Documentation in Play: A Framework for Collaborative Evidence Construction in Inquiry-Based Games

MICHELE NEWMAN, University of Washington, USA
RUNHUA ZHAO, University of Washington, USA
STEVEN J HU, University of Washington, USA
ILENA B DALLA GASPERINA, University of Washington, USA
NISHA DEVASIA, University of Washington, USA
TRAVIS WINDLEHARTH, foundry10, USA
JENNIFER RUBIN, foundry10, USA
JIN HA LEE, University of Washington, USA
JASON C YIP, University of Washington, USA

Fig. 1. Bomi (left) and Walter (right) look at the player, asking if they will believe or not believe the information they found on the abandoned Star Corp ship.

Authors' Contact Information: Michele Newman, Information School, University of Washington, Seattle, Washington, USA, mmn13@uw.edu; Runhua Zhao, Information School, University of Washington, Seattle, Washington, USA, runhz@uw.edu; Steven J Hu, Information School, University of Washington, Seattle, Washington, USA, steven.j.hu@outlook.com; Ilena B Dalla Gasperina, Human Centered Design and Engineering, University of Washington, Seattle, Washington, USA, ilenadg@uw.edu; Nisha Devasia, Human Centered Design and Engineering, University of Washington, Seattle, Washington, USA, ndevasia@uw.edu; Travis Windleharth, foundry10, Seattle, Washington, USA, travis@foundry10.org; Jennifer Rubin, foundry10, Seattle, Washington, USA, jen@foundry10.org; Jin Ha Lee, Information School, University of Washington, Seattle, Washington, USA, jinhalee@uw.edu; Jason C Yip, Information School, University of Washington, Seattle, Washington, USA, jcyip@uw.edu.

This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2026 Copyright held by the owner/author(s).
ACM 2573-0142/2026/11-ARTGAMES017
https://doi.org/10.1145/3831324

Games are sites where players construct meaning through their interactions with the game system and with one another. Yet, while existing work describes why players make meaning, less is known about the mechanisms by which players collectively decide what counts as evidence during play. We present a 2.5-year Research through Design study, conducted through co-design with youth, that investigates these evidentiary meaning-making processes via the design of a misinformation-based game. Drawing on formal game analysis and qualitative analysis of design documentation and co-design sessions, we develop the Documentation-in-Play (DiP) Framework, presenting four mechanisms researchers and designers can use to observe and design for how players produce and stabilize evidence in collaborative play. We also present design heuristics for scaffolding these mechanisms, organized across cognitive, social, material, and time-based dimensions. While our design probe centers a misinformation game for youth, DiP is positioned as a transferable lens for any inquiry-oriented game context where interpretation, validation, and inscription are central to play.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; *Collaborative and social computing*.

Additional Key Words and Phrases: Play, Meaning-making in games, Misinformation, Collective Sensemaking, Game-based learning, Research through design, Participatory design

ACM Reference Format:
Michele Newman, Runhua Zhao, Steven J Hu, Ilena B Dalla Gasperina, Nisha Devasia, Travis Windleharth, Jennifer Rubin, Jin Ha Lee, and Jason C Yip. 2026. Documentation in Play: A Framework for Collaborative Evidence Construction in Inquiry-Based Games. *Proc. ACM Hum.-Comput. Interact.* 10, 7, Article GAMES017 (November 2026), 32 pages. https://doi.org/10.1145/3831324

1 Introduction

Games are sites of active meaning-making [71]. As players encounter rules, feedback, and narrative, they interpret and revise their understanding as play unfolds [73, 78, 85]. For example, a player might read a sudden change in a game's music as a cue that danger is near, using it to adjust their strategy and deepen their grasp of the game world. Game designers deliberately shape this process [79] through feedback loops [74, 80] and by embedding emotional and symbolic cues [5, 29, 35, 53]. Furthermore, these elements help to focus players' attention, prompting them to interpret, construct, and share knowledge about the game together [6, 31, 41, 82]. In this sense, how a game is designed determines not just what players do, but how they come to know.

While all games require some level of interpretation, some games push this concept further by elevating meaning-making to a core challenge. To illustrate this idea, consider social deduction games like *Among Us* [G6], where there is no single obvious truth. Players reconstruct events from fragments gathered during play, recalling who was seen where or flagging inconsistencies in alibis. They then argue for their interpretation against competing accounts. Progress depends not on individual reasoning in isolation, but on the group's ability to collaboratively build, debate, and stabilize a shared understanding of what happened. These games make collaborative sensemaking [88] the essence of play, not simply a byproduct.

Moreover, skills like reasoning under uncertainty, evaluating evidence, and negotiating shared understanding are central to both this style of play and civic life [38, 42, 87, 95]. Recognizing these parallels, designers have created misinformation games to "*inoculate*" [68] players against misleading information by immersing them in ambiguous environments that require critical evaluation [4, 22, 67]. In these games, players are indeed asked to consider the validity of misinformation; however, they often do so through having players practice discrete reasoning skills, resulting in a more solitary, linear play structure [17, 21, 22, 30, 48].

In turn, these games model misinformation as a cognitive concern [47]. This distinction matters because misinformation is not a problem that can be addressed individually. It is deeply social,

emotional, and cultural [52, 76, 87, 88, 95]. As a result, games focused only on individualist and cognitivist approaches address just part of the problem [87].

Newer misinformation games like *Werewolf* [1], *Among Us* [89], and escape rooms [10, 14, 18, 65] have introduced collaboration, but typically focus on social manipulation or puzzle-solving, not on structured, collective evidence-building [75, 76, 87, 88]. In *Werewolf*, for instance, players debate and accuse, but the main goal is manipulation rather than building a shared account of events. The game lacks ways to build, revisit, or contest evidence over time. Even in *Among Us* [89], the focus is on finding the imposter, not reflecting on how information is processed and shared. As a result, even collaborative misinformation games often lack mechanics for deeper investigation or collective meaning-making [17, 21, 30, 48]. Thus, collaborative sensemaking practices like creating group frames [75] are often unsupported, forcing players to invent or intuit them, if at all.

Taken together, these challenges reveal a gap in current game design: *few misinformation games are structured to support the collective production, documentation, and circulation of knowledge.* Addressing this limitation means designing experiences where players actively build, evaluate, and share evidence as a group to learn the skills that are fundamental to digital literacy and civic participation [25, 42, 52]. When games are intentionally structured in this way, they can help players develop the abilities needed to navigate complex information environments [87].

Yet, if collaborative sensemaking is to be genuinely supported, we must attend to how group members decide what information is valuable and how it is shared, challenged, and stabilized [69, 75]. A central challenge in collaborative sensemaking is not just interpreting information, but producing and negotiating what counts as evidence within a group. To analyze this process in play and game design, we draw on library and information science and introduce **documentation** as our analytical lens. In everyday use, documentation means writing something down. In this context, it is the process by which a group decides what is worth keeping, what counts as evidence, and what can be trusted [8, 12, 27, 55]. Crucially, documentation does not simply capture information that already exists. *It produces it* [27]. A musical cue is just a sound until a player marks it as significant [8, 55]. Once documented, these moments become stable, shareable, and arguable.

In this way, we believe this approach offers a complement to other approaches of meaning-making common in game research. Meaning-making often describes how players interpret a game world when they reflect on it [46, 58], but interpretation alone can remain private, belonging to a single player. Documentation asks what happens when players must make that interpretation visible and convincing to others. It lets us study not just how players think, but how thinking becomes evidence [8, 16, 27, 51, 55].

To explore these ideas, we designed a study to observe how documentation and the collaborative construction of evidence in play unfold in practice. We adopted a Research through Design (RtD) [97] approach and a co-design process with youth [24]. Over 2.5 years, we developed and studied *Starbound Secrets: Down the Rabbit Hole*, a hybrid digital-analog misinformation game in which players must turn evolving, ambiguous information into evidence to solve a mystery. This game deliberately models the experience of falling down a misinformation rabbit hole [64], an analogy that was derived via work with our youth co-designers. Co-designing with youth allowed us to embed evidence construction directly into the game's mechanics and observe in real time how players negotiated meaning [50] and were especially valuable, as their developing interpretive strategies made the work of collaborative evidence construction particularly visible [42].

Explicitly stated, our goal is not to present the game as an intervention for misinformation literacy; instead, it is to use the game design experience and final game artifact as a probe to address the following research questions:

(1) How do youth use the game's goals, actions, and components to construct, materialize, and revise interpretations as part of their evidence work during collaborative play?
(2) What cognitive, social, and material functions do these practices serve in enabling players to collectively establish what counts as evidence during play?
(3) How can these functions inform the design of games that scaffold collaborative evidence construction, particularly in games that model information-rich contexts such as misinformation?

To address these research questions, we make three key contributions:

- **A Theoretical Lens (RQ1):** We introduce the Documentation-in-Play (DiP) Framework (Section 5.1) as a conceptual lens for understanding how players construct evidence during collaborative play. The framework identifies four recurring mechanisms observed through a multi-year Research through Design process informed by co-design activities, iterative game development, and playtesting. Grounded in document theory [8, 20, 27, 55], DiP connects these mechanisms to cognitive, social, and material dimensions of documentation, offering a structured way to analyze how evidentiary practices emerge during play, seeking to give researchers and designers mechanisms to observe that arise between player's and the game's system.
- **An Empirical Case Study (RQ2):** We present a 2.5-year Research through Design case study of *Starboud Secrets*, a misinformation game, showing how evidence construction practices emerged and were iteratively shaped through design, co-design, and playtesting.
- **Game Design Heuristics (RQ3):** We distill our findings into 12 design heuristics for supporting evidence construction in collaborative, inquiry-based games, grounded in Lankoski and Björk [50]'s formal aspects of gameplay (Section 5.2).

While our case study centers on a misinformation game designed for youth, we believe the DiP Framework is not limited to this context. In this vein, rather than claiming a universal model of collaborative play, we position DiP as a conceptual lens that is grounded in empirical work and document theory [8, 27, 55].

Specifically, the DiP framework is derived to help provide a way to connect player documentation practices to specific design features and interaction structures in games where interpretation, validation, and inscription are central to play. We propose DiP as a vocabulary for analyzing and designing for situations in which players must construct shared understanding from ambiguous information [69, 71]. By naming and formalizing these dynamics, we offer designers and researchers a structured way to reason about evidence construction as an intentional design feature of play rather than an incidental outcome.

2 Related Work

2.1 Meaning-Making in Games

Our work builds on a long tradition in game studies examining how players derive meaning from systems, narratives, and experiences [6, 44, 70, 80]. These approaches treat gameplay as a process of meaning-making [71, 85], where players interpret mechanics, narrative, and play dynamics [15, 26].

However, most research privileges either cognitive or social perspectives, often treating them separately. Cognitive approaches focus on how players' mental models align with game elements [7, 63, 91], highlighting interpretation, flow, and alignment with personal goals [29, 53]. Yet, this can reduce play to decoding systems, rather than constructing knowledge through active engagement [34]. Recent work connects design choices to meaning-making by exploring how mechanics such as difficulty or digital collections shape interpretation and player identity [2, 81]. Social approaches highlight how players negotiate, share information, and build trust within multiplayer environments

[41]. These practices support psychological needs like autonomy and relatedness [3, 79, 83, 84], showing that play is fundamentally social as well.

Yet, these cognitive and social accounts offer less clarity on how meaning becomes stabilized, shared, and contested within a group [16]. While they explain how players interpret and negotiate meaning, they provide less insight into how traces of play become durable evidence that can be referenced across time and participants. Understanding this process is important for explaining how groups construct shared interpretations during play [11, 82].

2.2 Document Theory and Evidence Construction

Therefore, our approach draws on **document theory** to address this question. Document theory, which examines how meaning is constructed, stabilized, and preserved through documents, broadens the definition of a document beyond simple text on paper. Originating in library and information science [12, 27], document theory defines **documentation** as the interplay of cognitive, social, and material dimensions [55]. Material artifacts, such as notes, screenshots, or shared documents, play an important role in this process by helping make knowledge durable and communicable. Rather than treating documents as passive records, document theory emphasizes how artifacts become evidence that can support and explain meaning [8, 55]. This perspective provides a way to understand how artifacts created during play anchor meaning and transform fleeting experiences into shared evidence [11, 12, 20, 55, 72]. Unlike approaches that primarily examine how artifacts support cognition [37, 39, 40, 66], document theory focuses on how artifacts become persistent records that can be referenced, shared, and challenged. In doing so, it helps explain how meaning becomes stabilized within groups and how moments of play become enduring reference points for discussion, memory, and debate [12, 55].

To clarify what we mean by evidence, we follow Suzanne Briet, who argued that even an antelope in a zoo becomes a document when it is framed and categorized to produce knowledge [8]. More broadly, document theorists argue anything that transforms an object, event, or trace into a material representation can function as a document [12]. We therefore define **evidence construction** as the process of taking traces from play and using them to represent, justify, or support meaning. In this view, meaning-making is not simply interpretation but the active selection, creation, and negotiation of evidence. This perspective is particularly useful in contexts where players must collaboratively evaluate, justify, and contest information. One such context is misinformation, where meaning emerges through the interpretation, sharing, and negotiation of evidence [38, 52, 76, 95].

2.3 Justification for the Case: Misinformation Games

When dealing with misinformation, individuals must evaluate claims, negotiate credibility, and construct evidence for what they believe [76, 88]. Paired with the increasing goal of creating misinformation games [87], these challenges make misinformation a useful domain for examining evidence construction in play. In addition, this framing aligns with evidence that youth and adults rely on analytical reasoning, prior knowledge, emotions, and social context to make sense of misinformation [11, 13, 43, 52, 59, 87, 95].

As Konstantinou and Karapanos [49] note, simply presenting corrective information is rarely sufficient. Instead, collaborative reasoning prompts critical thinking and justification [32]. In practice, youth interpret claims through cultural frames, negotiate credibility socially, and externalize judgments [14, 57, 75, 76]. Misinformation, therefore, is as much a social and cultural issue as a cognitive one [95]. This complexity is precisely what makes misinformation an important and challenging problem for our game to model. It allows us to explore how mental models, social negotiation, and material practices interact in evidence construction.

Despite this, while recent games examine misinformation, most focus on cognitive evaluation. Most seem to be single-player, quiz-based, and brief [17, 21, 21, 22, 30, 48]. In response, newer games have begun integrating approaches that foreground the social, embodied, and collaborative aspects of sensemaking in misinformation environments. As noted in the introduction, this includes adaptations of *Among Us* [89], *Werewolf* [1], and escape rooms [10, 14, 18, 65]. Others use narrative to support embodied sensemaking [19, 22, 23]. However, most still fall short of capturing the complexity of real information ecologies, making this a productive domain for studying how design shapes meaning-making.

3 Methodology

3.1 Methodological Rationale

3.1.1 Rational for Research Through Design. This study is guided by the methodological paradigm of Research through Design (RtD) [97], which emphasizes generating knowledge through iterative design. In our work, co-designing a digital-analog misinformation game was both the context for engagement and the main way to explore how youth and adults constructed evidence together. This approach allowed us to observe complex social dynamics as they unfolded and to connect those insights directly to the evolving design artifact.

3.1.2 Game Design as Inquiry. We aimed to understand how game design can support and illuminate how evidence is constructed. For this reason, we chose to situate our research within the design of a game itself rather than focusing on wrap-around educational materials or other information literacy tools. Prior work has established that games are effective for exploring phenomena like misinformation and meaning-making [4, 54, 67]. Therefore, designing a game from the ground up best served our research goals. Our concepts for the game were derived from a synthesis of empirical observations from DS1 (pilot) and relevant literature about misinformation and games [4, 14, 38, 64, 76, 87]. From this, we derived three design goals that structured the development of the game:

(1) **Instructional goal:** Reflecting children's experiences of being "lost" in information, the game is designed to prompt players to navigate and interpret information collaboratively, highlighting evidence construction across cognitive, social, and material processes.
(2) **Experiential goal:** Drawing from children's accounts of confusion and overwhelm in the face of misinformation [76, 95], players are intentionally immersed in uncertainty and disorientation, enabling analysis of the cognitive, social, and material functions that support collective sensemaking.
(3) **Gameplay goal:** Building on the "rabbit hole" metaphor [64] and misinformation scholarship, the game culminates in a structured decision point ("tipping point"), offering a concrete context to study how gameplay can scaffold collaborative evidence work in information-rich environments.

3.1.3 Rationale for Cooperative Inquiry. We used Cooperative Inquiry [24], a form of participatory design that involves youth as active collaborators throughout the design process. This approach centered youth perspectives and creative agency in our study. By treating youth as co-designers, we surfaced more authentic insights and fostered shared ownership of the process and outcomes [62, 96]. Cooperative Inquiry also helped us translate participants' ideas into actionable design principles, showing how youth make sense of and justify their design suggestions across different contexts.

3.2 Research Sites, Participants, and Data Collection

Every session in the study shaped the development and testing of *Starbound Secrets*, but here we summarize the main participant groups, each named for their design team. Having two co-design groups (KidsTeam SPL, n=39; KidsTeam UW, n=21) let us verify children's ideas during design and test the game in a typical environment, like a library. Including the STEM DBR Lab (n=20) at foundry10 allowed us to observe how players experienced the game outside the specific design context.

3.2.1 Ethics and Recruitment. Our study was approved by our university's Institutional Review Board for ethics as well as the Institutional Review Board for the research institute we partnered with for design sessions 12–14. For all teams, we collected parental consent and child assent for child participants and consent for adult participants.

Session activities were age-appropriate and conducted in child-friendly environments [9], and were adapted to work with our mix of children ages 5–17, with adjustments made to the activities depending on the average age of children in the inter-generational co-design team. In addition, we used common activities across studies that used the Cooperative Inquiry method [24, 33, 61, 95]. Participants' privacy was protected with pseudonyms and secure data handling. Recruitment used snowball sampling via emails, flyers, and posters.

3.2.2 KidsTeam SPL. Our primary site for this work was an intergenerational co-design group of youth (aged 5–17), graduate students, librarians, and researchers that met at a branch of the local public library system. For these sessions, we used two camcorders and an attached microphone to capture the sessions. During the sessions, one of the researchers would move the camcorders around the room to capture various groups during design time. Along with the videos, the graduate students also wrote analytic memos reflecting on the sessions they attended. We primarily worked with *KidsTeam SPL* to design the games at the start of the project. During year 2, when the game was built in *Minecraft*, we also conducted two testing sessions with the group (DS8, DS9).

3.2.3 KidsTeam UW. This intergenerational co-design group comprises children (aged 7–12), research assistants (undergraduate and graduate students), and researchers who met in a space on a university campus. On average, we had between 3 and 4 design groups that had 2–3 children and 2 adults. During this time, we recorded the main circle time on a *Meeting Owl* camera and a Bluetooth speaker/microphone. During the design time in groups, we would record on the internal microphone and cameras on iMac computers. After each session, we would debrief with the research assistants. We worked primarily with *KidsTeam UW* to design the game and test out early prototypes (DS6, DS7).

3.2.4 STEM DBR Lab. Lastly, for our final three sessions, we partnered with a research institute, *foundry10*, to test the game with children aged between 11–14. For these tests, we had two groups of two children test for DS12 and three groups test in groups of 2–4 in DS13 and DS14. To record these tests, we also used camcorders and an attached microphone. In addition, for each testing group, a research team member conducted observations and reflective memos.

3.3 Procedure and Timeline

This study spanned from April 2022 – January 2025. The sessions were conducted in one geographic location in the United States. Our study was completed with three intergenerational co-design groups as outlined in Section 3.2.

3.3.1 Procedure of Design Sessions. All design sessions lasted approximately 90 minutes and followed a systematic structure common to Cooperative Inquiry [24]. We started all sessions with

a snack time (15 minutes) to help with relationship building. Afterward, we conducted a circle time to prepare the co-design team for design time by explaining the session's goals and activities (15 minutes). Next, we broke the participants and researchers into design teams. These teams were comprised of 2–3 children and 2–3 adults, depending on the number of participants at the session. We then either engaged in co-design activities or game testing (45 minutes) before returning for a group discussion altogether (15 minutes).

Furthermore, we distinguish between generative design sessions and playtesting sessions. All sessions used Cooperative Inquiry methods, but specific techniques were adapted to suit each session's goals. Generative sessions produced design directions and artifacts, which informed later design iterations. These were then evaluated through playtesting. Table 1 summarizes the design sessions.

3.3.2 Design Process and Iteration. Figure 2 presents the evolution of the game design. This structure, more modular than linear, emerged from an initial design insight in early co-design sessions, where children described misinformation as a feeling of being lost. This informed a teen-built prototype (v0.1) and subsequently a divergence into spatial and narrative design trajectories, which were later integrated into a pre-alpha implementation (v0.4). Subsequent development refined these into two interdependent subsystems of design: an alpha gameplay system focused on mechanics, roles, and navigation, and a beta reflection system focused on structured debriefing and meaning-making. The final system integrates both into a unified end-to-end experience.

3.4 Overview of the Final Artifact

Our design culminated in the creation of *Starbound Secrets: Down the Rabbit Hole*, an interactive live experience designed for teams of 2-3 players, guided by a game host. The final game platform choice arose in response to early co-design sessions where children noted a familiarity and wished to play the game in *Minecraft*. Our designed game artifact seeks to model the rabbit hole effect, in which players initially encounter believable pieces of evidence that become increasingly outrageous over the course of the game [64]. Players ultimately reach "the tipping point," based on their interpretations during the game. The time is broken up as follows:

(1) **Gameplay** (Approx. 45 minutes): A play experience built in a designed *Minecraft* dungeon with additional analog materials and game host facilitation.
(2) **Debrief** (Approx. 30 minutes): The facilitated discussion among players after the game to reflect on the experience.

Players begin their journey aboard a personal rocket ship, accompanied by Bomi, their loyal AI Cat Companion. Together, they are on a mission to locate the Dragon's Tear, a mythical object rumored to possess the power to heal any illness. Their quest is personal: Missy, their best friend, has fallen ill with Space Pig Flu, and the Dragon's Tear may be her only hope. The story begins as the players intercept a transmission about a long-lost Star Corp vessel that once sought the Tear; the duo uses their skills to track down the ship. They beam to the ship and find its systems offline. The ship's AI, Walter, has lost part of his memory due to data corruption and is asking for help restoring power.

Once the ship is reactivated, players begin exploring its corridors, piecing together fragments of its story. Scattered throughout are conflicting sources of information that offer competing narratives about the FooFoo aliens, the mysterious beings who possess the Dragon's Tear. An example of a Star Corp transmission is shown in Figure 3. Some sources portray them as dangerous invaders; others hint at a misunderstood species caught in a web of misinformation.

Eventually, after two verification puzzles developed around the information found within the game, players unlock the cryogenics chamber, where Bomi and Walter stand before the frozen

Version	Session	Team	Date	Type	Activity
Insight	DS1	KidsTeam UW (Children)	2022-04-26	Generative	Children brainstorm misinformation as "being lost" using Big Paper [33]. Following, two teen co-designers built initial Minecraft prototype (v0.1) before DS2.
v0.1	DS2	KidsTeam SPL (Children and Teens)	2022-10-19	Playtest + Generative	Playtest of Minecraft maze prototype (Fig. 6) and redesign ideation using Big Paper [33].
v0.2	DS4	KidsTeam SPL (Teens)	2023-03-04	Generative	Maze structure design using Bags-of-Stuff [94].
v0.3	DS3	KidsTeam SPL (Teens)	2023-03-04	Generative	Rabbit-hole narrative generation using Big Paper [33].
v0.3	DS5	KidsTeam UW (Children)	2023-03-30	Generative	Space-western narrative development using Comicboarding [60].
v0.4	DS6	KidsTeam UW (Children)	2023-04-27	Playtest	Pre-Alpha low-fidelity story/puzzle testing using *Google Slides* mock up with Stickies [33].
v0.5–0.7	DS7	KidsTeam UW (Children)	2023-07-27	Playtest	*Minecraft* world playtest with paper maps. Observation notes by adult facilitators.
v0.5–0.7	DS8	KidsTeam SPL (Children/Teens)	2023-10-25	Playtest	Introduced distributed roles (navigator, cartographer, narrator). Observation notes by adult facilitators.
v0.5–0.7	DS9	KidsTeam SPL (Children/Teens)	2023-11-01	Playtest	Added additional scaffolding via structured notes and partially completed maps and Stickies [33] for evaluation of the game.
v0.5–0.7	DS10	KidsTeam UW (Children)	2024-01-23	Playtest	Full *Minecraft* system playtest with roles and scaffolding using stickies [33] and observation notes.
v0.8–0.9	DS11	KidsTeam UW (Children)	2024-02-13	Generative	Co-designed debrief materials using Comicboarding [60]).
v0.8–0.9	DS12	STEM DBR Lab (Children)	2024-04-09	Playtest	Testing with the trustworthiness debrief generated from DS11 activity integrated into play. Observations from facilitators.
v0.8–0.9	DS13	STEM DBR Lab (Children)	2024-06-05	Playtest	Revised timing/structure of reflection activities. Observations from facilitators.
v1.0	DS14	STEM DBR Lab (Children)	2025-02-13	Playtest	Full integrated system (gameplay + reflection + narrative). Observations from facilitators.

Table 1. Design sessions, co-design teams, and session type with activity, aligned with system build phases in Figure 2.

Captain as seen in Figure 1. At this pivotal moment, players must interpret the clues they have gathered and make a collective decision: *should they believe the information that FooFoo aliens are dangerous and take the Tear from them, or wake up the Captain against Walter's warning to gather*

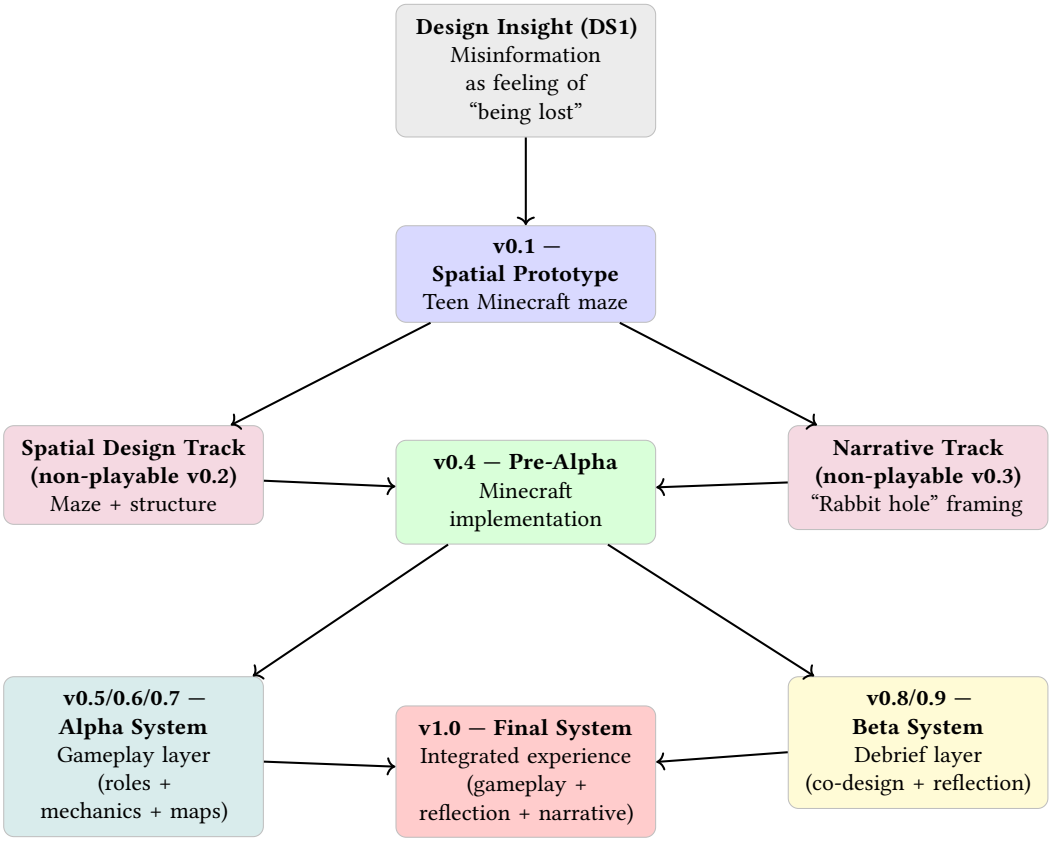


Fig. 2. Flow of game design builds, showing our divergence from an initial, teen-generated prototype into spatial and narrative tracks, convergence into a pre-alpha build, and subsequent development of parallel gameplay (alpha) and reflection (beta) systems that are integrated into the final system.

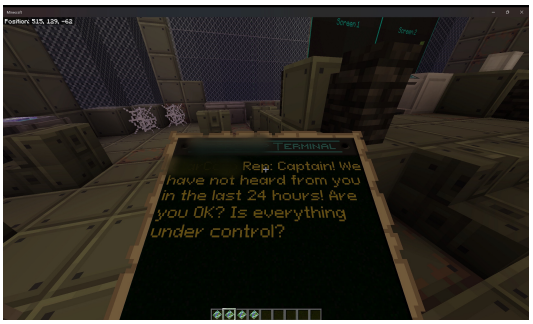


Fig. 3. An Example of a transmission to a member of the abandoned ship's crew from Star Corp

more information? In the game, Walter (representing the interests of Star Corp) and Bomi represent these two choices, such that if they wake the captain they might destroy the ship as Walter has frozen for them player's safety, or they might take the Dragon's Tear from the seemingly hostile aliens following Bomi's suggestion.

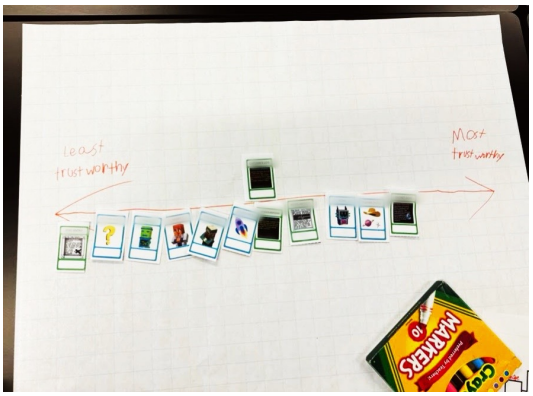


Fig. 4. An Example of the Trustworthiness Activity

After the choice, the game offers two possible endings. In one, the Captain is awakened and reveals that the FooFoo aliens are peaceful, and Walter and Star Corp simply want to steal the Tear, disregarding the aliens' safety, and the Dragon's Tear is not a cure for Space Pig Flu. However, the FooFoo aliens know another way to cure Missy. In the other, players choose not to wake him and take the Dragon's Tear, learn that the Tear does not cure Missy, only to be pursued by Star Corp operatives who seek to reclaim the ship and take the Tear from the players.

The final version of our game comprises three parts that work together. They are as follows:

(1) **The *Minecraft* World**: A world file built in *Minecraft* where players explore the digital space. Over the course of the game, they move around an abandoned spaceship, unlocking doors, finding pieces of information, and meeting new characters.
(2) **The Trustworthiness Activity** (Figure 4) : Designed by a group of children from *KidsTeam UW* during DS11, this activity is paired with the digital game experience. Each time a group encounters a new character or piece of information, they receive a card with a picture and title of what they encountered. During this time, players stop exploring and are prompted to have a conversation, placing the card on a line from least trustworthy to most trustworthy.
(3) **The Debrief**: After players reach the tipping point and make their final choice, they then reflect on how their attitudes, emotions, and decision-making processes regarding misinformation might have influenced their game experience. First, players reflect on the reasoning behind their choices during the game in their teams to understand the dynamics in a smaller setting. Afterward, all groups come together to compare their reasoning and make broader connections to their everyday lives.

3.5 Data Sources

We collected data from multiple sources to capture the complexity of the co-design process and game development. Sources included: video recordings of design sessions, artifacts generated (sketches, prototypes, notes), iterative game files and design documents, analytic memos, meeting notes, and materials from game testing. These provided a comprehensive record of both process and outcomes. Physical artifacts were photographed; digital artifacts were captured via screen recordings or screenshots. All session materials were compiled into *Google Slides* decks, and design documentation and game files were organized in *Google Drive*.

3.6 Data Analysis

Our analysis had two phases. We combined game formal analysis with qualitative coding [77] of videos and artifacts to triangulate findings about design and player practices [93]. While data from generative sessions and playtests were coded concurrently, we kept these distinctions to focus on design direction versus evaluation and refinement.

Phase 1: Analysis of Final Game and Design Process. We grounded our analysis in a formal examination of the game, focusing on its mechanics and embedded scaffolding strategies for evidence building. We completed this analysis using Lankoski and Björk's formal analysis system [50]. During this phase, we also compared the final game design with earlier iterations using game design documents such as game files and design briefs, focusing on major changes such as the change to a more open spatial design (v0.4) or the updating of the debrief (v0.9). These comparisons provided a way to contextualize and refine the open codes generated in Phase 2, by comparing our observed player experiences to the actual designs. These comparisons contextualized and refined the open codes into categories generated in Phase 2 by relating player observations to changes across major design iterations. For example, the final code *Prior Knowledge* emerged from observing that children consistently drew on existing knowledge across iterations, although the specific knowledge shifted from personal experiences (v.01–v0.3) to popular cultural understandings after the introduction of the spaceship theme (v0.4). A detailed breakdown of the final design via an analysis using Lankoski and Björk [50] is provided in Appendix A.

Phase 2: Analysis of Design Videos, Artifacts, and Session Memos. In the second phase, to code our data, we used an inductive approach to qualitative coding [77]. To conduct our analysis, we cleaned our video data by analyzing each session. To start our process, a team member reviewed each session video in five-minute segments, summarizing each segment and noting both quotes from participants and observations of participant behavior. The same team member then suggested open codes with reasoning, which we combined to generate our initial thematic codes. To ensure the reliability and accuracy of our memos, we implemented a consensus process [36], where a second team member reviewed the memos, offering suggestions and feedback. Following the research team, we refined our codebook through discussions a total of three times. For example, the quote "*the children were able to identify that Buzzfeed is a social media company whose priorities are not providing factually accurate information*" was applied with the code "*knowledge of pop and digital culture*." Our final codebook is shown in Appendix B.

3.7 Validity and Selection of Data

Given the large volume of data, we selected representative examples to illustrate our analytic process. We used triangulation [93], drawing from multiple data sources to corroborate findings. As we generated memos, coded, and selected data, we validated interpretations through a consensus process [36].

After completing the coding process, the first author systematically reviewed the analytic memos and the codebook, selecting vignettes that exemplified each identified practice through a deductive analysis, compared with our formal analysis [77]. Once examples were selected, two members of the research team again agreed on the examples through a consensus process [36], collectively reviewing each memo to confirm that the chosen examples accurately represented the design choices and player observations.

4 Findings

We organize our findings around three interrelated dimensions of documentation work that emerged during the design process, each illuminating aspects of evidence construction described earlier: (1) exploration as evidential reasoning, (2) social coordination as evidence validation, and (3) artifacts as evidential anchors. These dimensions did not appear in a fixed sequence; instead, they became increasingly entangled, together shaping how players constructed and stabilized evidence in the final game, consistent with documentation practice theory [12, 27, 55]. The following sections illustrate these player practices and the design choices informed by our iterative, generative, and playtesting process. Where relevant, we explicitly bold design decisions using Lankoski and Björk's [50] framework of goals, actions, and components. Please note that we use pseudonyms for all participants.

4.1 Exploration as Evidential Reasoning

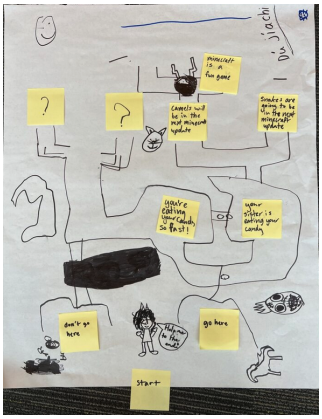
Throughout our design, we observed that introducing open and exploratory environments, paired with goals that prompted exploration (e.g., explore the ship vs. find this many clues), and populated with multi-modal, ambiguous cues, encouraged players to use their agency to apply their mental models for evidence construction through interaction with and exploration of the game system. This approach shifted player practices of evidence construction: rather than simply deciding what a clue means, players began to justify their interpretations of the game environment as evidence for their actions within the game. In other words, exploration became a process of evidential reasoning, where players constructed and justified meaning through their choices and interactions with game components via goals which directed them to act on these interactions.

Our earliest design (v0.1) embedded players' binary model misinformation directly into a maze-based game structure. This was because, in DS1–DS2, children frequently framed misinformation in simplified terms such as Charlie (child), who in DS1 suggested misinformation was akin to "*lies*." After DS1, two teen co-design participants translated this framing into binary decisions at crossroads in v0.1, such as choosing between *Buzzfeed* and *National Geographic* as reliable sources. As players explored the maze, they encountered chests (see Figure 5) that asked them to consider which source was more trustworthy, simulating this binary logic within the game design, both in the choice of evidence and the direction to take next. However, this structure consistently encouraged players to treat these interactions as selection problems rather than opportunities for interpretation. For example, the players often moved quickly through decision points without comparing or revisiting environmental cues, suggesting that though they were engaging with the evidentiary choice, they were not necessarily using these moments and information to construct evidence for continued play. For example, players rarely paused to reflect on differences between types of content or consider how credibility might shift across contexts (e.g., *Buzzfeed News* vs *Buzzfeed Entertainment*). In this case, the earlier design aligned closer to their understanding of cues in an inquiry-based game (misinformation=trustworthy or not trustworthy). However, the children's construction of evidence lasted only in the single choice, informing their actions (chose right or left), not supporting more exploratory engagement with a variety of possible cues across the game's system.

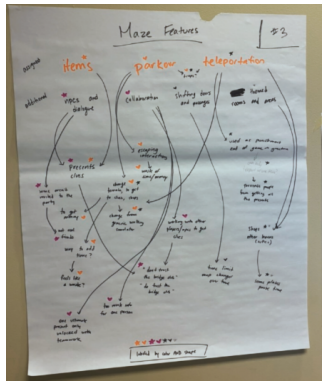
Ultimately, as we reflected on these early sessions, in accordance with our RtD method, we shifted our design away from binary structures toward open-ended spatial and narrative systems (v0.2, v0.3). Using our co-design process, we introduced the idea of a more complex maze to the teens in DS3 and DS4. They focused back in on the concept of feeling "lost," designing a space that was less of a maze and more of an exploration-based experience. For example, one of the groups created a game in which players had to navigate and explore a series of possible characters on a

Fig. 5. Allan (child) playing the original teen created prototype of the *Minecraft* Maze (v0.1) in DS2

ship, with hidden rooms and shifting reliability of clues depending on whom they should trust. In addition, *the teens told our design team that this type of game experience needed to be a mystery, not just in the narrative, but also in how the information was presented and verified.* We implemented these insights in a dungeon-like *Minecraft* environment, embedding distributed information cues rather than centralized decision gates (Figure 6).



(a) Binary Maze Design (DS2)



(b) More Complex Maze Design (DS4)

(c) Final Maze Design as Dungeon Layout

Fig. 6. Three iterations of the game's environmental, maze design moving from a gated maze, to a conceptualization of the maze as an environment, to our final dungeon design.

Prior research already indicates that players apply their mental models in play [7, 75, 91]. Therefore, after DS4, we turned to identifying what cues would allow this type of shifting information state within the game structure. In DS5, during narrative co-design discussions, Alan (child), an avid *Minecraft* player, described enjoying progression and reward systems in the game, explaining that part of the fun in *Minecraft* comes from "*leveling up my stuff*." He suggested incorporating "*netherite things*," drawing on familiar in-game value systems to imagine how the experience could be made engaging. Following, we incorporated these *Minecraft* references, in DS6 pre-alpha playtesting.

During this playest, Jessie (child), who had never played *Minecraft*, then connected a redstone puzzle in *Minecraft* to "*a shrine in Breath of the Wild, like a puzzle where you have a bunch of objects, and you move them around to uncover a button.*" In this case, Jessie utilized their preexisting knowledge of solving puzzles in another game to construct an interpretation that made sense to them for the goal of the game's system. The mental model used in sensemaking, therefore, was not only used to align with the game system, but as a frame for interpretation that justified their actions in the game [76].

To demonstrate this change, we can turn to a situation we observed during our playtesting in DS12 (v0.8). In this situation, one of the testing groups was exploring the ship, looking for a keycard to unlock the captain's quarters. This was a sub-goal of the current stage of the game. Broadly, their **goal** at this point was to "understand what happened to the crew" as they had been transported onto the empty ship where no other humans or characters existed besides the corrupted AI Walter. While exploring, they were trying to find where the keycard might be wandering around the ship, causing Mari (teen) to reflect on the **components** on the layout and visual design of the ship. They made a connection to the space ships in the TV show *Star Trek*, applying their mental model to decide where a likely place for the keycard to the bridge might be. They joked, "*If I were a Star Trek nerd, where would I put it?*," inferring that the game designers might be fans of the series based on the spaceship setting and layout. Following the players' suggestion to try to find the bridge, a common setting on the show, they adjusted their **actions** with the game systems. Guided by an open-ended goal in an open-ended space, players not only attempted to align their mental model with the game [7, 91], they used their mental model as evidence of why the game was designed in the way it was [8].

In summary, supporting players in applying their mental models as justification for their in-game choices requires guiding them through goals that prompt consideration of game components (e.g., characters, narrative beats, environmental design) as potential evidence. The game should then provide opportunities for players to act on their mental models, engaging directly in the process of justification. This highlights that **exploration functions as evidential reasoning: when players explore possible solutions, they construct and justify meaning within the game environment, focusing on their own interpretation rather than relying solely on the game system**.

4.2 Social Coordination as Evidence Validation

Building on the role of ambiguity in shifting player inference strategies, we found that explicitly distributing interpretive roles among players created structured opportunities for social validation of evidence. Assigning distinct responsibilities (reading, note-taking, and exploration) clarified how players engaged with the game's **goals** by directing them to take charge of specific **components** (e.g., the computer, the text, or the analog materials). In turn, players undertook coordinated **actions** to align their models of the game during evidence gathering. This shift moved cognitive work beyond individual sense-makeng to collaborative documentary practice [27, 37, 55, 75, 94], with evidence becoming something to be collectively validated as no person held all the answers.

Initially, we assumed that simply placing several children at a single computer with a facilitator would foster collective meaning-makeng. However, alpha playtesting (DS7-DS10) revealed a critical flaw: without defined roles, the player controlling the digital environment dominated evidence construction, leading to cognitive overload for that player and disengagement for others. For example, a participant named Victor (child) "speedran'' the game, relying on his *Minecraft* expertise, while his teammates remained passive observers. Important information was missed, and facilitation alone failed to create a meaningful division of labor. We realized that social participation required structural scaffolding, not just adult guidance.

To address this, we introduced explicit roles: Navigator (movement), Narrator (reading), and Cartographer (note-taking). This design change led to more negotiated and active collaboration, as players now relied on each other's input to validate observations and piece together findings. In KidsTeam SPL (DS8, DS9), for example, Dianne (Cartographer) recorded room labels while Markus (Navigator) led movement. When Dianne became disoriented, she asked Markus to retrace his steps, which prompted group discussion and clarification. Markus subsequently began pausing to check in with the group as he continued to play, ensuring that Dianne and the others agreed on their location in the game. Through this role-based coordination, the group collectively constructed and validated evidence, makeng interpretation a joint endeavor that required players to align their mental models and negotiate whose understanding should guide the group, consistent with collaborative gaming research [82].

This restructuring also changed the role of the adult facilitator. In DS13 (STEM DBR Lab, all teens), Sara, Mark, and Lauren added a facilitator to the "untrustworthy'' scale, reasoning that direct instructions undermined their collaborative inquiry. When the group became stuck during play, the facilitator tried to move them along by directly telling them what to do—an approach the children collectively found suspicious. In contrast, another group viewed facilitators as trustworthy when they used open-ended questions. Empowering players through game infrastructure shifted the perception of facilitation: players valued guidance that supported, rather than directed, their evidence building.

Taken together, these findings show that distributed roles did more than manage cognitive load. They instantiated a documentary practice in which evidence became a shared, negotiated object, with validation emerging from group discussion and consensus by encouraging players to align their cognitive work with each other as well as the game. **By distributing responsibilities tied to the game's goals, key components, and role-specific actions, the game's structure ensured that evidence was not just discovered but collectively scrutinized and validated through socially coordinated achievement.**

4.3 Artifacts as Evidential Anchors

Although cognitive interpretations and social coordination were essential, they were not enough to sustain evidence construction throughout extended play. Across both generative design sessions and playtests, we found that players relied on material artifacts, often requested as analog tools for children in our case, to anchor their discussions in lasting records. These inscriptions served as evidential anchors: they made individual reasoning visible to the group and allowed interpretations to persist and evolve over time. In this way, material artifacts became the central means by which players shaped, stabilized, and shared meaning through evidence within the game.

We first observed this practice early in our design process, even as our initial focus was on the digital gameplay experience. During a pre-alpha playtest in DS6, two children, Victor and Reiner, annotated the slide deck with their judgments about the game's characters. As they encountered an early version of Walter, the ship's AI, they added comments such as "*This thing looks bad*" and "*you are lying*" and shared their evolving suspicions with one another. When Walter appeared again, Reiner typed out "*RIP Sad robot. So sad and cruel. No one cares yay!!*" referencing earlier annotations. Over the course of the session, these accumulated comments became a resource the group drew on to justify their reasoning, as shown in Figure 7.

As our design evolved, children and teens increasingly asked for ways to annotate their understanding of the game world. This signaled that making group interpretations tangible was essential for building and sustaining evidence. Our alpha testing confirmed that externalizing knowledge through artifacts stabilized group meaning and coordinated sensemaking over time (see Dianne and Markus, Section 4.2).

Fig. 7. Reiner and Victor annotate on a slide during our low-fidelity Google Slides test in DS6

The value of this approach was especially clear in DS11, as we developed the debrief with KidsTeam UW. Mike (child) suggested players should interpret new information as they encountered it, perhaps by answering guiding questions. Cherry and Steph (children) proposed that players should rate the trustworthiness of characters and information on a scale. As they described their ideas, they negotiated and externalized interpretations in real time for the co-design group, demonstrating the collaborative potential of such tools. We concluded that material elements needed to serve a dual function such that they would help make individual interpretations visible and supporting ongoing revision as the game progressed.

These ideas were combined to create the "Trustworthiness Activity'' (Figure 4). In DS12 (STEM DBR Lab), we saw how externalization tools supported the evolution of discussion. For example, in DS12, the group's broad **goal** was to determine who among the ship's characters could be trusted. As Avery, Jamie, and Darren encountered new characters, they used physical character cards as anchors for their evolving reasoning about trustworthiness. In this context, the analog card served as evidence for the representation of the in-game textbfcomponents of characters and information objects. Jamie, prompted by the card, asked, "*Wait, guys! Do you think they are trustworthy or untrustworthy?*," and offered, "*I think Bomi is trying to sabotage us all.*" The group discussed and negotiated, and then Jamie placed the card under "*untrustworthy*," turning this assessment into a material anchor for the group's collective memory. Later, when they got stuck and called on Bomi for support, Jamie announced, "*I'm moving Bomi up [on the trust scale]. He's being helpful.*" This sequence shows how the group's **actions**, anchoring judgments to physical artifacts and updating them through interaction, enabled them to record, revisit, and revise their interpretations as new evidence emerged. In this way, material anchors provided stability and continuity for reasoning, connecting **goals**, **components**, and **actions** in the ongoing construction of evidence.

Across these iterations, material inscriptions enabled players to transform fleeting insights into shared, durable artifacts that function as anchors for collective reasoning. By recording and updating their interpretations, groups could continually return to, compare, and refine these anchors, building on earlier discussions rather than treating each clue in isolation. This iterative anchoring process supported a deepening, evolving understanding of the game world that was based in the exploration of possible interpretations within the environment as judged by the social work of validation done by the players.

5 Discussion

Our findings reveal that evidence construction in play can be supported by designing for ambiguity in space and narrative to prompt mental model application as a foundation for evidence-making (4.1), distributing the interpretive load across multiple players to encourage social coordination (4.2), and providing material artifacts that anchor evidence both across time and as collective spaces for inquiry (4.3).

When we reflect on these design shifts through the lens of document theory, a coherent pattern emerges: players engage in documented evidence-building across documentation dimensions through an iterative process in play. In the sections that follow, we formalize this pattern through the **Documentation-in-Play (DiP) Framework**, which identifies four mechanisms through which players' evidentiary practices gain credibility and stability within the game space. We then offer design recommendations grounded in this framework, providing actionable guidance for designers and scholars interested in supporting collaborative inquiry in games. Lastly, we discuss our theoretical contribution, urging scholars and designers to view games as document systems when designing for player-driven evidence construction.

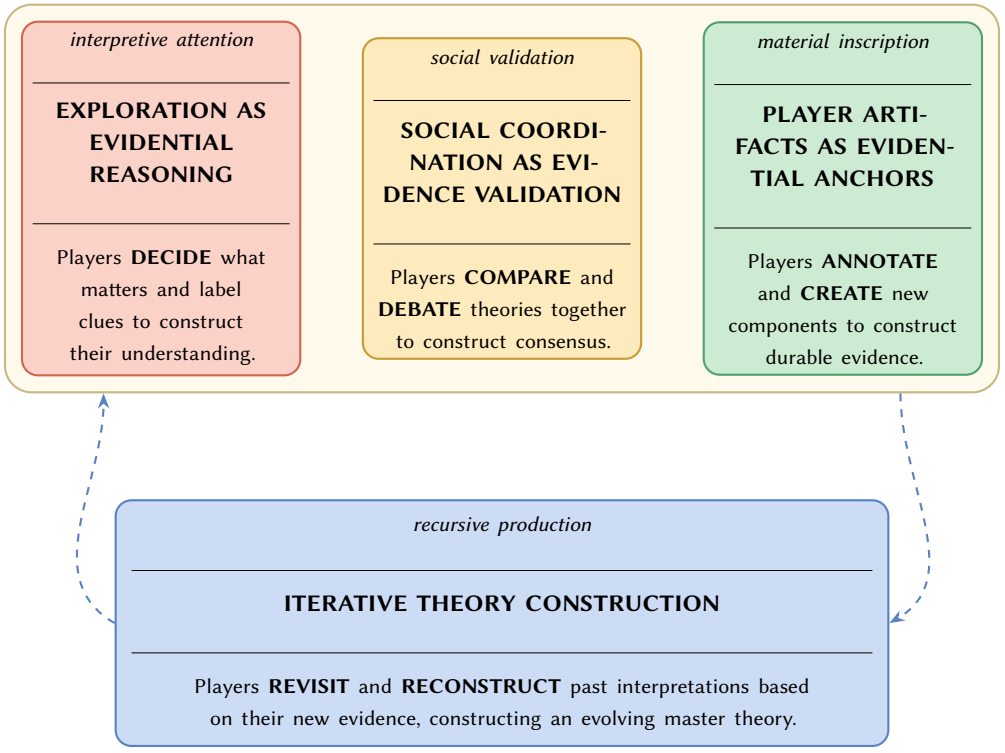


Fig. 8. **The Documentation in Play (DIP) Framework** presents evidence construction as a play loop. The top layer shows three co-present dimensions of meaning-making operating simultaneously, with mappings from our findings onto three mechanisms of *interpretive attention*, *social validation*, and *material inscription*. Inspired by our observations of players engaging in this documentation behavior of the course of play, we add *recursive production* as a feedback into all three dimensions, where players add onto their evidence construction in an iterative manner over time.

5.1 The Documentation in Play Framework: A "Play Loop" of Collaborative Evidence Construction

In this section, we introduce the **Documentation in Play (DiP) Framework** as a tool for designers and researchers of misinformation and collaborative inquiry games. The framework is grounded in observations from our playtests/co-design and informed by document theory, It is intended as a conceptual lens for interpreting evidentiary practices in play rather than a universal model of gameplay.

The DiP Framework consists of two interrelated parts, shown in Figure 8. The first, the top yellow box of the figure, is the documentation layer. It maps our findings onto the cognitive, social, and material dimensions from document theory [27, 55] and defines three concurrent player mechanisms we observed supporting evidence construction during our empirical work: interpretive attention, social validation, and material inscription. Although document theory considers these dimensions simultaneous [20, 27, 55], our findings show that certain foundational game features [50] can foreground particular mechanisms. In our playtests, we observed that these documentation practices continually reshape the game's information state and to capture this, we add a fourth mechanism, recursive production, which describes how players' ongoing engagement with interpretation, validation, and inscription reconstructs evidence throughout play. In the following sections, we define each mechanism and demonstrate its role in collaborative evidence-building, and how designers and researchers might consider using these mechanisms to design and understand evidence construction in games seeking to model ambiguous information environments.

5.1.1 Mechanism 1: Interpretive Attention. **Interpretive attention** is the process by which players notice, select, and make sense of cues in the game world, drawing on their own prior knowledge and experiences [29, 53, 88]. This mechanism is central to evidence-building as it determines which in-game information is considered meaningful, plausible, or worthy of further investigation, or in other words, what counts as potential evidence [8, 11]. In *Starbound Secrets*, for example, players interpreted ambiguous environmental features by connecting them to familiar cultural references (4.1), such as *Star Trek* or other games, using these expectations of play as evidence of why they made choices in the game [72, 90].

For design: Building in ambiguity encourages players to exercise agency in evidence construction [28]. While ambiguity is well-established in narrative design [22], our findings suggest that ambiguity in the physical environment and systemic structures is equally important. By populating the game world with cues that invite multiple plausible readings across all design layers, designers shift interpretive control from themselves to players. Rather than prescribing what counts as evidence, designers create space for players to draw on their own frames of reference and construct evidence collaboratively. Specifically, this matters for games designed to support collective-sensemaking in complex information environments because ambiguity gives players agency in deciding what counts as relevant evidence, a central step in determining what is trustworthy and in developing their own interpretive frames [75].

5.1.2 Mechanism 2: Social Validation. **Social validation** is the process by which players collectively negotiate what counts as credible evidence, align their interpretations, and stabilize shared claims through interaction and communication [34, 58, 86, 88]. In our study, this process determined which pieces of evidence gained or lost significance as they were discussed over time as only those elements that circulated among players and were collectively recognized as meaningful became stable reference points for decision-making [8, 27]. This dynamic negotiation was especially evident in how players evaluated both in-game characters and one another as sources of information.

Facilitator comments, for instance, were sometimes treated as helpful guidance and other times as suspicious (4.2), depending on the context and group discussion.

For design: Social validation requires structures that make reasoning visible and subject to collective judgment. Designers can support this by distributing information so no single player has complete knowledge (forcing negotiation), creating explicit decision points where players must reach consensus, and providing tools that externalize and make interpretations revisable (tied to Mechanism 3). The goal is not just discussion but *collaborative authentication* of evidence, making it possible for the group to collectively decide what counts as credible, and to change that decision as new information emerges. In doing so, role assignment can structure not only coordination of action but also coordination of interpretation by distributing interpretive authority, shaping where and who has the power to decide truth in the game.

5.1.3 Mechanism 3: Material Inscription. **Material inscription** is the process by which players stabilize, externalize, and manipulate information through persistent traces in the game world. This anchors meaning by transforming players' evolving interpretations into durable forms that can be revisited, revised, and used as evidence in future play [20, 27, 90]. For example, as Victor and Reiner annotated directly on the prototype slides, these inscriptions became a shared record for the group, guiding further discussion and reasoning about which characters could be trusted (4.3).

For design: Material inscriptions only stabilize evidence when they are visible, shared, and revisable by all players. Rather than functioning as static record-keeping tools, they serve as external structures for collective reasoning, allowing players to surface, compare, and refine interpretations over time. When inscriptions are private or hidden, they fragment shared understanding; when they are public and editable, they become anchors for collaborative sensemaking. This allows players to construct ad-hoc structures for reasoning during play that reflect and support their own interpretive processes.

5.1.4 Mechanism 4: Recursive Production. **Recursive production** captures the ongoing, iterative interplay among interpretive attention, social validation, and material inscription. Rather than occurring in isolation, these mechanisms continually inform and reshape one another in play [45, 55]. In our game, as players updated their interpretations on the trust scale and maps, this inscription prompted further negotiation about earlier clues, which in turn led to re-interpretation of the scale itself. For example, moving a character up or down the trust scale following a discussion would reframe how players evaluated future evidence, creating a dynamic, evolving information state (4.3).

This recursive process is what makes game-based inquiry unique: meaning is not fixed, but constantly produced and reproduced through cycles of interpretation, inscription, validation, and *interaction*. This aligns with document theory's view of documentation as performative and ongoing [27], and with understandings of games and interactive media as sites of visible, real-time knowledge production [56]. When viewed through this lens, players are not solving a puzzle with a predetermined answer but are continuously constructing and reconstructing the evidence that grounds their understanding of the game world, a feature that is increasingly designed for in considerations around reflective play [46, 58].

For design: Recursive production means building systems that continually loop players back to earlier evidence and interpretations. Designers should create moments where new information recontextualizes old evidence, where players revisit and update shared inscriptions, and where the group's evolving understanding becomes visible. This might include: periodic reflection prompts ("What did we believe then? What do we believe now?") [58], mechanics that force comparison between past and present judgments, or debrief cycles that formalize the recursive process. The goal is to make evidence-building an *ongoing performance* rather than a sequence of discrete moments

[27]. Taken together, recursive structures position gameplay as a sustained process of revising and stabilizing meaning over time, supporting repeated engagement with the practices of interpretation and judgment.

5.1.5 Framework Scope and Application. The Documentation-in-Play (DiP) Framework, grounded in empirical observation and document theory, brings these dimensions together as interacting processes that unfold over time. We position this as a **design-oriented lens** for examining how systems and player practices coincide to scaffold evidence construction. From this perspective, collaborative inquiry games can be seen as environments where players engage in forms of epistemic practice that resemble documentation, interpretation, and validation under uncertainty.

The DiP Framework offers two practical purposes. For designers, it provides a set of observable mechanisms observed over our RtD process to design toward and evaluate against. If players are exploring the game environment and labeling clues as important, comparing their ideas with other players, and then capturing these in some sort of material means (digital or analog), then the game can be see as scaffolding evidence construction through this documentation lens. For researchers, it offers a vocabulary drawn from document theory and mapped specifically to the game context.

In both cases, we present DiP as a transferable lens whose scope is analytical and generative in that it names mechanisms that may appear in varying degrees of combinations across games, players, and contexts. Therefore, while DiP was generated within the misinformation game context, we have tied our mechanisms to specific design features through our empirical work including exploration of ambiguous space, role asymmetry, and persistent, player-generated artifacts. We therefore expect the framework to extend to other inquiry focused context where interpretation, validation, and inscription are central to play.

5.2 Design Heuristics for Supporting Evidence-Building

Drawing on the DiP Framework and grounded in Lankoski and Björk's model of game analysis [50], we organize design heuristics across three design primitives (goals, components, actions) and show how each can scaffold [92] evidence-building across cognitive, social, and material dimensions [55]. In addition, to support the importance of continual updates to the documentation in both document theory and in our design work, we add an additional set of heuristics for time-based scaffolds.

In the following section, we illustrate these heuristics using our game design and provide commercial games as illustrative analogues that exhibit similar interaction patterns, though not necessarily the same underlying intent or structure. The resulting design recommendations are intended for games in which interpreting ambiguous information, constructing evidence, and coordinating meaning are central gameplay activities. We summarize these in Figure 9.

5.2.1 Cognitive Scaffolds: Prompting Interpretive Attention. Design experiences where players must continuously construct and articulate their interpretations of ambiguous information, treating their readings as evidence for understanding.

- **C-Goals (Gesture, Don't Prescribe):** Frame goals around interpretation and judgment, not collection or completion. Instead of "find these three clues," use goals such as "decide who to trust" or "investigate what happened." In *Starbound Secrets*, the goal "understand the fate of the ship's crew" prompted players to construct evidence frames; in *Outer Wilds* [G9], "solve the mystery of the solar system" lets players decide what counts as crucial evidence to progress in the game.
- **C-Components (Ambiguous Cues):** Populate the environment with cues that support multiple interpretations. The spaceship's architecture in our game invited *Star Trek* references, but also *Minecraft* references; the bridge, helm, and cargo bay were recognized not because

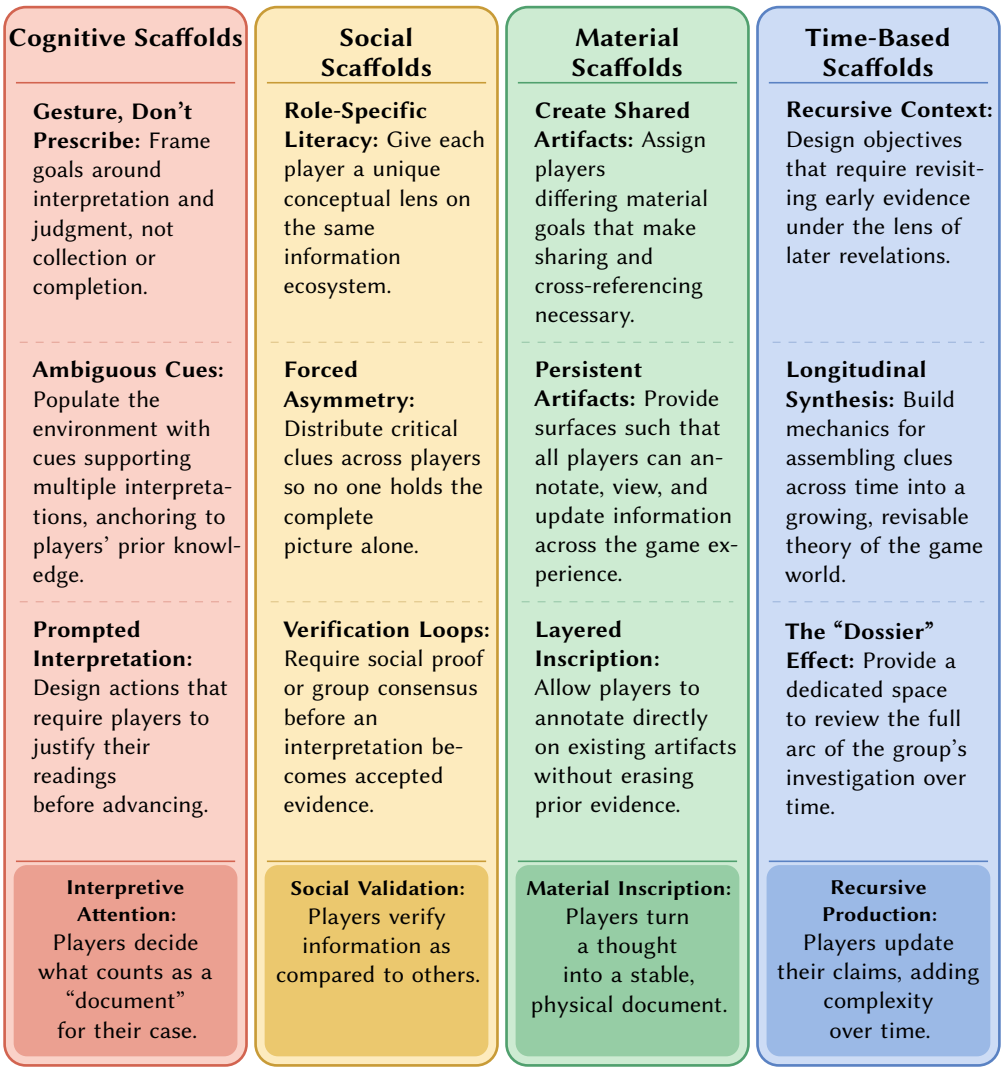


Fig. 9. Design guidelines organized by scaffold type. Each footer names the evidence construction mechanism from the Documentation in Play Framework which that scaffold supports. Full discussion with game examples can be found in Sections 5.2.1–5.2.4

we labeled them, but because players' prior knowledge made them meaningful. *Return of the Obra Dinn* [G1] does this throughout via incomplete visual information that forces players to construct evidence from ambiguity.

- **C-Actions (Prompted Interpretation):** Design player actions that require players to articulate their interpretations aloud, annotate their reasoning, or justify environmental readings. In our game, the role-based structure forced this: the narrator had to read clues aloud; the cartographer had to justify what deserved documentation. In *Kentucky Route Zero* [G3] and *Oxenfree* [G10], dialogue choices make interpretive acts visible, where players continuously

narrate their reading of character motivation, emotional state, and narrative meaning. Specifically, *Oxenfree* makes interpretation urgent and iterative so that players must use their intuition of interpretations to reshape understanding of earlier events.

5.2.2 Social Scaffolds: Enabling Validation. Structure how players collectively evaluate and authenticate evidence through distributing roles, making negotiation and consensus-building central to gameplay rather than incidental.

- **S-Goals (Role-Specific Literacy):** Assign roles that give each player responsibility for interpreting a distinct aspect of the evidence landscape. Each role provides partial access to system information, requiring players to interpret local signals and negotiate their meaning within the group. We engaged in this by providing players roles that asked them to do different information processing tasks: reading, writing, and navigating the space. Games such as *Overcooked!* [G5] or *PlateUp!* [G7] has players occupy specialized kitchen roles (e.g., chopping, cooking, plating, serving), each of which provides limited visibility into the overall order flow. Successful play depends on translating these partial views into a coordinated understanding of what the system requires at any moment.
- **S-Components (Forced Asymmetry):** Distribute information so no single player has complete knowledge. Puzzle elements require comparing distributed observations across roles, where different players access different tools, spaces, or signals. Meaning emerges through aggregating and negotiating these partial views into shared evidence. For example, in *Phasmophobia* [G8], players perceive different ghost-related cues (e.g., EMF readings, visual evidence, audio responses), but no single player can access all information unless shared by the players. The group must collectively interpret these traces and reconcile them into a shared diagnosis of where and what the ghost is in the location.
- **S-Actions (Verification Loops):** Implement game loops that require social cross-checking before a group can officially "accept" evidence. In *Among Us* [G6], Emergency Meetings are formal verification loops; in our game, the trust scale, where players receive a card to place on the scale for each new character or information piece serve this function as updating these required group discussion and consensus.

5.2.3 Material Scaffolds: Anchoring Evidence. Provide visible, persistent tools for externalizing and revising collective evidence, transforming discussion into durable documentary records.

- **M-Goals (Create Shared Artifacts):** Design goals that explicitly require creating a shared artifact that encodes their joint understanding and judgments. Progress in the game is then measured by the creation and refinement of the shared artifact. The artifact serves as both the output of collaboration and the evidence of how that collaboration unfolded. In *Starbound Secrets*, this is given to the players through the use of the trustworthiness activity and maps. Therefore, the artifact becomes the evidence from the collaboration and exploration. *Portal 2* [G13] encourages this type of shared artifact through the environment itself. Each portal placement by a player is an action that inscribes partial reasoning of the solution for the other player to build upon.
- **M-Components (Persistent Artifacts):** Provide physical or digital records that players can annotate, update, and contest that last through the game. In our game, these included maps, trust scales, and cards of the characters and information pieces. The artifact should be visible at the table (or on screen for all) and editable by all. Other examples can include *Phoenix Wright*'s Court Record [G2], which lets players compare and contrast all their information at one time as well as *Love and Deep Space*'s [G12] journal that lets player's decide what is valuable for them to use to capture their evolving relationship with the love interest.

- **M-Actions (Layered Inscription):** Allow players to physically annotate, add to, or revise inscriptions without erasing original data. When engaging on the map or trust scale, as players encountered new characters and information, they added it to the scale so that they had to layer their interpretations over previous ones. While facilitators often took pictures our trust scales through the games, the nature of the analog components means that the state changes and is more difficult for players to recreate. This means that this heuristic may be particularly well suited for digital games where the system can capture these varied states. Future designs could consider additional ways to maintain this state across analog contexts as well.*The Case of the Golden Idol* [G4] shows a similar idea, where players gradually build and refine their interpretation of a case instead of replacing earlier guesses.

5.2.4 Time-Based Scaffolds: Encouraging Recursive Production of Meaning. Design systems that periodically loop players back to earlier evidence, prompting continuous reconstruction of understanding and making the recursive nature of evidence-building explicit.

- **T-Goals (Recursive Context):** Design goals that explicitly ask players to re-evaluate earlier evidence in light of new information. In our game this happens through prompting with physical cards. *Outer Wilds* [G9] does this through time loops; debrief sessions do this explicitly in many learning games.
- **T-Components (Longitudinal Synthesis):** Provide mechanics for reviewing and comparing evidence across time. A log that tracks when you learned what; a timeline of discoveries that grows as play unfolds. The goal is to make the evolution of understanding visible. One example of this might include *Breath of the Wild*'s [G11] Hero's Path mode, which allows players to review their movements across the world over time, making prior exploration decisions visible and supporting reflection on how understanding of the game world has evolved.
- **T-Actions (The "Dossier" Effect):** Provide a central, dedicated space to visually review the historical accumulation, growth, and timeline of investigation. Players need a moment to step back and see how their understanding has changed, not just in the moment, but in visible, comparative form. *Phoenix Wright's* Court Record is another example of this concept.

6 Limitations and Future Work

Our study, though conducted with a variety of children, teens, and adult co-designers over 2.5 years, engaged children only in a single geographic location. Therefore, our work offers contributions to theory rather than statistical generalization [93]. Additionally, not all children attended every session due to scheduling and the length of the study.

While this work explores design mechanisms for evidentiary work as a form of meaning-making in games, our empirical analysis is based on a single custom-designed game focused on misinformation. This limits the extent to which our findings can be generalized across other game types and contexts. Future work should examine whether these mechanisms appear in other genres or whether they are specific to games designed around information and interpretation.

The children in our study also knew one another, a factor which may make it more likely that their experiences represent the specific group. Conducting more co-design work and interviews with teachers, parents, and youth, in addition to observing children and adults in unstructured play environments, could expand this work to a more holistic perspective. In addition, exploring these ideas outside the co-design process may offer further clarity and refinement of the suggested mechanisms. In future work, we hope to further evaluate the efficacy of our game in how its design supports or hinders meaning-making, and to identify the nuances of how these different mechanisms intersect more clearly across play domains.

7 Conclusion

Our study used a Research through Design process of co-designing with youth to surface game primitives and player practices that illuminate how evidentiary work unfolds during collaborative, inquiry-focused play. In this paper, we introduce the Documentation in Play (DiP) Framework as a language for designers and game scholars to discuss the mechanisms through which evidence becomes credible, stable, and shareable.

Our paper argues that designers who wish to engage players in taking on agency in determining the meaning they form from a game should support players in engaging in interpretation, validating those interpretations with others, and materializing them for use in the game and by other players. In In this sense, the game becomes both a document and a system of documentary practice, in which meaning, strategy, and interpretation are continually produced, revised, and recorded through play itself [20]. This approach shifts players from users of a system to active practitioners of epistemic work, demonstrating that "information" in the game is not a stable property of code or design, but is produced through labor, negotiation, practice, and performance [27, 56]—a labor that designers can intentionally scaffold and support.

Acknowledgments

The authors would like to thank the librarians and staff at Seattle Public Library for their dedication and care throughout this project. We also wish to express our sincere gratitude to the child co-designers, undergraduate students, and graduate students whose invaluable contributions enriched this study. Lastly, we thank foundry10 for their support of the study and our game, as well as for sharing their time and expertise. This work was supported in part by IMLS grants LG-252291-OLS-22 and LG-256571-OLS-24.

References

[1] Abigail D Adams. 2023. Werewolf on campus: A case study in inoculation theory and gamified one-shot library instruction. *Public Services Quarterly* 19, 4 (2023), 255–272. doi:10.1080/15228959.2023.2252322

[2] Dennis Ang and Alex Mitchell. 2019. Representation and frequency of player choice in player-oriented dynamic difficulty adjustment systems. In *Proceedings of the annual symposium on computer-human interaction in play*. Association for Computing Machinery, New York, NY, 589–600. doi:10.1145/3311350.3347165

[3] Tiziano Antognozzi, Alessandro Crociata, and Pier L Sacco. 2025. The Ludic Paradox: Why Behavioral Change Through Games Requires Agency, Not Control. doi:10.31235/osf.io/jdyhb_v1

[4] Melisa Basol, Jon Roozenbeek, and Sander van der Linden. 2020. Good News about Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News. *Journal of Cognition* 3, 1 (2020), 2. doi:10.5334/joc.91

[5] Johan Blomberg. 2018. The semiotics of the game controller. *Game Studies* 18, 2 (2018), 311–323. https://gamestudies.org/1802/articles/blomberg

[6] Ian Bogost. 2007. *Persuasive Games: The Expressive Power of Videogames*. The MIT Press, Cambridge, MA.

[7] Andy Boyan, Rory McGloin, and Joe A Wasserman. 2018. Model matching theory: A framework for examining the alignment between game mechanics and mental models. *Media and Communication* 6, 2 (2018), 126–136. doi:10.17645/mac.v6i2.1326

[8] Suzanne Briet. 2006. *What is documentation?: English translation of the classic French text*. Bloomsbury Academic, London, United Kingdom.

[9] Peter C Brown, Henry L Roediger III, and Mark A McDaniel. 2014. *Make it stick: The science of successful learning*. Harvard University Press, Cambridge, MA.

[10] Josef Buchner. 2025. Playing an augmented reality escape game promotes learning about fake news. *Technology, Knowledge and Learning* 30, 1 (2025), 425–445. doi:10.1007/s10758-024-09749-y

[11] Michael Buckland. 2017. *Information and society*. The MIT Press, Cambridge, MA.

[12] Michael K Buckland. 1991. Information as thing. *Journal of the American Society for information science* 42, 5 (1991), 351–360. https://doi.org/10.1002/(SICI)1097-4571(199106)42:5%3C351::AID-ASI5%3E3.0.CO;2-3

[13] Shelly Chaiken and Alison Ledgerwood. 2012. A theory of heuristic and systematic information processing. *Handbook of theories of social psychology* 1 (2012), 246–266.

[14] Yeonhee Cho, Chris Coward, Jacob Lackner, Travis Willingham Windleharth, and Jin Ha Lee. 2023. The use of an escape room as an immersive learning environment for building resilience to misinformation. *Journal of Librarianship and Information Science* 57 (2023). Issue 2. doi:10.1177/09610006231208027

[15] Tom Cole and Marco Gillies. 2021. Thinking and doing: Challenge, agency, and the eudaimonic experience in video games. *Games and Culture* 16, 2 (2021), 187–207. doi:10.1177/1555412019881536

[16] Mia Consalvo and Nathan Dutton. 2006. Game analysis: Developing a methodological toolkit for the qualitative study of games. *Game studies* 6, 1 (2006), 1–17. https://gamestudies.org/0601/articles/consalvo_dutton

[17] Ruth S Contreras-Espinosa and Jose Luis Eguia-Gomez. 2023. Evaluating video games as tools for education on fake news and misinformation. *Computers* 12, 9 (2023), 188. doi:10.3390/computers12090188

[18] Chris Coward, Nisha Devasia, and Jin Ha Lee. 2025. Escape Rooms for Misinformation Education: A Case Study of Co-design with Two Communities. In *Conference Proceedings of DiGRA 2025: Games at the Crossroads*. doi:10.26503/dl.v2025i2.2470

[19] Ann Cunliffe and Chris Coupland. 2012. From hero to villain to hero: Making experience sensible through embodied narrative sensemaking. *Human relations* 65, 1 (2012), 63–88. doi:10.1177/0018726711424321

[20] Ronald E Day. 2019. *Documentarity: Evidence, ontology, and inscription.* The MIT Press, Cambridge, MA.

[21] Scott DeJong. 2023. Playing with fake news: state of fake news video games. *The International Journal of Games and Social Impact* 1, 1 (2023), 94–111. doi:10.24140/ijgsi.v1.n1.05

[22] Nisha Devasia and Jin Ha Lee. 2024. The role of narrative in misinformation games. *Harvard Kennedy School Misinformation Review* 5 (2024). Issue 5. doi:10.37016/mr-2020-158

[23] Nisha Devasia, Runhua Zhao, and Jin Ha Lee. 2025. Does the Story Matter? Applying Narrative Theory to an Educational Misinformation Escape Room Game. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, 1–15. doi:10.1145/3706598.3713131

[24] Allison Druin. 1999. Cooperative inquiry: developing new technologies for children with children. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, 592–599. doi:10.1145/302979.303166

[25] Tianqi Duan. 2023. Preparing Youth With Media Literacy for Future Digital Civic Engagement. In *Proceedings of the 7th International Conference on Education and Multimedia Technology* (Tokyo, Japan) *(ICEMT '23)*. Association for Computing Machinery, New York, NY, USA, 315–320. doi:10.1145/3625704.3625718

[26] Carlo Fabricatore, Dimitar Gyaurov, and Ximena Lopez. 2019. An exploratory study of the relationship between meaning-making and quality in games. *Multimedia Tools and Applications* 78, 10 (2019), 13539–13564. doi:10.1007/s11042-019-7232-1

[27] Bernd Frohmann. 2004. Documentation redux: Prolegomenon to (another) philosophy of information. *Library trends* 52, 3 (2004), 387–407. http://hdl.handle.net/2142/1683

[28] William W. Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 233–240. doi:10.1145/642611.642653

[29] James Paul Gee. 2008. Video games and embodiment. *Games and culture* 3, 3-4 (2008), 253–263. doi:10.1177/1555412008317309

[30] René Glas, Jasper van Vught, Timo Fluitsma, Teresa De La Hera, and Salvador Gómez-García. 2023. Literacy at play: an analysis of media literacy games used to foster media literacy competencies. *Frontiers in communication* 8 (2023), 1155840. doi:10.3389/fcomm.2023.1155840

[31] Seth Glickman, Nathan McKenzie, Joseph Seering, Rachel Moeller, and Jessica Hammer. 2018. Design Challenges for Livestreamed Audience Participation Games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 187–199. doi:10.1145/3242671.3242708

[32] Anuradha A Gokhale. 1995. Collaborative Learning Enhances Critical Thinking. *Journal of Technology Education* 7, 1 (1995), 22–30. doi:10.21061/jte.v7i1.a.2

[33] Mona Leigh Guha, Allison Druin, and Jerry Alan Fails. 2013. Cooperative Inquiry revisited: Reflections of the past and guidelines for the future of intergenerational co-design. *International Journal of Child-Computer Interaction* 1, 1 (2013), 14–23. doi:10.1016/j.ijcci.2012.08.003

[34] Stuart Hall. 2014. Encoding and decoding the message. *The discourse studies reader: Main currents in theory and analysis* (2014), 111–121.

[35] Christian Hayes. 2020. Meaning-Making in Indie Games: The Intersection of Narrative and Interactivity in Non-Traditional Videogames. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Canada) *(CHI PLAY '20)*. Association for Computing Machinery, New York, NY, USA, 21–23. doi:10.1145/3383668.3419956

[36] Clara E Hill, Sarah Knox, Barbara J Thompson, Elizabeth Nutt Williams, Shirley A Hess, and Nicholas Ladany. 2005. Consensual qualitative research: An update. *Journal of counseling psychology* 52, 2 (2005), 196. doi:10.1037/0022-0167.52.2.196

[37] James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput.-Hum. Interact.* 7, 2 (June 2000), 174–196. doi:10.1145/353485.353487

[38] Philip N Howard, Lisa-Maria Neudert, Nayana Prakash, and Steven Vosloo. 2021. Digital misinformation/disinformation and children. *UNICEF. Retrieved on February* 20 (2021), 2021. https://www.unicef.org/innocenti/reports/digital-misinformation-disinformation-and-children

[39] Ying-Shao Hsu, Ting-Ling Lai, and Wei-Hsiu Hsu. 2015. A design model of distributed scaffolding for inquiry-based learning. *Research in Science Education* 45, 2 (2015), 241–273. doi:10.1007/s11165-014-9421-2

[40] Edwin Hutchins. 1995. *Cognition in the Wild.* The MIT press, Cambridge, MA.

[41] Kokil Jaidka, Hansin Ahuja, and Lynnette Hui Xian Ng. 2024. It Takes Two to Negotiate: Modeling Social Exchange in Online Multiplayer Games. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 85 (April 2024), 22 pages. doi:10.1145/3637362

[42] Henry Jenkins, Ravi Purushotma, Margaret Weigel, Katie Clinton, and Alice J Robison. 2009. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century.* The MIT Press, Cambridge, MA. doi:10.7551/mitpress/8435.001.0001

[43] S Mo Jones-Jang, Tara Mortensen, and Jingjing Liu. 2019. Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American behavioral scientist* 65, 2 (2019), 371–388. doi:10.1177/0002764219869406

[44] Jesper Juul. 2011. *Half-real: Video games between real rules and fictional worlds.* The MIT press, Cambridge, MA.

[45] Sarah Kember and Joanna Zylinska. 2012. *Life after new media: Mediation as a vital process.* The MIT Press, Cambridge, MA. doi:10.7551/mitpress/8796.001.0001

[46] Rilla Khaled. 2018. Questions over answers: Reflective game design. In *Playful disruption of digital media.* Springer, 3–27. doi:10.1007/978-981-10-1891-6_1

[47] Kristian Kiili. 2005. Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education* 8, 1 (2005), 13–24. doi:10.1016/j.iheduc.2004.12.001

[48] Kristian Kiili, Juho Siuko, and Manuel Ninaus. 2024. Tackling misinformation with games: a systematic literature review. *Interactive Learning Environments* 32, 10 (2024), 7086–7101. doi:10.1080/10494820.2023.2299999

[49] Loukas Konstantinou and Evangelos Karapanos. 2025. Behavior Change Interventions Combating Online Misinformation: A Scoping Review. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems.* 1–19. https://doi.org/10.1145/3706598.3713127

[50] Petri Lankoski and Staffan Björk. 2015. Formal analysis of gameplay. In *Game Research Methods: An Overview.* Carnegie Mellon University: ETC Press, Pittsburgh, PA, 23–35. doi:10.1184/R1/6686765

[51] Bruno Latour and Steve WooIgar. 1986. *Laboratory Life: The Construction of Scientific Facts* (2 ed.). Princeton University Press, Princton, NJ. http://www.jstor.org/stable/j.ctt32bbxc

[52] Stephan Lewandowsky, Ullrich K. H. Ecker, Colleen M. Seifert, Norbert Schwarz, and John Cook. 2012. Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest* 13, 3 (2012), 106–131. doi:10.1177/1529100612451018 arXiv:https://doi.org/10.1177/1529100612451018 PMID: 26173286.

[53] Craig A Lindley and Charlotte C Sennersten. 2008. Game play schemas: from player analysis to adaptive game mechanics. *International Journal of Computer Games Technology* 2008, 1 (2008), 216784. doi:10.1155/2008/216784

[54] Ioana Literat, Yoo Kyung Chang, and Shu-Yi Hsu. 2020. Gamifying fake news: Engaging youth in the participatory design of news literacy games. *Convergence* 26, 3 (2020), 503–516. doi:10.1177/1354856520925732

[55] Niels Windfeld Lund. 2024. *Introduction to documentation studies.* Facet Publishing, London, United Kingdom.

[56] Lev Manovich. 2013. *Software takes command.* Bloomsbury Academic, London, United Kingdom.

[57] Weijie Mao, Yunhuo Cui, Ming M Chiu, and Hao Lei. 2022. Effects of game-based learning on students' critical thinking: A meta-analysis. *Journal of Educational Computing Research* 59, 8 (2022), 1682–1708. doi:10.1177/07356331211007098

[58] Josh Aaron Miller, Kutub Gandhi, Matthew Alexander Whitby, Mehmet Kosa, Seth Cooper, Elisa D. Mekler, and Ioanna Iacovides. 2024. A Design Framework for Reflective Play. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24).* Association for Computing Machinery, New York, NY, USA, Article 519, 21 pages. doi:10.1145/3613904.3642455

[59] Luis C. Moll, Cathy Amanti, Deborah Neff, and Norma Gonzalez. 1992. Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice* 31, 2 (1992), 132–141. doi:10.1080/00405849209543534

[60] Neema Moraveji, Jason Li, Jiarong Ding, Patrick O'Kelley, and Suze Woolf. 2007. Comicboarding: using comics as proxies for participatory design with children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07).* Association for Computing Machinery, New York, NY, USA, 1371–1374.

doi:10.1145/1240624.1240832

[61] James Newman. 2011. (Not) playing games: player-produced walkthroughs as archival documents of digital gameplay. *International Journal of Digital Curation* 6, 2 (2011), 109–127. doi:10.2218/ijdc.v6i2.206

[62] Michele Newman, Kaiwen Sun, Ilena B Dalla Gasperina, Grace Y Shin, Matthew Kyle Pedraja, Ritesh Kanchi, Maia B Song, Rannie Li, Jin Ha Lee, and Jason Yip. 2024. " I want it to talk like Darth Vader": Helping Children Construct Creative Self-Efficacy with Generative AI. In *Proceedings of the 2024 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, 1–18. doi:10.1145/3613904.3642492

[63] Don Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books, New York, NY.

[64] Derek O'Callaghan, Derek Greene, Maura Conway, Joe Carthy, and Pádraig Cunningham. 2015. Down the (white) rabbit hole: The extreme right and online recommender systems. *Social Science Computer Review* 33, 4 (2015), 459–478. doi:10.1177/0894439314555532

[65] Raymond Pun. 2017. Hacking the research library: Wikipedia, Trump, and information literacy in the escape room at Fresno State. *The Library Quarterly* 87, 4 (2017), 330–336. doi:10.1086/693489

[66] Sadhana Puntambekar. 2022. Distributed scaffolding: Scaffolding students in classroom environments. *Educational Psychology Review* 34, 1 (2022), 451–472. doi:10.1007/s10648-021-09636-3

[67] Jon Roozenbeek and Sander van der Linden. 2020. Breaking Harmony Square: A game that "inoculates" against political misinformation. *Harvard Kennedy School (HKS) Misinformation Review* 1 (2020). Issue 8. doi:10.37016/mr-2020-47

[68] Jon Roozenbeek, Sander Van Der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky. 2022. Psychological inoculation improves resilience against misinformation on social media. *Science advances* 8, 34 (2022). doi:10.1126/sciadv.abo6254

[69] David W Shaffer. 2006. Epistemic frames for epistemic games. *Computers & education* 46, 3 (2006), 223–234. doi:10.1016/j.compedu.2005.11.003

[70] Miguel Sicart. 2008. Defining Game Mechanics. *Game Studies* 8, 2 (2008). https://gamestudies.org/1103/articles/sicart

[71] Miguel Sicart. 2019. Play in the information age. *Philosophy & Technology* 32, 3 (2019), 517–534. doi:10.1007/s13347-018-0311-9

[72] Olle Sköld. 2018. *Documenting videogame communities: A study of community production of information in social-media environments and its implications for videogame preservation*. Ph. D. Dissertation. Department of ALM.

[73] Olle Sköld. 2018. Understanding the "expanded notion" of videogames as archival objects: A review of priorities, methods, and conceptions. *Journal of the Association for Information Science and Technology* 69, 1 (2018), 134–145. doi:10.1002/asi.23875

[74] Florence Smith Nicholls and Michael Cook. 2025. Archaeological Gameworld Affordances: A Grounded Theory of How Players Interpret Environmental Storytelling. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, 1–20. doi:10.1145/3706598.3714036

[75] K Starbird. 2024. Facts, frames, and (mis) interpretations: understanding rumors as collective sensemaking. https://www.cip.uw.edu/2023/12/06/rumors-collective-sensemaking-kate-starbird/

[76] Kate Starbird, Emma Spiro, Isabelle Edwards, Kaitlyn Zhou, Jim Maddock, and Sindhuja Narasimhan. 2016. Could this be true? I think so! Expressed uncertainty in online rumoring. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, 360–371. doi:10.1145/2858036.2858551

[77] Anselm L Strauss. 1987. *Qualitative analysis for social scientists*. Cambridge university press, Cambridge, United Kingdom.

[78] Melanie Swalwell, Angela Ndalianis, and Helen Stuckey (Eds.). 2017. *Fans and videogames: histories, fandom, archives*. Routledge, New York, NY. doi:10.4324/9781315563480

[79] Karen Tanenbaum and Theresa J Tanenbaum. 2009. Commitment to meaning: A reframing of agency in games. *UC Irvine: Digital Arts and Culture* (2009). https://escholarship.org/uc/item/6f49r74n

[80] Salen Tekinbas Katie and Eric Zimmerman. 2003. *Rules of play: Game design fundamentals*. MIT press.

[81] Phoebe O. Toups Dugas, Nicole K. Crenshaw, Rina R. Wehbe, Gustavo F. Tondello, and Lennart E. Nacke. 2016. "The Collecting Itself Feels Good": Towards Collection Interfaces for Digital Game Objects. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 276–290. doi:10.1145/2967934.2968088

[82] Phoebe O. Toups Dugas, Jessica Hammer, William A. Hamilton, Ahmad Jarrah, William Graves, and Oliver Garretson. 2014. A framework for cooperative communication game mechanics from grounded theory. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play* (Toronto, Ontario, Canada) *(CHI PLAY '14)*. Association for Computing Machinery, New York, NY, USA, 257–266. doi:10.1145/2658537.2658681

[83] April Tyack and Elisa D. Mekler. 2020. Self-Determination Theory in HCI Games Research: Current Uses and Open Questions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3313831.3376723

[84] April Tyack and Peta Wyeth. 2021. "The Small Decisions Are What Makes it Interesting": Autonomy, Control, and Restoration in Player Experience. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 282 (Oct. 2021), 26 pages. doi:10.1145/3474709

[85] Rita Santoyo Venegas. 2021. Digital play as an epistemic experience. In *Play, Philosophy and Performance*. Routledge, London, United Kingdom, 216–231. doi:10.4324/9780429323737

[86] Lev S Vygotsky. 1978. *Mind in society: The development of higher psychological processes*. Vol. 86. Harvard university press. doi:10.2307/j.ctvjf9vz4

[87] Stacey Wedlake, Chris Coward, and Jin Ha Lee. 2024. How games can support misinformation education: A sociocultural perspective. *Journal of the Association for Information Science and Technology* 75, 13 (2024), 1480–1497. doi:10.1002/asi.24954

[88] Karl E Weick. 1995. *Sensemaking in organizations*. Sage publications, Thousand Oaks, CA.

[89] Garrison Wells, Agnes Romhanyi, and Alaina Klaes. 2024. Sus: Modifying Among Us for Misinformation Discernment. In *Proceedings of the 19th International Conference on the Foundations of Digital Games* (Worcester, MA, USA) *(FDG '24)*. Association for Computing Machinery, New York, NY, USA, Article 58, 4 pages. doi:10.1145/3649921.3656990

[90] Niels Windfeld Lund. 2010. Document, text and medium: concepts, theories and disciplines. *Journal of documentation* 66, 5 (2010), 734–749. doi:10.1108/00220411011066817

[91] Travis W Windleharth. 2021. *Mental Models, Meaning, & Games: An Interpretative Phenomenological Analysis of Player Meaning Making in a Complex STEM Themed Video Game*. University of Washington, Seattle, WA. http://hdl.handle.net/1773/48520

[92] David Wood, Jerome S Bruner, and Gail Ross. 1976. The role of tutoring in problem solving. *Journal of child psychology and psychiatry* 17, 2 (1976), 89–100.

[93] Robert K. Yin and Donald T. (Donald Thomas) Campbell. 2018 - 2018. *Case study research and applications : design and methods* (sixth edition. ed.). SAGE Publications, Inc., Thousand Oaks, California.

[94] Jason Yip, Tamara Clegg, Elizabeth Bonsignore, Helene Gelderblom, Emily Rhodes, and Allison Druin. 2013. Brownies or bags-of-stuff? domain expertise in cooperative inquiry with children. In *Proceedings of the 12th International Conference on Interaction Design and Children* (New York, New York, USA) *(IDC '13)*. Association for Computing Machinery, New York, NY, USA, 201–210. doi:10.1145/2485760.2485763

[95] Jason Yip, Michele Newman, Runhua Zhao, Darae Kim, Jan R. Lim, Matthew Kyle Pedraja, Swati Sachdeva, Xiaoyu Zheng, Yifang Zhou, Chris Coward, and Jin Ha Lee. 2026. Rethinking Misinformation: A Holistic Community Model for Youth Resilience through Socioemotional Learning and Sociocultural Design. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. Association for Computing Machinery, New York, NY, USA, Article 46, 21 pages. doi:10.1145/3772318.3790876

[96] Jason C. Yip, Elizabeth Foss, Elizabeth Bonsignore, Mona Leigh Guha, Leyla Norooz, Emily Rhodes, Brenna McNally, Panagis Papadatos, Evan Golub, and Allison Druin. 2013. Children initiating and leading cooperative inquiry sessions. In *Proceedings of the 12th International Conference on Interaction Design and Children* (New York, New York, USA) *(IDC '13)*. Association for Computing Machinery, New York, NY, USA, 293–296. doi:10.1145/2485760.2485796

[97] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 493–502. doi:10.1145/1240624.1240704

Ludography

[1] 3909 LLC. 2019. *The Return of Obra Dinn*. Game [macOS, Windows, Switch, PlayStation 4, Xbox One]. Lucas Pope.

[2] Capcom Production Studio 4. 2001. *Phoenix Wright: Ace Attorney*. Game [Game Boy, Advance, Nintendo DS, Windows, Wii, iOS]. Capcom Co., Ltd, Osaka, Japan.

[3] Cardboard Computer. 2013. *Kentucky Route Zero*. Game [Windows, macOS, Linux, PlayStation 4, Xbox One, Nintendo Switch, iOS, Android]. Annapurna Interactive, Los Angeles, California.

[4] Color Gray Games. 2022. *The Case of the Golden Idol*. Game [macOS, Windows, Nintendo Switch, Android, iOS, Xbox One, Xbox Series X/S, PlayStation 5]. Playstack, London, England.

[5] Ghost Town Games. 2016. *Overcooked!* Game [PlayStation 4, Windows, Xbox One, Nintendo Switch, Xbox Series X/S, PlayStation 5]. Team17, Wakefield, England.

[6] Innersloth. 2018. *Among Us*. Game [Android, iOS, Windows, Nintendo Switch, PlayStation 4, PlayStation 5, Xbox One, Xbox Series X/S]. Innersloth, Redmond, Washington.

[7] It's Happening. 2022. *PlateUp!* Game [Windows, Nintendo Switch, PlayStation 4, PlayStation 5, Xbox One, Xbox Series X and Series S]. Yogscast Games, London, England.

[8] Kinetic Games. 2019. *Phasmophobia*. Game [Windows, PlayStation 5, Xbox Series X/S, Nintendo Switch 2]. Kinetic Games, Southhampton, England.

[9] Mobius Digital. 2019. *Outer Wilds*. Game [Windows, Xbox One, PlayStation 4, PlayStation 5, Xbox Series X/S, Nintendo Switch]. Annapurna Interactive, Los Angeles, California.
[10] Night School Studio. 2016. *Oxenfree*. Game [Windows, macOS, Linux, PlayStation 4, Xbox One, Nintendo Switch, iOS, Android]. Night School Studio, Glendale, California.
[11] Nintendo EPD. 2017. *The Legend of Zelda: Breath of the Wild*. Game [Wii U, Nintendo Switch, Nintendo Switch 2]. Nintendo, Kyoto, Japan.
[12] Papergames. 2001. *Love and Deep Space*. Game [Android, iOS]. Infold Games, Singapore.
[13] Valve. 2011. *Portal 2*. Game [Windows, macOS, Linux, PlayStation 3, Xbox 360, Nintendo Switch]. Valve, Bellevue, Washington.

A Formal Analysis of v 1.0 of *Starbound Secrets*

For the purposes of this paper, our formal analysis of v1.0 (Tested in DS14) is presented via a discussion of the primary primitives, components, actions, and goals as described by Lankoski and Björk [50].

Goals orient players toward evidentiary activity. The overarching objective, in our case, identifying who to trust, directs players to attend to cues in the environment and to verify their interpretations through exploration. Because the goal does not prescribe a single correct path, it invites players to negotiate which interpretations feel valid and what information should count as evidence . Importantly, these goals were reinforced by the game's social dynamics, namely the facilitator and the character Bomi, who continually ask players to focus on these goals to keep them engaged.

Actions operate as evidentiary negotiation and methods of inscription. During the game, players gather, test, and manipulate information by interacting with components across digital, social, and physical media. In this context, as the goals orient players toward specific interpretive tasks once directed. For example, in the game, the design asks to move characters and information on the trust scale. This action serves both to allow players to prompt negotiation with other players (actions in the game's social space) and to inscribe their discussion in the game's physical components (actions in the game's analog space). In addition, the system performs actions such as revealing information, which then opens new actions for the players, thereby shifting which evidence becomes available.

Components function as inscriptions that present players with traces to interpret. In the game, some components are fixed, allowing us as designers to guide and suggest types of interpretations. For example, we could control the design of non-player characters, such as Bomi, to help trigger specific interpretations. In addition, through the design process with youth, it became increasingly important to include *dynamic* components, i.e., components that players could explicitly control and change over time, such as the Trustworthiness Scale. In this way, the components were both designed and player-generated, allowing for players to both interpret and build new interpretations in the game.

SUMMARY: The combination of these primitives gives rise to the game's rules and constraints as a system that seeks to regulate when and how evidence becomes available (Actions and Components), and then explicitly directs them to engage and change the system (Goals). These constraints shape the sequence interpretation and determine what players come to treat as credible or actionable within the game. Taken together, these design elements position the game as a system that structures the production, contestation, and revision of evidence as play occurs over time.

B Final Codebook

Category	Code	Description
Prior Knowledge	*Minecraft* Knowledge	Participants make connections to their previous knowledge of *Minecraft* as a game
	Game Literacy	Participants make connections to their previous knowledge of video games
	Knowledge of Pop and Digital Culture	Participants make connections to their previous knowledge of digital culture
	Information Literacy	Participants make connections to their previous knowledge of misinformation or experiences with information
Social Reasoning	Motivation of Players	Participants note their motivation for performing an action
	Social Trust	Participants note their motivation for trusting a character or information in a specific way
	Validating	Participants prompt one another about what they encounter during the game to verify interpretations
	Social Dynamics	Participants talk with others to support thinking about a character or information in a specific way
Tools	Externalizing Thinking	Participants note or use an analog tool to support thinking about a character or information in a specific way
	Modal Bridges	A component of the game helps players switch between modes (digital, physical, and social)

Received February 2026; revised June 2026; accepted July 2026