linking this gap to the clarity measure discussed above. The strict three-correct-answers unlock requirement, by contrast, appears well calibrated: the median participant answered only two distinct quiz questions incorrectly in the whole session, and frustration remained low.

4.3.5 Threats to Validity

Several threats to the validity of these results must be acknowledged.

Internal validity. The single-group pre–post design cannot isolate the cause of the knowledge gain: the improvement may stem from the game mechanics, from the encyclopedia texts alone, or from the testing effect of answering the same questions twice. Without a control condition, the gain can be attributed to the intervention as a whole but not to its game-specific components. In addition, one participant's four-point score drop was traced to disengaged responding, illustrating that unsupervised remote participation permits low-effort data; this case was retained in the analysis, making the reported effect conservative.

Construct validity. Knowledge was measured with five multiple-choice questions, which capture factual recall rather than deeper understanding or the ability to apply shader concepts. The experience measures are self-reports on four-point scales; while the substantial correlation between perceived and measured learning ($\rho = 0.58$) supports their credibility, self-reports remain susceptible to demand characteristics—participants knew the game was the author's work.

External validity. The sample was self-selected, recruited through personal and university networks, predominantly male, and considerably more experienced with the subject matter than the intended target audience: over half of the participants achieved the maximum knowledge score before playing. The resulting ceiling effect means the aggregate gain likely understates the effect for genuine beginners, but it equally means the results cannot be claimed to generalize to that audience without a dedicated study. Finally, knowledge was measured immediately after play; nothing can be concluded about longer-term retention.

4.3.6 Participant Dropout

Of the 64 participants who completed the pre-game questionnaire, 23 (36 %) did not complete the full protocol. While some attrition is expected in an unsupervised, unpaid online study, a dropout of this magnitude warrants comment, particularly as it interacts with the discoverability problems reported in the feedback. The most frequently raised issue—that the control bindings and the game menu were accessible only via a key (Tab) that was not signposted after the initial instructions—plausibly contributed to the dropout: several participants described feeling lost or being unable to find how to proceed, and the same menu that was hard to locate also contained the link back to the post-game questionnaire. It is therefore likely that some participants who became stuck either abandoned the study or returned to the survey before completing the intended play session. This interpretation is consistent with the telemetry of those who did reach the post-game questionnaire: five of the 41 submitted after completing only a single quest, and five played for under five minutes. The dropout is thus not only a threat to the representativeness of the sample—those who persevered may differ systematically from those who left—but also, in part, a measurable symptom of the same usability shortcomings identified as the primary target for future work.

CHAPTER 5

Conclusion

This thesis presented *Shwizard*, a serious game developed in Unity WebGL to teach foundational shader concepts to players without technical background, and evaluated it in a pre-post study with 64 participants (41 complete cases).

5.1 Summary of Findings

Several conclusions can be derived from the data. First, the game demonstrably teaches. Knowledge scores rose significantly (M = 3.71 to M = 4.59 of 5, $p = 0.001$, $d = 0.55$), with gains on every question, and the effect is likely understated due to a strong ceiling effect in a technically experienced sample—over half of the participants started at the maximum score, while those who started below it almost universally improved. Second, self-perception tracked reality: perceived learning correlated substantially with measured gain ($\rho = 0.58$), lending credibility to the self-report instruments. Third, clarity of the learning content emerged as a pivotal experience variable, strongly negatively associated with frustration ($\rho = -0.76$) and positively with every other experience measure—in an educational game, comprehensible content *is* the user experience. Fourth, the central design metaphor worked: the ability to cast spells and see shader effects applied in real time was the most praised element and, per participant accounts, the mechanism that made abstract concepts concrete. Finally, a majority of participants (56 %) judged the game more effective than a textbook, while the modest recommendation score (NPS −24) reflects the honest limits of a research prototype.

These findings align with the broader serious games literature, which reports positive but design-dependent learning effects [4, 5], and extend it to a previously unaddressed niche: Shader education for complete beginners.

5.2 Future Work

The main limitations identified in Section 4.3.5—the absence of a control condition, the ceiling effect, a self-selected and predominantly male sample, and the lack of a delayed retention measure—directly define the agenda for future work. The most immediate priority is a controlled comparison against a non-game version of the same content, which would isolate the contribution of the game mechanics from that of the educational texts. A study with genuine beginners would provide a valid test of the design goal, and a delayed post-test would measure retention rather than immediate recall.

On the design side, the open-ended feedback points to a concrete, user-grounded roadmap of improvements, ordered here roughly by how frequently participants raised them:

- **Persistent control hints (13 mentions)** By far the most common request: the key bindings are shown only once at startup and cannot be recalled afterwards. Displaying them in a pause menu or as a permanent on-screen element would address the single most frequent complaint. Since the same discoverability gap plausibly contributed to the study's dropout rate (Section 4.3.6), improving it may also raise completion rates in future evaluations.
- **Quest orientation (10 mentions)** A visible indicator of the current quest and of remaining progress—a HUD counter or a quest sidebar—would resolve the disorientation several participants reported.
- **Tighter integration of learning and play (6 mentions)** Instead of front-loading all encyclopedia reading before play, participants suggested unlocking spells gradually as their topics are studied, or embedding the concepts more directly into quests, so that the game does not resemble a textbook followed by a sandbox.
- **More and longer content (7 mentions)** Additional quests, more varied spell effects, and a longer overall experience were repeatedly requested—consistent with the game's status as a prototype.
- **Explicit quest–spell links (5 mentions)** Several participants found it unclear which topic a given quest required, and suggested that quest givers name or prompt the relevant concept directly.
- **A mouse-sensitivity setting (4 mentions)** A recurring usability issue that a single options entry would resolve.

A more fundamental direction, raised by several of the more experienced participants, questions the code-free premise itself: they suggested that deeper understanding would come from letting players read, assemble, or tweak actual shader logic—for instance adjusting a vertex shader to make a flag wave—rather than only observing effects. This would move the game closer to the "learning by doing" that the current design deliberately avoids for accessibility, and represents a distinct design branch rather than an incremental improvement; whether it can be done without reintroducing the code barrier that *Shwizard* set out to remove is an open question.

Overall, the data support the conclusion that a serious game can make an intimidating technical subject accessible and measurably teach its fundamentals within a single twenty-minute session—provided the content is clear, and the game makes the invisible visible.

APPENDIX A

Example Encyclopedia Topic

The following is the complete in-game encyclopedia entry for the *Water Shader* topic, together with its three quiz questions, as an example of the educational content. The correct answers are marked in bold.

Water Shader

A water shader combines several techniques to fake the look of moving water—all without simulating real physics. It is one of the most common shader effects in games and a great example of how multiple simple ideas can combine into something convincing.

As explained in the "Noise" topic, movement is usually created using noise functions. Color also plays a big role. Real water looks darker in deep areas and lighter in shallow ones, so the shader blends between colors based on depth. Reflections and a small amount of transparency are often added on top to further sell the illusion.

Water shaders are a perfect example of what you might call creative cheating—none of it is real physics, but the result looks convincing enough that players never stop to question it. This was especially important when computer hardware was still much weaker than it is today. Understanding that real-time graphics is largely built on clever illusions is one of the most useful mindset shifts a shader beginner can make.

Quiz Questions

1. What creates the rippling movement in a typical water shader? A video of real water playing in the background / Physics simulation running on the CPU / **A noise function that shifts over time** / Randomly repositioning vertices every frame

2. Why do water shaders blend between different colors? To match the color of nearby objects automatically / **To simulate how real water looks darker in deep areas and lighter in shallow ones** / Because the GPU can only display two colors at once / To make the water shader run faster
3. What does "creative cheating" mean in the context of water shaders? Copying another game's shader code directly / Using real-world water physics to calculate every ripple / **Using simple tricks that look convincing without simulating reality** / Hiding low-quality graphics behind visual effects

APPENDIX B

Questionnaire Instruments

B.1 Pre-Game Questionnaire

Demographics: age (open numeric); gender (Man / Woman / Non-binary / Prefer not to say / Prefer to self-describe); current occupation (multi-select: student, employed, self-employed, unemployed).

Prior experience: experience with shaders (never heard of shaders / heard of but never used / experimented a little / use regularly); familiarity with 3D computer graphics concepts (not at all / slightly / moderately / very familiar).

Knowledge questions (each with four options including "I don't know"; correct answers shown here):

1. What does a fragment shader do? — *Determines the color of each pixel on screen*
2. What is a vertex in a 3D model? — *A point that defines the shapes of the geometry*
3. What does "distance falloff" mean in lighting? — *Light gets dimmer the further it travels from the source*
4. What is ambient light? — *A constant brightness applied to all surfaces*
5. What is a noise function used for in shaders? — *Creates smooth, organic-looking randomness*

Attitudes (four-point forced choice): interest in learning about shaders (not at all / slightly / moderately / very interested); preference for learning through games versus traditional methods (strongly prefer traditional / somewhat prefer traditional / somewhat prefer games / strongly prefer games).

Open question: "In your own words: what do you think a shader is?"

B.2 Post-Game Questionnaire

Knowledge questions: identical to the five pre-game questions.

Experience ratings (four-point forced choice with anchored labels): perceived learning (nothing new / a little / quite a bit / a lot); engagement (not engaging at all / slightly / quite / very engaging); clarity of information (very unclear / somewhat unclear / mostly clear / very clear); overall game experience (very poor / below average / good / excellent).

Workload: NASA-TLX items on 0–10 scales (mental demand, temporal demand, performance, effort, frustration).

Comparison: "Did the game format help you understand the topics better than you think reading a textbook would have?" (textbook much more effective / textbook slightly more effective / both about equal / game slightly more effective / game much more effective).

Recommendation: likelihood of recommending the game, 0–10.

Open questions: unclear or confusing topics; most enjoyed aspect; suggested improvements; "In your own words: how would you now explain what a shader is to a friend?"

APPENDIX C

Study Information and Consent Form

C.1 Study Information

Investigator: Jonathan Seyr (e12119032@student.tuwien.ac.at)

Supervision / Principal Investigator: Proj.Ass.in Annalena Ulschmid, MSc BSc (annalena.ulschmid@tuwien.ac.at, +43 1 58801 18646), TU Wien – Faculty of Informatics, Institute of Visual Computing & Human-Centered Technology, Favoritenstraße 9–11 / E193-02, A-1040 Vienna, Austria

Data Protection Officer: Mag. Christina Thirsfeld (datenschutz@tuwien.ac.at, +43 1 58801 41000), TU Wien, Karlsplatz 13/018, A-1040 Vienna, Austria

C.2 Purpose and Procedure of the Study

The aim of this study is to investigate whether an educational game in the field of computer graphics can serve as a meaningful learning tool – either as a supplement or as an alternative to traditional teaching formats such as lectures. As part of the study, you will be asked to play an online game and afterwards complete a short survey. The purpose of the study is not to evaluate your performance. We therefore kindly ask you to answer questions only when you believe you know the correct answer.

The expected duration of participation is approximately 20 minutes.

C.3 Voluntary Participation

Participation in this study is entirely voluntary. You may withdraw your consent at any time and without giving reasons, with no disadvantages arising for you as a result. You may discontinue the study at any time.

C.4 Potential Risks and Benefits

Participation in this study involves no known risks. The study contributes to scientific knowledge in the field of computer science education.

C.5 Data Protection and Confidentiality

The collection, processing, and storage of your personal data is carried out in accordance with the General Data Protection Regulation (GDPR) and applicable national data protection law. The following data will be collected as part of the study:

- Demographic information: age, gender, level of computer graphics education, occupation
- Measurement data from playing the game: time, completion, correctness of answers / failed attempts

Your data will be processed in pseudonymized or anonymized form. This means that your information is separated from your name and assigned a code. Only the principal investigator can link the data back to your person, and only via a separately and securely stored code list. The results of the study will be published exclusively in anonymized form; no conclusions can be drawn about your identity.

The anonymized data collected and the results derived from them will be used for scientific publications in relevant academic journals. If you wish to have the data collected about you in the course of this study deleted, you may send an e-mail to the principal investigator no later than 01.07.2026, stating your user ID.

C.6 Your Rights Under the GDPR

You have the right at any time to:

- Access the data stored about you (Art. 15 GDPR)
- Rectification of inaccurate data (Art. 16 GDPR)
- Erasure of your data (Art. 17 GDPR)

- Restriction of processing (Art. 18 GDPR)
- Data portability (Art. 20 GDPR)
- Object to processing (Art. 21 GDPR)
- Lodge a complaint with a data protection supervisory authority (Art. 77 GDPR)

You may withdraw your consent at any time with effect for the future. The lawfulness of the processing carried out prior to the withdrawal remains unaffected. Please note that subsequent deletion of your data is only possible as long as your data can still be attributed to you personally (i.e., before full anonymization has taken place).

Overview of Generative AI Tools Used

The AI assistant Claude (Anthropic) was used as a writing aid throughout this thesis, primarily for revising text according to the author's specifications and for writing evaluation scripts. Some graphical assets in the game were also created using Claude. The study concept, game design, game implementation, data collection, and all final decisions regarding the content and its interpretation are the author's own. All AI-assisted content was reviewed, verified, and edited by the author, who takes full responsibility for the entire thesis.

List of Figures

List of Tables

Acronyms

GLSL OpenGL Shading Language. 5, 11

HUD Heads-Up Display. 7, 11

IQR Interquartile Range. 13, 14, 17

NASA-TLX NASA Task Load Index. 12, 14, 15, 18, 19, 30

NPS Net Promoter Score. 12, 18, 21, 24

Bibliography

[1] M. Zyda, "From visual simulation to virtual reality to games," *Computer*, vol. 38, no. 9, pp. 25–32, 2005. [Online]. Available: https://doi.org/10.1109/MC.2005.297

[2] P. G. Vivo and J. Lowe, "The book of shaders," https://thebookofshaders.com, accessed: 2026-07-07.

[3] I. Quilez and P. Jeremias, "Shadertoy," https://www.shadertoy.com, accessed: 2026-07-07.

[4] P. Wouters, C. van Nimwegen, H. van Oostendorp, and E. D. van der Spek, "A meta-analysis of the cognitive and motivational effects of serious games," *Journal of Educational Psychology*, vol. 105, no. 2, pp. 249–265, 2013. [Online]. Available: https://doi.org/10.1037/a0031311

[5] Y. Zhonggen, "A meta-analysis of use of serious games in education over a decade," *International Journal of Computer Games Technology*, vol. 2019, p. 4797032, 2019. [Online]. Available: https://doi.org/10.1155/2019/4797032

[6] M. Muratet, P. Torguet, J.-P. Jessel, and F. Viallet, "Towards a serious game to help students learn computer programming," *International Journal of Computer Games Technology*, vol. 2009, p. 470590, 2009. [Online]. Available: https://doi.org/10.1155/2009/470590

[7] M. Moradi, N. F. B. M. Noor, and R. B. H. Abdullah, "The effects of problem-based serious games on learning 3D computer graphics," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 46, pp. 989–1004, 2022. [Online]. Available: https://doi.org/10.1007/s40998-022-00526-0

[8] P. Backlund and M. Hendrix, "Educational games – are they worth the effort? A literature survey of the effectiveness of serious games," in *Proc. 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, Poole, UK, 2013, pp. 1–8. [Online]. Available: https://doi.org/10.1109/VS-GAMES.2013.6624226

[9] M. Muratet, P. Torguet, F. Viallet, and J.-P. Jessel, "Experimental feedback on Prog&Play: A serious game for programming practice," *Computer*

Graphics Forum, vol. 30, no. 1, pp. 61–73, 2011. [Online]. Available: https://doi.org/10.1111/j.1467-8659.2010.01829.x

[10] M. Lysenko, "Shader school," https://github.com/stackgl/shader-school, accessed: 2026-07-24.

[11] Unity Technologies, "Unity engine," https://unity.com, version 2022.3 LTS, accessed: 2026-07-24.

[12] F. F. Reichheld, "The one number you need to grow," *Harvard Business Review*, vol. 81, no. 12, pp. 46–54, 2003.

[13] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.

[14] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 5th ed. London: SAGE Publications, 2018.

[15] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, 2006, pp. 904–908. [Online]. Available: https://doi.org/10.1177/154193120605000909