education sciences

MDPI

Article

When Farming Becomes Playable: A Qualitative Case Study of Triadic Gameplay Talk and Prompted Reflection on Hay Day in Fifth-Grade Geography

Michael Morawski * and Dina Pommerening

Institute of Geography and Primary Education, Faculty of Human and Social Sciences, University of Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany; dina.pommerening@uni-wuppertal.de
* Correspondence: morawski@uni-wuppertal.de

Abstract

Children increasingly encounter agriculture through mediated images and interfaces rather than sustained contact with farms. This exploratory qualitative case study examined (RQ1) what fifth-grade triads articulated while operating *Hay Day*, (RQ2) which game–reality comparisons and medium-related evaluations appeared in a prompted debriefing, and (RQ3) which design conjectures can be drawn from the contrast between these non-equivalent data production settings. Twenty-one students worked in seven triads during a four-lesson geography project. The corpus comprised audio-recorded triadic gameplay verbalizations with 14 rotating speaker-players, process notes, and written group responses. Structured qualitative content analysis combined theory-informed categories with inductively refined subcategories. Gameplay discourse centered on actions, resource dependencies, waiting, production, exchange, and progression; timers and level restrictions were salient but rarely interpreted as authored design choices. The debriefing elicited critique-oriented statements about omitted or simplified aspects of farming, but its materials supplied many relevant categories. Accordingly, thematic uptake is distinguished from student elaboration, and the debriefing is treated as prompt-responsive evidence rather than as spontaneous conceptual change. *Hay Day* is conceptualized as a playable agricultural model linking actionable game relations to agricultural relations, actors, spatialities, and consequences rendered absent or non-actionable. The findings support testable design conjectures, not a causal game effect.

Keywords: commercial video games; geography education; agricultural literacy; student conceptions; livestock farming; game literacy; prompted reflection; qualitative content analysis

Check for updates

Academic Editors: Simone Bassanelli, Federico Bonetti, Eftychia Roumelioti and Annapaola Marconi

Received: 16 July 2026
Revised: 1 October 2026
Accepted: 5 October 2026
Published: 13 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

For many children, agriculture is encountered at a distance. Food arrives through retail systems, animals appear in advertising and entertainment, and farming is often known through fragments rather than sustained experience of production. Recent reviews continue to report uneven agricultural literacy among school students, including difficulty connecting production, processing, distribution, consumption, and environmental consequences (Cosby et al., 2022; Manning et al., 2024). A food-systems account of literacy also distinguishes functional knowledge from the critical capacity to examine social, economic, environmental, and media conditions (Ares et al., 2024). Geography education has a particular stake in this problem because agriculture is simultaneously spatial, economic, ecological, political, and cultural.

Digital games are not marginal to the media lives of German children; recent national data document their routine use across mobile devices, computers, and consoles (Medienpädagogischer Forschungsverbund Südwest, 2025). For geography education, the question is therefore not whether such games belong to students' lifeworlds, but how their representations of socially relevant geographical issues can be made available for disciplinary analysis (Lux & Budke, 2020).

Digital games now belong to this representational landscape. Children do not merely watch farms in games; they cultivate, feed, harvest, collect, sell, upgrade, and wait. Commercial farming games therefore make certain agricultural relations available through action. Their models are compelling precisely because they are easy to operate: crops become feed, feed becomes eggs or milk, products become currency, and currency becomes expansion. Yet what is playable is also selective. Illness, mortality, slaughter, manure, labor, land conflict, regulation, market concentration, and ecological externalities may be absent, softened, or translated into variables that serve progression.

This selectivity matters educationally. Research on children's ideas about livestock farming has documented both informed and romanticized understandings of animal husbandry, farm size, feeding, milking, reproduction, and animal lives (Folsche & Fiebelkorn, 2024; Szczepanski et al., 2023). Studies of animal ethics further show that students' reasoning is shaped by embodied metaphors and tacit moral assumptions, not simply by the amount of factual knowledge they possess (Tramowsky et al., 2022). Stakeholders concerned with food sustainability and farm-animal welfare consequently argue for early, cross-curricular, practically grounded provision, while also noting teachers' uncertainty about controversy and disciplinary knowledge (Jones et al., 2025). A cheerful farming game can enter this field of prior ideas as a powerful but partial account of what farming is.

The educational-games literature offers no warrant for assuming that interactivity is intrinsically educative. Reviews of both purpose-built learning games and entertainment games used in curricula report heterogeneous outcomes and repeatedly return to the same conditions: curricular alignment, teacher mediation, task design, feedback, and opportunities to explain what happened during play (Clark et al., 2016; Cole et al., 2024; Ke, 2016; Martinez et al., 2022; Mayer, 2019; Plass et al., 2015; Wouters et al., 2013). Recent work on games and metacognition similarly stresses that reflective processes must be made visible and supported (Checa-Romero & Gimenez-Lozano, 2025). Game-literacy research goes further by treating games as cultural and designed texts whose rules, representations, platforms, and uses are open to analysis; systematic work on digital literacy games likewise identifies review and debriefing as important conditions for critical engagement (Bacalja, 2025; De La Hera et al., 2024; Pavloff-Pelkonen et al., 2026). Rules matter, but they do not determine meaning on their own: players interpret, appropriate, and sometimes resist what a game affords (Arjoranta, 2022; Sicart, 2011). Operational mastery therefore cannot be read as direct uptake of a game's worldview.

This distinction is especially relevant to geography. Digital environments organize space, agency, movement, scale, and consequence; they are not neutral containers for curricular content. Current geography-education scholarship calls for forms of geomedia use that combine reflection, argumentation, and participation rather than reducing digital competence to technical handling (Jekel et al., 2025). Research on immersive technologies reaches a related conclusion: technological novelty acquires educational value only through subject-specific framing and carefully designed tasks (Edler et al., 2025). Recent work on teachers' use of games in sustainability education also foregrounds mediation, adaptation, and the pedagogical meanings teachers construct around the game (Wahlström et al., 2026).

Agricultural serious games constitute a growing field, but their purposes and audiences differ markedly from the present case. A recent systematic review identified a

substantial corpus of games designed for learning, participation, research, or co-design in agriculture, while also noting weak and uneven evidence concerning real-world impact (Dernat et al., 2025). *Farmily*, for example, was purpose-built and evaluated as a persuasive serious game for novice home gardeners (García de Alba-Chávez et al., 2024). *Hay Day* is different. It is a commercial free-to-play game, not an instructional simulation. Its relevance lies precisely in the tension between the everyday attractiveness of its farm and the disciplinary incompleteness of its agricultural world.

The present study examines that tension in one fifth-grade geography class. It asks what triads articulated while operating *Hay Day*, what comparisons and evaluations appeared in a later explicitly prompted debriefing, and what design knowledge can be derived from the contrast between these two data production settings. The article uses a playable agricultural model as a domain-specific sensitizing concept for linking (a) relations that become actionable through a commercial game's rules and feedback with (b) agricultural relations, actors, spatialities, and consequences that remain non-actionable or absent. This framing does not treat *Hay Day* as a scientifically validated simulation or assume that rules determine player interpretation.

The evidential scope is deliberately bounded. Gameplay verbalizations and debriefing artifacts were produced under different tasks, modes, timings, social arrangements, and degrees of prompting. The study therefore addresses what discourse was produced under each elicitation condition and what design conjectures can be derived from their contrast; it does not estimate conceptual change, latent reflective capacity, or an autonomous game effect. The implications of these boundaries are consolidated in Section 5.5.

2. Conceptual Background

2.1. From Game-Based Learning to Critical Game Literacy

Game-based learning research has long examined how challenge, feedback, agency, and situated action can support learning (Clark et al., 2016; Ke, 2016; Mayer, 2019; Plass et al., 2015; Wouters et al., 2013). These affordances remain important, but they become analytically insufficient when a commercial game is brought into a classroom. The learning situation is then composed not only of player and game, but also of curricular aims, teacher instructions, peer arrangements, platform economies, classroom time, assessment practices, and learners' previous experiences. Bacalja's postdigital account is useful here because it resists both technological boosterism and simple rejection: claims about games have to be tested within the particular relations through which play is organized in school (Bacalja, 2025). This relational view also accords with the socio-didactical understanding of digital artifacts as constituents of the teaching situation rather than neutral additions to it (Rezat & Sträßer, 2012).

A critical game-literacy perspective accordingly asks more than whether students enjoyed or understood a game. It asks how rules and representations work together; which positions the player is invited to occupy; which actions are made efficient, desirable, or impossible; and how commercial and platform conditions enter the experience. At the same time, game meaning cannot be inferred mechanically from rules alone, because play is interpretive and situated (Arjoranta, 2022; Sicart, 2011). Recent reviews and educational design studies treat game and digital literacy as objects of explicit pedagogy rather than incidental benefits of exposure, with guided discussion and debriefing serving as means of moving from use toward critical interpretation (De La Hera et al., 2024; Pavloff-Pelkonen et al., 2026). For the present study, this means distinguishing three accomplishments without presuming a fixed sequence among them: operating the game, interpreting the agricultural representation, and evaluating the relation between the two.

2.2. Debriefing as Pedagogy and Elicitation

Debriefing is often described as the bridge between experience and learning (Crookall, 2010). The metaphor is helpful only if the two shores are specified. In this study, the first shore is the player's practical problem: What must we do now to continue? The second is a disciplinary problem: What kind of agriculture has the game made, and what has it left out? Moving between them requires a change in epistemic position. Students must temporarily suspend the game's definition of success and treat the system itself as evidence.

Reflection does not arise automatically from a pause in play. Reviews of reflection and metacognition in games point to the importance of prompts, dialogue, written externalization, and opportunities to revise an initial interpretation (Checa-Romero & Gimenez-Lozano, 2025; Villareale et al., 2020). Geography interventions using reflection diaries and peer coaching likewise show that carefully staged support can extend attention beyond immediate play (Baßeng & Budke, 2024, 2025). Yet elicitation is not neutral. Research comparing qualitative methods with children shows that different tasks generate different kinds and depths of response (Pollak et al., 2023), while think-aloud research emphasizes that task wording materially shapes the resulting protocols (Vanicek & Popelka, 2023). Debriefing must therefore be treated both as pedagogy and as a condition of data production.

2.3. Agricultural Literacy, Student Conceptions, and Ethical Absence

Agricultural literacy is not a list of farm facts. It involves understanding interconnected systems of production, processing, distribution, consumption, resource use, and societal decision-making (Cosby et al., 2022; Manning et al., 2024; Volk et al., 2025). Food literacy scholarship similarly distinguishes functional, relational, and critical competencies and locates children within complex food systems rather than treating knowledge as an isolated possession (Ares et al., 2024). For younger learners, these systems are interpreted through pre-existing images and narratives. Folsche and Fiebelkorn (2024) found that elementary students' conceptions of pig and dairy-cow husbandry varied in relation to urban and rural experience and included misconceptions about herd sizes, feeding, and housing. Related work on milk production found realistic ideas alongside pastoralized or incomplete accounts (Szczepanski et al., 2023).

Livestock farming also poses ethical questions that cannot be reduced to technical accuracy. Reasoning about animal welfare involves ideas about sentience, human benefit, responsibility, necessity, and acceptable harm (Tramowsky et al., 2022). The absence of suffering or slaughter from a game is therefore not a simple factual omission; it alters the moral architecture of the represented system. A game in which animals remain cheerful and productive, regardless of serious negative consequences, can make care appear as a frictionless input into production. Geography education can reopen that architecture by connecting farm spaces with actors, flows, environmental burdens, economic pressures, and competing values.

2.4. Playable Agricultural Model as a Sensitizing Concept

The term playable agricultural model is used here as a domain-specific sensitizing concept for a commercial game system that translates selected agricultural relations into interfaces, rules, resources, timings, feedback, and rewards. The concept builds on, rather than replaces, established accounts of games as simulations and procedural or rule-based representations (Bogost, 2007; Frasca, 2003). Its added analytical value is disciplinary: it requires the researcher to map what the player can enact onto an external agricultural system and to examine not only representational selection but also the spatial, material, institutional, ethical, and economic relations that are made non-actionable. In other words, procedural analysis asks how the game system works; the playable-agricultural-model

lens additionally asks what kind of agricultural system becomes actionable, what remains outside action, and what counterevidence geography education must introduce. This distinction also avoids treating procedural structure as a deterministic message, consistent with critiques of strong proceduralism (Sicart, 2011).

Hay Day, released by Supercell in 2012, is a free-to-play mobile farming game (Alha, 2020; Supercell, 2012). Its introductory sequence includes crop cultivation, feed production, hens, later cows, the collection of eggs and milk, orders, currencies, waiting times, and level-gated expansion (Alha, 2020; Google Play, n.d.; Klatt, 2024). The farm is colorful, orderly, and highly controllable. Animals are individualized and expressive, yet severe illness, death, and slaughter do not structure the early gameplay. The game is therefore useful in class not because it faithfully simulates farming, but because its selections are unusually visible once the player's goals are placed under examination.

For geography education, the concept therefore directs attention to four linked analytical moves: (1) identify represented spatial and material relations; (2) reconstruct the actions and feedback through which those relations become operational; (3) identify actors, scales, flows, externalities, and consequences that remain absent or non-actionable; and (4) design comparisons with disciplinary evidence that make those selections arguable. These moves connect critical game literacy with geography's system and structure–function–process concepts and with reflexive geomedia education (Fögele et al., 2021; Jekel et al., 2025; Morawski, 2023; Morawski & Wolff-Seidel, 2023; Rinschede & Siegmund, 2022). The research capability added by the concept is thus a traceable bridge from mechanic to agricultural relation to disciplinary counterevidence.

2.5. Research Questions

Three questions guided the analysis:

RQ1. What did the triads articulate while operating *Hay Day*, particularly in relation to livestock, production, exchange, and design constraints?

RQ2. Which game–reality comparisons, evaluations, and comments on the medium appeared in the explicitly prompted debriefing?

RQ3. What design conjectures can be drawn from the task-conditioned contrast between the two data-production settings, given their non-equivalence?

3. Materials and Methods

3.1. Design and Scope of the Case

The study employed an exploratory qualitative case-study design with a design-oriented purpose (Anderson & Shattuck, 2012; Sandoval, 2014; Yin, 2018). The case was bounded by one fifth-grade class, one project day, the introductory sequence of *Hay Day*, and a curricular focus on dairy cow and laying hen husbandry. The aim was analytical rather than causal: to reconstruct group-produced discourse under two different elicitation conditions and to use the observed patterns to refine a future instructional design.

The chronology is important for interpreting the analysis. The classroom intervention was conducted in 2025 as a supervised university student research project within the broader Gaming & Geography context. The classroom sequence, recording, transcription, and initial coding were completed within that source project. The present article is a later secondary reanalysis and theoretical reframing of the retained, anonymized material by the author team; it generated no new participant data. The current research questions and the playable agricultural model framing were formulated during article preparation after fieldwork and were not preregistered. Accordingly, the analysis is presented as an exploratory reanalysis rather than as a confirmatory test of hypotheses fixed before data collection.

Two forms of data generation were connected. During gameplay, seven triads produced socially organized concurrent verbalizations through rotating speaker-player roles. A later station-based debriefing elicited written group comparisons and evaluations. The phases are not equivalent measurements: task, timing, response mode, social organization, and degree of prompting changed together. Their frequencies are therefore neither compared as pre-/post-test scores nor interpreted as measures of stable individual capacity.

3.2. Participants and Curricular Context

The project took place on 8 July 2025 at a Gymnasium in Wuppertal, North Rhine-Westphalia, Germany. The class consisted of 24 students: 13 boys and 11 girls. All participants whose data were included in the research dataset were minors. Written informed consent from a parent or legal guardian was required for participation in the research component, together with the project's data-protection documentation. The parent/legal-guardian consent form covered participation in the research component, audio recording, research analysis, and anonymized scientific use/publication of the resulting materials. The required documentation was returned for 21 pupils; three girls without the required documentation worked in a separate instructional group, whose data were neither recorded nor analyzed. The corpus therefore represents 21 students (13 boys and 8 girls) working in seven triads.

The project was linked to the North Rhine-Westphalian geography curriculum on food-production chains—production, processing, transport, and trade—and to media-competence objectives concerning information evaluation, design features, identity, and self-regulated media use (Medienberatung NRW, 2020; Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2019). Planning had assumed that the class had already studied the relevant agriculture unit. The opening lesson showed that this assumption was incorrect. The limited disciplinary foundation is important for interpreting the students' comparisons and is addressed as a design limitation rather than being concealed as a minor implementation detail.

3.3. Game Selection and Classroom Exposure

Hay Day was selected because its opening levels make agricultural actions available within a short classroom period. At the shared starting point, students encountered fields, a feed mill, hens, and a small farm infrastructure; cows became available through progression. University-owned iPads were used because applications could not be installed on the school devices, and each device began at the same level to reduce variation in available actions.

Hay Day carried a German age label of 12 during the project (Klatt, 2024) and is a commercial free-to-play title with optional purchases and chance-related reward mechanisms (Google Play, n.d.; Klatt, 2024). USK age labels indicate the age from which a title is considered unobjectionable from a child-protection perspective; they are not pedagogical recommendations (Unterhaltungssoftware Selbstkontrolle, n.d.). The fifth-grade cohort included pupils for whom this label was therefore a safeguarding concern. The session was teacher/researcher supervised and time-limited on university devices, but the retained project documentation does not contain a separate approval protocol for use below the age label and does not preserve a technical audit of account status, purchase restrictions, network access, social functions, randomized rewards, or the exact game build. We therefore do not retrospectively claim that particular commercial functions were disabled or encountered. This documentation gap is treated as a study limitation and as a non-negotiable requirement for replication: future school use should employ an age-appropriate or explicitly institutionally approved configuration, record the software build, and disable or document purchases, social communication, and chance-based commercial features.

3.4. Teaching and Data-Generation Sequence

The opening disciplinary orientation drew on established accounts of livestock husbandry and hygiene, laying hen management, and dairy cow production systems (Deutsche Landwirtschafts-Gesellschaft, 2020; Hoy et al., 2016; Tergast et al., 2022). These sources functioned as instructional counterpoints to the game rather than as outcome measures.

The project day comprised four 45-minute lessons. Table 1 makes visible that instruction, method training, gameplay, and debriefing were separate stages and that only the latter two generated the focal research material. This sequencing is essential for interpreting later contrasts because the debriefing introduced a new task and new response supports rather than repeating the gameplay elicitation condition. Figure 1 summarizes the overall research design and analytic focus.

Table 1. Overview of the teaching and data generation sequence.

Lesson	Instructional Function	Main Activities	Research Material
1	Disciplinary orientation	Refresh selected knowledge of dairy farming and introduce the project.	Class discussion and worksheet
2	Method preparation	Model and practise triadic concurrent verbalization; assign speaker-player, technology, and note-taking roles.	Practice verbalizations and role cards
3	Supervised gameplay	Two planned 12 min rounds from a common starting level; a different speaker-player in each round.	Audio from 14 speaker-players embedded in seven triads; process notes
4	Prompted debriefing	Four station tasks on the game world, game–reality comparison, feelings/perceived learning, and game critique.	Written group responses; artifacts from three stations retained

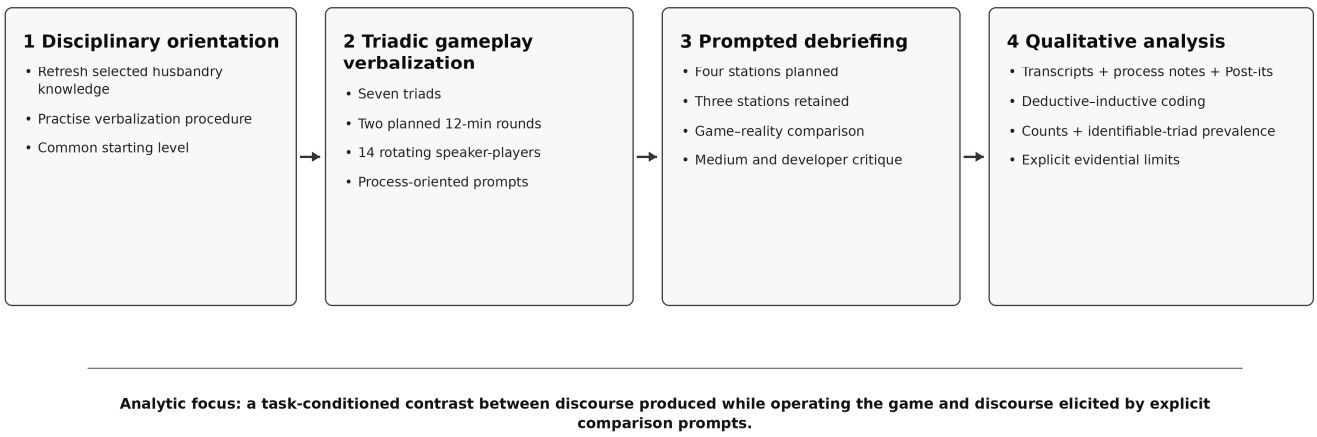


Figure 1. Research design and analytic focus. The figure visualizes the change from disciplinary orientation to triadic gameplay verbalization, prompted debriefing, and to qualitative analysis. Its central message is methodological: gameplay and debriefing are connected stages but not equivalent measures. Original artwork created for this article.

3.5. Data Corpus

The primary material comprised audio recordings generated by the seven triads during two planned 12 min gameplay rounds. Each triad had three concurrent roles: the speaker-player operated the game and was explicitly responsible for verbalizing what

was seen, intended, surprising, or used as a reason; the technology manager supported device handling and procedural continuity; and the process note-taker documented selected actions and observations. The speaker-player verbalized within the triad, not to an interviewer. The other two pupils remained interactional partners and could respond, question, suggest actions, or coordinate the task; such utterances were retained when audible. A different pupil became speaker-player in the second round. Consequently, 14 pupils received a guaranteed focal verbalization role, whereas the third pupil in each triad could contribute only through interaction and notes. The recordings are therefore analyzed as triadic discourse with unequal role-based opportunities to speak, not as equivalent individual think-aloud data for all 21 students. Some recordings were shorter or incomplete because not every group fully understood the procedure, and usable minutes per triad were not preserved.

The secondary material comprised process notes and written Post-it responses from the debriefing. Four stations had been planned, but artifacts from only three were retained. The stations explicitly asked students to describe the game world, compare it with livestock farming beyond the game, consider feelings/perceived learning, and comment on design choices. Crucially, the materials also supplied themes, examples, answer options, or sentence starters, including animal suffering and illness, slaughter, manure, limited space, strenuous work, large companies/citizens, and cheerful animal design. These supports were available but not compulsory. Because some Post-its lacked secure triad identifiers, the post-play material was analyzed as coded statements rather than as a complete seven-case dataset. The revised analysis treats prompt provenance as an evidential property rather than a minor limitation. Table 2 summarizes the scope, elicitation condition, and analytical role of each retained data source.

Table 2. Data corpus and analytical roles. The table distinguishes who or what generated each data source, how strongly it was elicited, and what kind of inference it can support. In particular, the debriefing artifacts are classified as prompt-responsive comparison/evaluation rather than spontaneous evidence of conceptual change.

Data Source	Scope	Elicitation Condition	Analytical Role
Triadic gameplay audio	Seven triads; two planned 12 min rounds; 14 rotating speaker-players	Process-oriented prompts; time-limited play; peer coordination	Immediate group-produced attention, decisions, and practical reasoning
Process notes	One designated note-taker per triad	Selective written record during play	Rough action sequence and support for interpreting unclear audio
Debriefing artifacts	Written group responses from three retained stations	Direct comparison and critique prompts; answer supports available	Prompt-responsive comparison, evaluation, and medium commentary
Instructional materials	Role cards, station sheets, disciplinary worksheet	Not outcome data	Documentation of how participation and response possibilities were structured

3.6. Analytical Procedure

In the 2025 source project, the audio recordings were transcribed in simplified form with f4 transcription software; the software version was not retained in the project records. The transcripts were checked against the recordings and anonymized (Döring, 2023). The source-project analysis followed structured qualitative content analysis (Hsieh & Shannon, 2005;

Kuckartz & Rädiker, 2024; Mayring, 2019). Deductive sensitizing categories were derived from a four-level framework for reflection on digital games: the game system, comparison with realities outside the game, self-reflection, and reflection on the medium (Lux & Budke, 2023). These categories functioned as analytical lenses rather than as a closed codebook. During close reading, inductive subcategories were developed to represent recurring forms of talk specific to the corpus.

The final coding architecture is shown in Table 3. For gameplay, inductively refined codes captured representations, perceived action consequences, action descriptions, exchange/distribution relations, self-reflective practical evaluations, and medium-related noticing of timers, waits, unlocks, and level thresholds. For debriefing, the coding captured understanding of the represented game world, game–reality comparison, and metareflection on strengths, limits, and developer choices. Passages could receive more than one code when they performed more than one analytical function. No formal segmentation rule had been recorded in the 2025 project; counts are therefore descriptive coded mentions rather than standardized units. To reduce false precision, gameplay findings also report the number of identifiable triads in which a category occurred where this could be reconstructed.

Table 3. Coding architecture linking research questions, deductive lenses, inductively refined code families, operational meanings, and primary evidence.

Research Question/Condition	Deductive Lens	Inductively Refined Code Family	Operational Meaning	Primary Evidence
RQ1/gameplay	Game system	Representation	What objects, animals, products, spaces, or interface elements are named or characterized.	Gameplay transcripts; process notes
RQ1/gameplay	Game system	Action and consequence	What is done, what must happen next, and which dependencies or consequences are articulated.	Gameplay transcripts
RQ1/gameplay	Game system/outside-game relation	Exchange and distribution	References linking production to customers, deliveries, currencies, orders, or farm growth.	Gameplay transcripts
RQ1/gameplay	Self-/medium reflection	Practical evaluation and medium-related noticing	Feelings, practical judgments, timers, waiting, unlocks, level thresholds, or other design constraints noticed during play.	Gameplay transcripts
RQ2/debriefing	Comparison with realities outside the game	Game–reality comparison	Comparisons concerning husbandry, production, actors, space, work, omissions, or consequences; interpreted with prompt provenance P1–P3.	Debriefing artifacts

Table 3. *Cont.*

Research Question/Condition	Deductive Lens	Inductively Refined Code Family	Operational Meaning	Primary Evidence
RQ2/debriefing	Self-reflection/reflection on medium	Perceived learning and metareflection	Comments on feelings/learning, strengths and limits, visual framing, constraints, or developer choices; interpreted with prompt provenance P1–P3.	Debriefing artifacts
RQ3/integration	Cross-condition comparison	Task-conditioned contrast and design conjecture	Comparison of what each elicitation condition made discussable, without treating phases as equivalent measures.	Joint display across gameplay and debriefing

The two-pass coding procedure described here refers to the 2025 source-project analysis and was conducted by one researcher. The first pass allocated relevant segments to the deductive main lenses; the second refined inductive subcategories, operational definitions, and anchor examples, and reapplied them across the corpus. Credibility and confirmability were strengthened through checking coded passages against the recordings, explicit code definitions and anchor examples, preservation of negative and contradictory cases, separation of descriptive counts from prevalence claims, and an audit trail linking research questions, data conditions, codes, and reported findings (Levitt et al., 2018; Tracy, 2010). These procedures improve transparency but do not substitute for independent coding; no interrater coefficient is reported.

3.7. Research Quality, Translation, and Project Chronology

Research quality was pursued through a bounded case description, explicit documentation of elicitation conditions, operational code definitions and anchor examples, descriptive rather than inferential use of counts, identifiable triad prevalence where reconstructable, and retention of contradictory cases (Hsieh & Shannon, 2005; Levitt et al., 2018; Tracy, 2010). The revised coding architecture makes the pathway from RQ to data condition to code family to result explicit. The comparison of gameplay talk and debriefing artifacts examines differences in discourse across settings; it does not constitute triangulated proof of a stable internal competence.

The project was simultaneously an instructional event and a research situation. Role cards, time limits, teacher instructions, researcher presence, peer interaction, and station prompts shaped what could be said. Research with children requires particular attention to how methods structure participation and voice (Pollak et al., 2023; Sun et al., 2023). Accordingly, the data are treated as situated, group-produced discourse rather than transparent reports of private beliefs. Quotations were translated from German by the authors, with wording kept close to the source while avoiding awkward literalism. The present article reanalyzes anonymized materials from the 2025 project; it does not claim that the current RQs or conceptual framing preceded fieldwork.

The distinction between original project work and current reanalysis is also relevant to the coding. The materials retained and available for article preparation comprised the gameplay audio recordings, anonymized transcripts, process notes, written debriefing artifacts retained from three stations, instructional materials, initial categories, and coded

summaries. During preparation of the present article, the author team re-examined these materials, clarified the deductive–inductive hierarchy, aligned categories with the current research questions, checked selected coded passages and anchor examples against the retained source material, and recalibrated claims to what the surviving documentation could support. Where the project record does not support reconstruction—for example, exact usable minutes, some group identifiers, formal segmentation rules, or statement-level prompt provenance—the article reports the uncertainty instead of inferring missing information.

No generative artificial intelligence tool was used during the 2025 classroom intervention, data collection, transcription, or original coding. During preparation of the revised manuscript, the authors used OpenAI ChatGPT (GPT-5.6 Sol) to assist with language editing, manuscript restructuring, and drafting and refining text in response to peer review. The tool was not used to generate or alter participant data, conduct transcription or coding, or produce empirical results. All AI-assisted text was reviewed, verified, and edited by the authors, who take full responsibility for the final content.

3.8. Ethics, Consent, and Safeguarding

The study involved minors in classroom-based educational research and originated from a supervised university student research project conducted in 2025. The University of Wuppertal Ethics Committee (EBUW) Guidelines for Applicants dated 13 February 2024 were in force before the project began. The Guidelines state that there is no obligation to submit an application to the EBUW and expressly include students conducting final theses or study-related research projects under the supervision of a member of the University of Wuppertal among those eligible to apply (University of Wuppertal Ethics Committee, 2024). Accordingly, a project-specific application to the EBUW was not a mandatory University requirement for this supervised student research project. This institutional rule is reported here as the policy applicable at the time; it is not presented as retrospective ethics approval or as a project-specific exemption. All participants whose data were recorded and analyzed were minors. Written informed consent was obtained from a parent or legal guardian for each of the 21 included pupils, together with the required data-protection documentation. The consent form used for the project was addressed to parents/legal guardians and required the signature of a parent/legal guardian; it covered the research component, audio recording, and anonymized scientific use/publication of the resulting material. Three pupils without the required documentation were taught in a separate, non-recorded instructional group and were not included in the analysis. The retained research materials used for the present article were anonymized and handled confidentially. The present article is a secondary reanalysis of the 2025 materials and involved no new participant contact or new data collection. A fully anonymized publication copy of the parent/legal-guardian consent form is provided as Supplementary File S3; signed originals are not reproduced.

Safeguarding is also relevant to the game itself. The class was in the fifth grade, while *Hay Day* carried an age label of 12. The activity was supervised, time-limited, and conducted on university-owned devices, but no separate below-rating approval protocol and no complete record of account, purchase, network, social-feature, or randomized-reward settings survives. The study therefore does not infer that these functions were inaccessible. For replication, the design principles are conditional on prospective institutional/school approval, age-appropriate software selection or documented justification, guardian/child information, and a recorded technical configuration that disables or controls commercial and social functions.

4. Results

4.1. Operating the Game Dominated the Gameplay Discourse

The triadic gameplay verbalizations were dominated by the practical work of keeping the session moving. Groups named objects, identified shortages, bought animals, produced feed, fed hens and cows, collected products, waited, and searched for the next available action. The animal-related coding yielded 41 mentions for cows and milk and 52 for hens and eggs. In both categories, action descriptions were the largest subcategory, with 29 coded mentions each. These mentions appeared in five identifiable triads for cows and six for hens. The pattern is not a measure of how many students held a conception; it characterizes the discourse produced while the triads were coordinating play.

The talk nevertheless contained causal statements about the game system. A group explained, "There is a feed mill on the farm so that the animals do not have to go hungry." Another group condensed the egg-production loop to, "Only with feed do you get eggs." Such statements articulated dependencies within the playable system: a resource had to be produced, an animal supplied, time allowed to pass, and a product collected. The evidence supports operational coherence in the talk, not a claim that agricultural understanding had been acquired. Table 4 provides the descriptive coding profile across gameplay and debriefing.

Table 4. Descriptive coding profile. Counts refer to coded mentions, not students; multiple coding was possible. "Identifiable triads" reports how many of the seven triads were explicitly traceable in the source results. Debriefing prevalence could not be reconstructed because some group identifiers were missing.

Phase	Main Category	Subcategory	Coded Mentions	Identifiable Triads
Gameplay	Cows and milk	Representation	7	3/7
Gameplay	Cows and milk	Perceived action consequences	5	4/7
Gameplay	Cows and milk	Action descriptions	29	5/7
Gameplay	Hens and eggs	Representation	12	5/7
Gameplay	Hens and eggs	Perceived action consequences	2	2/7
Gameplay	Hens and eggs	Action descriptions	29	6/7
Gameplay	Hens and eggs	Exchange and distribution relations	9	5/7
Gameplay	Self-reflection	Values, feelings, and practical evaluation	6	2/7
Gameplay	Medium-related noticing	Timers, waiting, unlocks, and level thresholds	26	5/7
Debriefing	Understanding the game world	Representation and consequences	5	Not reconstructable
Debriefing	Game–reality comparison	Production, husbandry, agency, and omissions	9	Not reconstructable
Debriefing	Metareflection	Strengths, limits, and developer choices	9	Not reconstructable

4.2. The Farm Appeared as a Production-and-Exchange Network

Nine coded mentions connected egg production to exchange or distribution. Students noticed customers, deliveries, orders, and the conversion of products into currencies.

One group reported, "A supplier came and took all the eggs. We received money and a diamond." The game thus extended attention beyond the animal enclosure toward a simplified chain linking production, sale, and farm growth.

4.3. Design Constraints Were Salient but Rarely Interpreted as Design

Twenty-six coded mentions concerned fixed production times, waiting, locked content, or level thresholds. Students repeatedly monitored how long eggs, milk, or feed would take and which possibilities were unavailable. These constraints were salient because they interrupted or directed action; they appeared in five identifiable triads.

What was almost entirely absent was a second interpretive step. Students did not ask why a timer rather than an agricultural process governed production, why severe negative outcomes had been removed, or how waiting, progression, and restricted availability related to pacing or monetization. The talk therefore registered the interface as a condition of play but rarely treated that condition as an authored representation. The 26 mentions are consequently interpreted as medium-related noticing, not as developed medium critique.

Explicit comparison with agriculture beyond the game was likewise scarce. One student wondered why the hens did not reproduce. This may indicate a mismatch between prior expectation and screen representation, but no explicit real-world comparison followed; it remains an analytically suggestive rather than conclusive instance.

4.4. Prompted Debriefing: Cued Themes and Student Elaboration

The written debriefing artifacts contained a different discursive register: five coded mentions concerned understanding the game world, nine compared the game with reality, and nine addressed strengths, limitations, or developer choices. These totals cannot be compared directly with gameplay frequencies because the latter task was shorter, written, collective, and explicitly content-focused. More importantly, the station materials themselves supplied many of the categories that appear in these responses. The revised analysis therefore separates thematic uptake from elaboration instead of treating the mere presence of a critical theme as evidence that students independently generated it.

Themes such as illness, animal suffering, slaughter, manure, limited space, demanding work, powerful companies/citizens, and the supplied ideas that animals should look cute and that the game should be fun were directly available in the debriefing materials. A response that selected, repeated, or minimally reformulated one of these supplied categories is therefore classified as P1 and demonstrates prompt uptake only. P2 is reserved for cases in which the retained response adds a relation, reason, consequence, or evaluation that is not itself supplied by the relevant prompt or response support. Because statement-level prompt provenance does not survive for every Post-it, a defensible corpus-wide P1/P2/P3 count cannot be reconstructed; the analysis therefore avoids pseudo-precision and assigns no spontaneous status to a theme merely because it appeared in student writing.

The same rule applies to agency. The statement "In the game the player decides; in reality the citizens decide" is classified conservatively as P1 because the task required a game–reality comparison and "citizens" was available in the station supports. A reference to large companies is treated in the same way where that actor category was supplied. These responses show that students used the scaffold to formulate a comparison, but the surviving record does not establish student-generated reasoning beyond the supplied categories. The article therefore does not use these examples as evidence of independent elaboration.

4.5. Prompt Provenance and Medium-Related Responses

Medium-related responses included observations that the game restricted how many barns or animals could be placed, and comments about why the animals were designed as they were. Because the debriefing explicitly invited design critique and supplied the

examples that the animals should look cute and that the game should be fun, responses reproducing that association are classified as P1 rather than P2. A stronger provenance classification is used only when the retained response demonstrably adds a relation not contained in the relevant support.

Several groups also reported that the activity had taught them about time and cost, equipment, cow feeding, or animal-appropriate husbandry. These are self-reports of perceived learning elicited by the debriefing, not independent evidence that the corresponding understanding was acquired or retained. The study contains no individual baseline or delayed assessment with which to validate those claims.

Some responses nevertheless moved beyond simple uptake of a supplied term. One group reasoned that livestock farming would cause few problems if farmers kept only as many animals as they could feed themselves. The station prompt itself supplied the herd-size comparison and examples of possible problems (including manure, limited space, animal suffering, and an unfinished 'too little . . .' cue), but it did not supply the conditional claim that problems would be avoided when herd size was matched to the farmer's own feed-production capacity. The response is therefore treated as P2, while remaining prompt-responsive rather than spontaneous. It also illustrates the limits of the elaboration: animal welfare and sustainability are reduced to one condition, while housing, health, breeding, labor, emissions, transport, slaughter, land use, and economic pressure disappear. Such cases are analytically useful because they show what students did with a scaffold, not because they demonstrate an unprompted misconception or learning gain.

Table 5 makes the conservative provenance decisions explicit by pairing selected responses with the relevant task/supports, the assigned provenance class, and the specific reasoning—if any—added by the students. Exact translated station wording is retained in Supplementary File S1; Supplementary File S2 provides the corresponding analytic-transparency record. The central interpretive rule is that scaffolds can make disciplinary comparison possible without converting scaffolded performance into evidence of spontaneous critical game literacy.

Table 5. Conservative prompt-provenance classification of selected debriefing evidence. Exact station wording is reproduced in Supplementary File S1.

Selected Response/Evidence	Relevant Task and Supplied Supports	Conservative Provenance Classification	Evidential Status
Mentions of illness/suffering, slaughter, manure, limited space, or demanding work	Game–reality/critique tasks; these theme families were supplied in the station materials	P1: thematic presence alone does not establish student-added reasoning	Prompt uptake only; not independent noticing
"In the game the player decides; in reality the citizens decide" and references to large companies	Game–reality comparison; citizens and large companies were available in the supports	P1: the comparison task and actor categories were supplied; no additional reasoning is claimed	Supported comparison under scaffolding
Responses stating that the developers made the animals this way because they should look cute/because the game should be fun	Design-critique prompt: "The developers made the animals this way because . . ." with supplied examples that they should look cute/the game should be fun.	P1: the response reproduces the supplied association	Prompt uptake/supported evaluation

Table 5. *Cont.*

Selected Response/Evidence	Relevant Task and Supplied Supports	Conservative Provenance Classification	Evidential Status
Reasoning that farming would cause few problems if farmers kept only as many animals as they could feed themselves	Game–reality prompt supplied the herd-size comparison and examples of possible problems (manure, limited space, animal suffering, and an unfinished "too little . . ." cue), but not the conditional self-sufficiency relation in the response.	P2: adds the conditional relation that problems would be avoided if herd size matched the farmer's own feed-production capacity	Prompt-responsive elaboration; not spontaneous conceptual change
P3 emergent reasoning	No complete statement-level prompt provenance survives for the full corpus	P3 is not assigned or counted corpus-wide	Potentially stronger evidence in principle, but not quantifiable here

4.6. Integrated Findings: A Task-Conditioned Contrast

Across the research questions, the evidence supports a task-conditioned contrast rather than a learning gain. During play, triads repeatedly articulated operational relations—dependencies, constraints, and progression. In the debriefing, critique-oriented language became more visible, but many of its central categories had been supplied by the materials. Thematic presence alone is therefore interpreted as prompt uptake; where the retained response demonstrably adds a relation beyond the relevant support, that elaboration provides stronger evidence of supported reasoning. Because task, response mode, timing, social organization, and prompting changed together, the study cannot attribute the contrast to play, debriefing, prior knowledge, peer coordination, or to any single mechanism. Table 6 integrates the empirical patterns with these evidential limits and the resulting design implications.

Table 6. Joint display of empirical patterns, evidential limits, and design implications.

Empirical Pattern	Evidence	Evidential Interpretation	Design Implication
Operational coordination dominated gameplay talk	Action descriptions occurred in five identifiable triads for cows and six for hens; causal dependencies were articulated.	The setting elicited coherent reasoning about operating the game, not evidence of agricultural learning.	Separate the player problem from the disciplinary problem and assess each with appropriate evidence.
The farm appeared as a simplified production-and-exchange network	Five identifiable triads linked eggs with customers, delivery, money, or diamonds.	A partial production chain became discussable, while labor, regulation, waste, and market power remained peripheral.	Add maps, material flows, actors, and counterevidence that extend beyond the game loop.
Constraints were salient without developed authorial critique	Timers, waiting, unlocks, and thresholds appeared in five identifiable triads.	Students registered constraints as conditions of action; they rarely interpreted why these had been designed.	Treat interfaces and commercial mechanics as evidence, but do not infer critical literacy from noticing alone.

Table 6. *Cont.*

Empirical Pattern	Evidence	Evidential Interpretation	Design Implication
Prompted debriefing elicited omitted dimensions	Written responses mentioned illness, slaughter, manure, work, space, agency, and visual design.	These are prompt-responsive accomplishments under a different task and mode, not spontaneous reflection or conceptual change.	Collect open responses before direct prompts and retain both as analytically distinct records.
The contrast is method-bound	Gameplay and debriefing differed in timing, mode, social organization, and prompting; one station was missing.	No latent "reflection gap" or causal game effect can be established.	Use equivalent individual measures, preserved exposure data, and delayed transfer tasks in future studies.

5. Discussion

5.1. Operational Coherence Is Not Agricultural Understanding

The gameplay discourse should not be read as evidence that the students lacked reflection. They continuously monitored rules, dependencies, shortages, and consequences, but they did so in relation to the practical problem the activity presented: how to continue. This is a form of situated reasoning. It becomes educationally limited only when operational coherence is mistaken for an adequate account of agriculture.

The distinction helps reconcile divergent positions in game-based learning. Reviews of educational and entertainment games report promise but also uneven evidence and strong dependence on context, curricular purpose, and instructional design (Clark et al., 2016; Cole et al., 2024; Ke, 2016; Martinez et al., 2022). A critical game-literacy perspective adds that rules do not transmit a single meaning; player interpretation and social use matter (Arjoranta, 2022; Sicart, 2011). In this case, operational discourse emerged from the interaction of game demands, process-oriented prompts, limited prior agricultural knowledge, time pressure, and triadic coordination—not from the game alone.

5.2. What the Playable Model Made Visible—And What It Made Easy to Forget

As a playable agricultural model, *Hay Day* offered a compact representation of production and exchange. Students could see that feed had to precede animal products, that products moved toward customers, and that sales supported further investment. These relations matter in agricultural literacy, which asks learners to connect, rather than merely name, parts of food systems (Cosby et al., 2022; Manning et al., 2024; Volk et al., 2025). The game's accessibility may therefore be pedagogically useful as an initial systems diagram enacted through play.

The same compactness produced the central risk. The model joined animal care to productivity without giving serious consideration to illness, death, slaughter, waste, or exploitative conditions. It individualized the farm as the player's domain and displaced collective institutions, labor, and unequal power. In ethical terms, the animals mattered because feeding them unlocked production; their vulnerability did not place an independent demand on the player. This is precisely where work on students' animal-ethics conceptions becomes relevant (Tramowsky et al., 2022). The representational problem is not simply that the game is inaccurate. It is that its omissions reorganize responsibility.

The later references to manure, space, work, illness, slaughter, agency, and visual framing show that the structured materials could support critique-oriented agricultural

comparisons. They do not show that the game independently "opened" these topics, because the debriefing itself supplied many of them. The theoretically relevant outcome is therefore conditional: when a selective playable model is juxtaposed with disciplinary counterevidence and explicit prompts, students can produce relations that are absent from ordinary gameplay discourse. Future designs must test whether those relations can also be generated openly and transferred beyond the prompted settings.

5.3. An Elicitation-Dependent Contrast, Not a Demonstrated Mechanism

In this case, the debriefing changed both the object of talk and the available vocabulary. Timers, cheerful animals, mortality, work, space, and concentrated agency became discussable partly because the activity explicitly asked students to suspend the player problem and examine the represented farm, and partly because the materials named several of these issues. This is pedagogically important but methodologically different from discovering a psychological mechanism. Figure 2 therefore represents an elicitation model for future testing, not an established sequence of cognitive change. More broadly, the form, directness, and timing of scaffolding should themselves be treated as design variables: open questions or images can first preserve opportunities for student-generated noticing, whereas more explicit statements or categories can subsequently support disciplinary comparison when the instructional goal requires it.

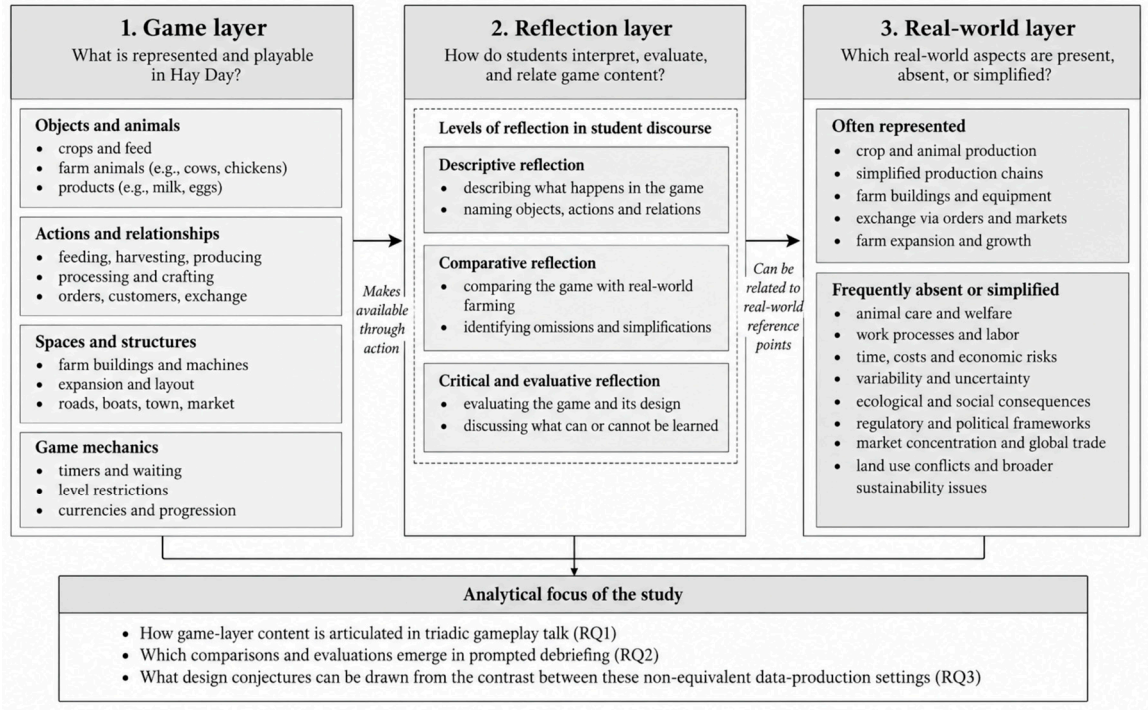


Figure 2. Heuristic account of the observed task-conditioned contrast and its evidential limits. The diagram represents a proposal for future testing, not a mechanism established by the present case; prompted-debriefing discourse is interpreted through the P1–P3 provenance distinction. Original artwork created for this article.

The provenance analysis also prevents critical-sounding phrases from being overinterpreted. P1 responses document uptake or minimal reformulation of provided language; P2 responses document an additional relation, reason, consequence, or evaluation under support; P3 would require a theme or line of reasoning not available in the relevant materials. The present records do not permit a reliable corpus-wide P3 estimate. A stronger design would therefore collect open individual noticing before directed prompts and retain exact prompt-to-response provenance.

5.4. Conceptual, Methodological, and Practical Contribution

The exploratory findings support six design conjectures rather than a universal recipe. Together they translate the study's conceptual and methodological contribution into testable relations among prior knowledge, open noticing, directed scaffolding, gameplay evidence, disciplinary counterevidence, represented/omitted actors, commercial mechanics, collective dialogue, and individual transfer.

Table 7 pairs each conjecture with an instructional implication and an observable indicator. The indicators are intentionally more demanding than student enjoyment or successful progression; they specify what future studies would need to observe to test whether supported critique transfers beyond *Hay Day* and beyond the immediate prompt. They may also serve as formative observation heuristics for teachers planning or reviewing game-based lessons. They are not validated assessment criteria, but they make the intended processes sufficiently explicit to support pedagogical observation and adaptation.

Table 7. Design conjectures, instructional implications, and observable indicators for future iterations.

Design Conjecture	Instructional Implication	Observable Indicator for Future Research
1. Build a disciplinary foothold before evaluation	Provide age-appropriate evidence on animal needs, husbandry systems, farm work, production chains, and spatial requirements.	Students use specific evidence—not only "realistic/unrealistic"—when evaluating a scene.
2. Mark the change of task	Explicitly distinguish the player problem ("How do we progress?") from the geography problem ("What does this farm represent?").	Students can formulate both questions and explain why they require different evidence.
3. Elicit openly before prompting directly	Collect unprompted observations and puzzles before introducing categories such as welfare, power, or environmental effects.	The record distinguishes spontaneous observations from prompt-responsive reasoning.
4. Juxtapose mechanics with counterevidence	Pair timers, currencies, unlocks, and production loops with farm data, maps, images, or testimonies.	Students explain how a mechanic selects, compresses, or omits a real relation.
5. Reintroduce distributed actors and consequences	Ask who works, decides, benefits, bears costs, and remains absent, including animals, workers, firms, regulators, and communities.	Students construct multi-actor accounts rather than attributing control to one benevolent player.
6. Assess individual transfer after collective dialogue	Follow group reflection with individual and delayed tasks using a different agricultural representation.	Students apply the distinction between playable system and represented system beyond *Hay Day*.

Practically, the findings refine several conditions emphasized in game-based-learning research. Curricular alignment requires a disciplinary foothold before critique; teacher mediation requires an explicit shift from the player problem to the geography problem; peer interaction can make operational reasoning and disagreement audible but does not provide equivalent individual evidence; feedback from the game should be juxtaposed with

agricultural counterevidence; reflection should proceed from open noticing to increasingly direct scaffolding. Agricultural knowledge and game literacy should therefore be connected rather than treated as separate add-ons. The study also suggests that prompt format and timing—questions, statements, images, or other comparison supports—should be selected in relation to the instructional purpose rather than treated as interchangeable forms of debriefing.

This distinction is especially important for a commercial title. Purpose-built agricultural serious games can align mechanics with explicit learning or participation goals, although their impact still requires evaluation (Dernat et al., 2025; García de Alba-Chávez et al., 2024). *Hay Day* was designed for entertainment and monetized play. Teachers therefore have to build the learning architecture around a system whose objectives lie elsewhere, making curricular framing, counterevidence, debriefing, and safeguarding constitutive rather than optional.

The age of the participants intensifies that responsibility. Fifth-grade students can contribute meaningful observations, but the study shows how strongly their participation is structured by roles, language, time, prompts, and technical conditions. Future implementations should prospectively document the software version, age-rating decision, consent and assent pathway, account configuration, purchase restrictions, network settings, and exposure to social or chance-based features, and should preserve individual as well as group-level evidence (Raneri et al., 2022; Sun et al., 2023).

5.5. Limitations and Research Agenda

The case is limited by its scale, interactional design, and documentation. Seven triads from one class cannot represent fifth-grade students generally.

The gameplay audio reflects 14 guaranteed speaker-player opportunities embedded in triadic interaction, not equivalent individual accounts from all 21 pupils. Several recordings were incomplete, the actual usable recording time per triad was not preserved, and one debriefing station was lost.

The absence of an individual baseline, comparison condition, or delayed measure does not prevent the study from addressing its descriptive and design-oriented research questions. Rather, these features delimit inference beyond those questions: the study can characterize situated discourse and derive testable design conjectures, but it cannot establish conceptual change, transfer, or an autonomous game effect. Likewise, because the debriefing constituted a different and more strongly scaffolded elicitation condition, the contrast with gameplay cannot be interpreted as a latent reflection gap.

The strongest measurement limitation within the study's stated scope is incomplete prompt provenance. The station materials supplied many themes later coded as critique, and the retained records do not permit complete statement-level P1/P2/P3 reconstruction. The original coding was conducted by one researcher without independent audit, formal segmentation rules were not preserved, some group identifiers were missing, and the planned prior agriculture unit had not actually been taught. Finally, the technical safeguarding configuration is incompletely documented. These conditions constrain reproducibility, generalization, and claims about independent student-generated reflection, even though they do not negate the descriptive and design-oriented contribution of RQ1–RQ3.

Future research should therefore prerecord the instructional and analytic plan where feasible; preserve exact exposure, prompts, screen/audio records, and technical settings; collect equivalent individual measures before and after play; elicit open observations before supplying categories; use independent or consensual qualitative coding; and include delayed transfer tasks with a second agricultural representation. Comparative designs

could then test whether commercial gameplay is more useful before disciplinary instruction as an elicitation device or after instruction as an object of model critique.

6. Conclusions

Hay Day made selected agricultural dependencies easy to act upon. During gameplay, triads articulated a coherent production loop involving feed, animals, products, customers, currencies, waiting, and expansion. This is evidence of operational reasoning within the playable system, not evidence that students acquired an adequate account of livestock farming.

The structured debriefing supported critique-oriented talk about illness, slaughter, manure, labor, space, agency, and visual framing, but many of these categories were directly supplied by the materials. Their mere appearance therefore demonstrates prompt uptake, not independent noticing. Where a retained response adds a relation not contained in the relevant support, it can be interpreted more strongly as prompt-responsive elaboration. The study nevertheless demonstrates neither conceptual change nor an autonomous game effect.

The playable agricultural model is best understood as a disciplinary sensitizing concept and research tool: it links what a commercial game makes actionable to the agricultural relations, actors, spatialities, and consequences that it renders absent or inconsequential. The case also offers a methodological lesson for game-based learning research: debriefing is simultaneously instruction and data production, so prompt provenance must be made visible. More broadly, educational use of commercial games should not be conceptualized as an interaction between learner and device alone; teacher mediation, peer dialogue, task framing, and disciplinary counterevidence jointly shape what becomes available for learning and critique. Testing the resulting design conjectures now requires equivalent individual measures, open-before-directed elicitation, preserved process data, independent coding, and prospectively documented ethics and safeguarding.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/educsci16101695/s1: Supplementary File S1 contains the English translations of the role cards and station prompts and the original coding materials; Supplementary File S2 provides the RQ–coding map, operational definitions, a conservative prompt-provenance evidence table, the chronology of source-project versus article-preparation procedures, and an analytic-transparency checklist; Supplementary File S3 provides a fully anonymized publication copy of the parent/legal-guardian consent form used for the minor participants, with all participant-specific entries and signatures removed. The original German classroom materials are available from the corresponding author subject to institutional and ethical conditions.

Author Contributions: Conceptualization, M.M. and D.P.; methodology, M.M. and D.P.; formal analysis, D.P. and M.M.; investigation, D.P.; data curation, D.P. and M.M.; writing—original draft preparation, M.M. and D.P.; writing—review and editing, M.M. and D.P.; visualization, M.M.; supervision, M.M.; project administration, M.M. and D.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: A project-specific application to the Ethics Committee of the University of Wuppertal (EBUW) was not mandatory under the institutional policy applicable at the time of the 2025 study. The official EBUW Guidelines for Applicants dated 13 February 2024 state that there is no obligation to submit an application to the EBUW and expressly cover supervised student final-thesis and study-related research projects. The study was conducted as supervised, non-clinical classroom-based educational research. Written informed consent from a parent or legal guardian and the required data-protection documentation were obtained for every minor participant whose data were included in the research dataset; pupils without the required documentation were not recorded or analyzed. The research materials used for the present analysis were anonymized. This

statement documents the institutional rule governing whether an EBUW application was obligatory; no retrospective ethics approval or project-specific exemption is claimed.

Informed Consent Statement: Written informed consent was obtained from the parents or legal guardians of all minor participants whose data were recorded and analyzed. The parent/legal-guardian consent form covered participation in the research component, audio recording, research analysis, and anonymized scientific use/publication of the resulting materials. The required consent and data-protection documentation had been returned for all 21 pupils included in the research dataset; three pupils without the required documentation participated in a separate, non-recorded instructional group and were not analyzed. A fully anonymized publication copy of the consent form is provided as Supplementary File S3; signed originals are retained confidentially and are not reproduced. No directly identifying participant information, images, or audio recordings are included in this publication.

Data Availability Statement: The qualitative data are not publicly available because they originate from minors in a small classroom setting and could permit indirect identification. De-identified excerpts, coding structures, and aggregated descriptive results are reported in the article and supplements. Further access can only be considered subject to institutional, legal, and ethical restrictions.

Acknowledgments: During the preparation of this manuscript, the authors used OpenAI ChatGPT (version GPT-5.6 Sol; OpenAI, San Francisco, CA, USA) for language editing, manuscript restructuring, and drafting and refining text in response to peer review. The authors reviewed, verified, and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. Supercell had no role in the design of the study; in the collection, analysis, or interpretation of the data; in the writing of the manuscript; or in the decision to submit the results.

References

Alha, K. (2020). *The rise of free-to-play: How the revenue model changed games and playing*. Tampere University.

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41*(1), 16–25. [CrossRef]

Ares, G., De Rosso, S., Mueller, C., Philippe, K., Pickard, A., Nicklaus, S., van Kleef, E., & Varela, P. (2024). Development of food literacy in children and adolescents: Implications for the design of strategies to promote healthier and more sustainable diets. *Nutrition Reviews*, *82*(4), 536–552. [CrossRef] [PubMed]

Arjoranta, J. (2022). How are games interpreted? Hermeneutics for game studies. *Game Studies*, *22*(3). Available online: https://gamestudies.org/2203/articles/arjoranta_how_are_games_interpreted (accessed on 15 July 2026).

Bacalja, A. (2025). Postdigital game-based learning: Complexity, continuity, and contingency. *Postdigital Science and Education*, *7*, 523–541. [CrossRef]

Baßeng, G., & Budke, A. (2024). Game on, reflection on: Reflection diaries as a tool for promoting reflection skills in geography lessons. *Education Sciences*, *14*(3), 316. [CrossRef]

Baßeng, G., & Budke, A. (2025). Peer coaching to reflect on digital games in geography lessons during the debriefing. *Education Sciences*, *15*(4), 445. [CrossRef]

Bogost, I. (2007). *Persuasive games: The expressive power of videogames*. MIT Press. [CrossRef]

Checa-Romero, M., & Gimenez-Lozano, J. M. (2025). Video games and metacognition in the classroom for the development of 21st century skills: A systematic review. *Frontiers in Education*, *9*, 1485098. [CrossRef]

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. [CrossRef] [PubMed]

Cole, C., Parada, R. H., & Mackenzie, E. (2024). A scoping review of video games and learning in secondary classrooms. *Journal of Research on Technology in Education*, *56*, 544–577. [CrossRef]

Cosby, A., Manning, J., Power, D., & Harreveld, B. (2022). New decade, same concerns: A systematic review of agricultural literacy of school students. *Education Sciences*, *12*(4), 235. [CrossRef]

Crookall, D. (2010). Serious games, debriefing, and simulation/gaming as a discipline. *Simulation & Gaming*, *41*(6), 898–920. [CrossRef]

De La Hera, T., Cañete Sanz, L., Navarro Sierra, N., Jansz, J., Kneer, J., Glas, R., & van Vught, J. (2024). Digital literacy games: A systematic literature review. *Frontiers in Communication*, *9*, 1407532. [CrossRef]

Dernat, S., Grillot, M., Andreotti, F., & Martel, G. (2025). A sustainable game changer? Systematic review of serious games used for agriculture and research agenda. *Agricultural Systems*, *222*, 104178. [CrossRef]

Deutsche Landwirtschafts-Gesellschaft. (2020). *DLG-Merkblatt 405: Legehennen* (3rd ed.). Deutsche Landwirtschafts-Gesellschaft. Available online: https://www.dlg.org/mediacenter/dlg-merkblaetter/dlg-merkblatt-405-legehennenhaltung (accessed on 15 July 2026).

Döring, N. (2023). *Forschungsmethoden und evaluation in den sozial- und humanwissenschaften* (6th ed.). Springer.

Edler, D., Wiktorin, D., Becker, G., Schweins, U., Reichert, K., Keil, J., & Dickmann, F. (2025). The potential of immersive virtual reality for teaching geography in schools. *KN—Journal of Cartography and Geographic Information*, *75*, 205–217. [CrossRef]

Folsche, E., & Fiebelkorn, F. (2024). Students' conceptions of keeping fattening pigs and dairy cows: An exploratory interview study with elementary school students in North-West Germany. *Journal of Biological Education*, *58*, 721–751. [CrossRef]

Fögele, J., Sesemann, O., & Westphal, N. (2021). *Mit Basiskonzepten die fachliche Tiefenstruktur des Geographieunterrichts gestalten*. Klett. Available online: https://hilpub.uni-hildesheim.de/handle/ubhi/10903 (accessed on 15 July 2026).

Frasca, G. (2003). Simulation versus narrative: Introduction to ludology. In M. J. P. Wolf, & B. Perron (Eds.), *The video game theory reader* (pp. 221–236). Routledge.

García de Alba-Chávez, C. A., Espinosa-Curiel, I. E., & Michel-Nava, R. M. (2024). Exploring the impact of a persuasive serious video game (Farmily) on promoting home gardening among novices: Design and randomized controlled trial. *JMIR Serious Games*, *12*, e60771. [CrossRef] [PubMed]

Google Play. (n.d.). *Hay Day*. Available online: https://play.google.com/store/apps/details?id=com.supercell.hayday (accessed on 19 August 2025).

Hoy, S., Gauly, M., & Krieter, J. (2016). *Nutztierhaltung und -hygiene* (2nd ed.). Eugen Ulmer.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. [CrossRef] [PubMed]

Jekel, T., Gryl, I., Lauffenburger, M., Kanwischer, D., Budke, A., & Schulze, U. (2025). Reflection, argumentation, and participation through geomedia: A model of emancipatory use for teacher training. *Journal of Geography*, *124*(1), 12–21. [CrossRef]

Jones, A., McCulloch, S. P., & Reiss, M. J. (2025). Stakeholder perceptions of primary school education about food sustainability and farm animal welfare in England. *Research in Science & Technological Education*, *43*, 718–737. [CrossRef]

Ke, F. (2016). Designing and integrating purposeful learning in game play: A systematic review. *Educational Technology Research and Development*, *64*, 219–244. [CrossRef]

Klatt, N. (2024). *Spielbeurteilung Hay Day: Freundliche Farmsimulation für Groß und Klein*. Spieleratgeber-NRW. Available online: https://spieleratgeber-nrw.de/spiel/hay-day/ (accessed on 15 July 2026).

Kuckartz, U., & Rädiker, S. (2024). *Qualitative Inhaltsanalyse: Methoden, Praxis, Umsetzung mit Software und künstlicher Intelligenz* (6th ed.). Beltz Juventa.

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology. *American Psychologist*, *73*(1), 26–46. [CrossRef] [PubMed]

Lux, J.-D., & Budke, A. (2020). Alles nur ein Spiel? Geographisches Fachwissen zu aktuellen gesellschaftlichen Herausforderungen in digitalen Spielen. *GW-Unterricht*, *160*, 22–36. [CrossRef]

Lux, J.-D., & Budke, A. (2023). Reflexives Spielen? Wie junge Spielende Repräsentationen gesellschaftlicher Themen in digitalen Spielen reflektieren. *MedienPädagogik*, *2023*, 188–211. [CrossRef]

Manning, J., Cosby, A., McDonald, N., & Fogarty, E. (2024). Primary and secondary school students' knowledge and perceptions of agriculture. *Journal of Agricultural Education*, *65*(2), 226–240. [CrossRef]

Martinez, L., Gimenes, M., & Lambert, E. (2022). Entertainment video games for academic learning: A systematic review. *Journal of Educational Computing Research*, *60*(5), 1083–1109. [CrossRef]

Mayer, R. E. (2019). *Computer games for learning: An evidence-based approach*. MIT Press.

Mayring, P. (2019). Qualitative content analysis: Demarcation, varieties, developments. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *20*(3), 16. [CrossRef]

Medienberatung NRW. (2020). *Medienkompetenzrahmen NRW*. Medienberatung NRW.

Medienpädagogischer Forschungsverbund Südwest. (2025). *KIM-Studie 2024: Kindheit, internet, medien* (2nd ed.). MPFS. Available online: https://mpfs.de/app/uploads/2025/05/KIM-Studie-2024.pdf (accessed on 15 July 2026).

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2019). *Kernlehrplan für die Sekundarstufe I Gymnasium in Nordrhein-Westfalen: Erdkunde*. MSB NRW.

Morawski, M. (2023). Reflexive Kompetenzen in digitalen Räumen von Videospielen—Auszüge einer quantitativen Studie über die Selbstreflexion von Schüler*innen in (geographisch-schulischen) Gamingkontexten. *GW-Unterricht*, *172*, 22–41. [CrossRef]

Morawski, M., & Wolff-Seidel, S. (2023). Gaming & geography (education): A model of reflexive analysis of space & action in video games. *European Journal of Geography*, *14*(3), 1–19. [CrossRef]

Pavloff-Pelkonen, E. M., Hakkarainen, K., & Korhonen, T. (2026). Educational design and implementation of game literacy in a multiliteracy learning environment. *Journal of Research on Technology in Education*, *58*(5), 1028–1046. [CrossRef]

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258–283. [CrossRef]

Pollak, I., Stiehl, K. A. M., Schrank, B., Birchwood, J., Krammer, I., Mitic, M., Rogers, J. C., & Woodcock, K. A. (2023). Engaging children in intervention development—A comparison of four qualitative methods and their suitability to elicit information relevant for early intervention development. *International Journal of Qualitative Methods*, *22*, 16094069231204775. [CrossRef]

Raneri, P. C., Montag, C., Rozgonjuk, D., Satel, J., & Pontes, H. M. (2022). The role of microtransactions in Internet Gaming Disorder and Gambling Disorder: A preregistered systematic review. *Addictive Behaviors Reports*, *15*, 100415. [CrossRef] [PubMed]

Rezat, S., & Sträßer, R. (2012). From the didactical triangle to the socio-didactical tetrahedron: Artifacts as fundamental constituents of the didactical situation. *ZDM Mathematics Education*, *44*, 641–651. [CrossRef]

Rinschede, G., & Siegmund, A. (2022). *Geographiedidaktik* (5th ed.). Brill Schöningh.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*(1), 18–36. [CrossRef]

Sicart, M. (2011). Against procedurality. *Game Studies*, *11*(3). Available online: https://gamestudies.org/1103/articles/sicart_ap (accessed on 15 July 2026).

Sun, Y., Blewitt, C., Edwards, S., Fraser, A., Newman, S., Cornelius, J., & Skouteris, H. (2023). Methods and ethics in qualitative research exploring young children's voice: A systematic review. *International Journal of Qualitative Methods*, *22*, 16094069231152449. [CrossRef]

Supercell. (2012). *Hay Day* [Mobile video game]. Supercell.

Szczepanski, L., Fiebelkorn, F., Ostermann, G., Altevogt, L., & Folsche, E. (2023). Student conceptions of the production of cow's milk—An exploratory interview study with 6th- and 10th-grade students. *Frontiers in Nutrition*, *10*, 1112183. [CrossRef] [PubMed]

Tergast, H., Hansen, H., & Weber, E.-C. (2022). *Steckbriefe zur Tierhaltung in Deutschland: Milchkühe*. Thünen-Institut für Betriebswirtschaft. Available online: https://literatur.thuenen.de/digbib_extern/dn065685.pdf (accessed on 15 July 2026).

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837–851. [CrossRef]

Tramowsky, N., Messig, D., & Groß, J. (2022). Students' conceptions about animal ethics: The benefit of moral metaphors for fostering decision-making competence. *International Journal of Science Education*, *44*, 355–378. [CrossRef]

University of Wuppertal Ethics Committee. (2024, February 13). *Evaluation of research projects by the Ethics Committee of the University of Wuppertal: Guidelines for applicants*. Available online: https://www.uni-wuppertal.de/fileadmin/data/forschung/PDFs/Handout_Ethics_Commitee_240213_2.pdf (accessed on 15 July 2026).

Unterhaltungssoftware Selbstkontrolle. (n.d.). *Die USK-Alterskennzeichen*. Available online: https://usk.de/die-usk-alterskennzeichen/ (accessed on 15 July 2026).

Vanicek, T., & Popelka, S. (2023). The think-aloud method for evaluating the usability of a regional atlas. *ISPRS International Journal of Geo-Information*, *12*(3), 95. [CrossRef]

Villareale, J., Biemer, C. F., Seif El-Nasr, M., & Zhu, J. (2020, November 2–4). *Reflection in game-based learning: A survey of programming games* [Conference paper]. Annual Symposium on Computer-Human Interaction in Play (pp. 672–678), Virtual Event, Canada. [CrossRef]

Volk, N., Rumble, J. N., & Alabi, S. T. (2025). Exploring the impact of EFTs on students' understanding of systems-based agricultural processes. *Journal of Agricultural Education*, *66*(2), 20. [CrossRef]

Wahlström, T., Windsor, S., & Svensson, M. (2026). Teachers' understandings of using a game in sustainability education—A case study from Sweden. *Education Sciences*, *16*(6), 975. [CrossRef]

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. [CrossRef]

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.