



Designing believable worlds: NPC AI and immersive evolution in *The Legend of Zelda* series

Esboña A J

Submitted: June 2, 2026, Revised: version 1, July 20, 2026, version 2, July 26, 2026

Accepted: July 28, 2026

Abstract

As *The Legend of Zelda* franchise approaches its 40th anniversary, it provides an opportunity to examine how advances in non-player character (NPC) artificial intelligence (AI) have contributed to increasingly immersive open-world experiences. Although previous work has described NPC AI architectures and theories of immersion, no quantitative framework has been proposed for systematically comparing their evolution within a long-running game franchise. This study addresses that gap by introducing a seven-dimension observational framework that operationalizes NPC AI through measurable behavioral surrogate variables. Three landmark titles—*Ocarina of Time* (1998), *The Wind Waker* (2002), and *Breath of the Wild* (2017)—were evaluated using standardized observations collected at three gameplay stages per title. Raw measurements were normalized and combined into an Environmental Coherence Index (ECI). The results demonstrate a progression from scripted finite-state behaviors toward systemic, environmentally responsive AI. A sensitivity analysis further revealed a biphasic evolutionary regime, in which early gains were dominated by refinements in navigation, interaction, and behavioral believability, whereas later gains were driven primarily by architectural innovations involving environmental simulation and emergent behavior. Although the proposed framework does not directly measure player immersion, it provides a reproducible quantitative methodology for comparing NPC AI across games and establishes a foundation for future quantitative studies of game AI evolution and immersion.

Keywords

Non-Player Character AI, Game artificial intelligence, Open-world games, NPC behavior, Player immersion, Believability, Environmental Coherence Index, Comparative game analysis, Behavioral metrics, Video game design

Author: Alonso J. Esboña, Colegio San Pedro, Calle Huron 409, La Molina, Lima, 15026, Peru.

alonso.esbona@gmail.com

Mentor: Rebecca Larkin, University of California, 510 Aldrich Hall, Irvine, CA 92697, USA. larkinr@uci.edu

1. Introduction

For over four decades, *The Legend of Zelda* franchise has continued to shape the framework of an adventure open-world video game, redefining their formula to look fresh by the release of the next title. This series, which is celebrating its 40th anniversary on February 21st, 2026, is one of the best examples in showing how technological developments have improved player experience over time. One of these developments is the creation of Artificial Intelligence (AI) for Non-Playable Characters (NPC), which has been crucial to the franchise's ability to create immersive settings.

NPC are fundamental to how players engage with virtual environments. Their movement, decision-making, and reaction are key factors whether a player feels the game world to be alive or to be artificial. This research investigates how the *Zelda* franchise demonstrates the evolution of open-world immersiveness through NPC AI. This is accomplished by analyzing three key titles in *The Legend of Zelda* franchise: *Ocarina of Time*, *The Wind Waker*, and *Breath of the Wild*. This paper studies how advances in NPC behaviors help build immersion, believability, and player engagement approaching these titles with a comparative mindset.

While Millington (1) has surveyed the general evolution of game AI architectures, and Jennett et al. (2) and Warpefelt and Verhagen (3) have developed theoretical frameworks for immersion and NPC believability respectively, no study has applied a structured comparative framework to a single long-running franchise to isolate how NPC architectural shifts link with

immersion outcomes. This paper addresses this gap by proposing a seven-dimension evaluation rubric (see Table 1a) applied to three temporally distinct *Zelda* titles. The research question is: To what extent does the architectural shift from scripted FSMs to systemic AI explain perceived immersion gains in an open-world franchise? This contribution has two main advantages: that it uses a reusable framework to analyze a long-running franchise that other researchers may use and apply to other franchises, and that it focuses more on how systemic advances contribute to coherences.

2. Methods

This study employed a structured comparative case-study methodology. Three inclusion criteria were determinative in selecting three titles from *The Legend of Zelda* franchise: First, each represented a distinct console generation, second, each held a Metacritic critical aggregate score >95, qualifying them as franchise-defining entries, and third, the temporal spacing between releases (1998, 2002, and 2017) permitted observation of generational AI development. Each title was evaluated across seven analytical dimensions such as NPC decision architecture, world simulation, movement and navigation model, and believability mechanism (see Table 1a). The analysis was comparative and illustrative rather than systematic. No formal risk-of-bias appraisal was conducted; instead, each dimension was measured through the observational surrogate measures described below.

NPCs play a central role in shaping the player's experience, particularly in open-world video games, where believable interactions contribute

to the perception of a living virtual world (1,4). Advances in AI have progressively transformed NPCs from rigid, scripted entities into adaptive agents capable of responding dynamically to players and their environments. Examining the evolution of these AI systems provides a foundation for systematically evaluating how *The Legend of Zelda* franchise has advanced NPC design and how these technological developments have contributed to increasingly immersive open-world experiences.

2.1 Background: NPC AI architectures

Before the mid-2000s, non-player characters (NPCs) in video games commonly relied on Finite State Machines (FSMs) as their primary decision-making architecture. According to Al Shehabi (4), FSMs represent behavior as a collection of discrete states, with transitions governed by predefined conditions encoded by the developer. This architecture is computationally efficient, making it particularly well suited for hardware-constrained systems. However, because NPC behavior is restricted to predetermined state transitions, FSM-based agents cannot reason beyond their scripted logic, resulting in predictable and repetitive behavioral patterns (1). Consequently, while FSMs provided reliable and efficient control, they limited the adaptability and perceived autonomy of NPCs, reducing their contribution to player immersion.

To overcome these limitations, developers increasingly adopted Behavior Trees (BTs) and Goal-Oriented Action Planning (GOAP). BTs organize NPC decision-making hierarchically, enabling agents to evaluate environmental conditions and select actions dynamically rather than relying solely on predefined state

transitions (1). According to Jeff Orkin (5), GOAP further extends NPC flexibility by allowing agents to formulate sequences of actions based on goals and the current state of the world instead of executing fixed scripts. Together, these architectures enable NPCs to respond more naturally to player actions and environmental changes, producing behaviors that appear more adaptive and context-sensitive. Such capabilities are particularly important in open-world games, where unpredictable player interactions frequently expose the limitations of rigid, scripted AI systems based on finite-state machines.

NPC movement systems evolved alongside advances in decision-making architectures. Earlier games frequently relied on waypoint-based navigation, in which NPCs traversed predefined routes between fixed locations. Although computationally efficient, this approach offered limited adaptability and, like finite-state machines, could reduce immersion when players or environmental changes disrupted scripted movement. According to Ian Millington (1), modern games increasingly employ navigation meshes (navmeshes), which represent traversable environments as interconnected polygons, allowing NPCs to calculate paths dynamically and avoid obstacles in real time. By enabling more flexible, context-aware navigation, navmeshes substantially improved the believability of NPC movement and became a foundational component of contemporary open-world AI systems.

Collectively, these technological advances have been major contributors to three concepts that underpin the present study: immersion, believability, and player engagement.

According to Jennett et al. (2), *immersion* refers to the player's sense of being mentally and emotionally absorbed in a virtual world.

Warpefelt and Verhagen (3) define *believability* as the extent to which NPC behaviors align with players' expectations of how a character should act within the context of the game world. Together, immersion and believability promote player engagement by fostering experiences that are perceived as coherent, responsive, and authentic. These concepts therefore provide a theoretical foundation for evaluating how

advances in NPC AI contribute to increasingly immersive open-world experiences.

2.2 Operationalizing the seven dimensions

The internal AI source code of these titles has not been publicly documented, hence the algorithms could not be measured directly. Each dimension was operationalized as the observable behaviors of village/town NPCs. The seven measures from which NPCs are judged are summarized in Table 1a.

Table 1a. Operational definitions and measurement criteria for the seven NPC AI dimensions

Dimension.name	Ceiling	Surrogate measure (what was counted)
D1. Decision Architecture	5	Number of distinct context-dependent responses observed to the same NPC approach
D2. World Simulation Scope	6	Number of environmental variables (of six tracked) that visibly influenced NPC behavior
D3. Movement & Navigation	1	Proportion of obstructed-path encounters in which the NPC successfully rerouted
D4. Emergent Behavior	6	Rate of unscripted emergent events per hour of observation
D5. AI Complexity Emphasis	7	Number of independent systems simultaneously influencing a single NPC
D6. Player–NPC Interaction	1	Proportion of a fixed set of player actions that elicited a unique NPC response
D7. Believability Mechanism	5	Mean appropriateness rating (1–5 scale) of NPC behavior given context

2.3 Sampling protocol

Observation was performed at three sampling points per title: early game (10%, first village), mid game (50%) and late game (90%). At each point a fixed 30-minute observation window was used. As mentioned before, ordinary town/village NPCs were the focus of the observation across the three titles. For this kind of NPCs, navigation was judged over the obstructed-path encounters. Interaction was measured with a set of five player actions (approach, draw weapon, initiate dialogue, jump

nearby, remain idle). Gathering all data together, believability was rated on a 1-5 scale.

2.4 Data collection

Due to time and hardware constraints, measurements were obtained from publicly available, continuous non-commentary gameplay videos rather than live play. Footage provided is re-observable records that any other researcher can verify against the same source, supporting reproducibility. All measurements were processed by a single observer (the

author). The three gameplay videos used as data sources for all the titles discussed are listed in the references (6-8)

2.5 Normalization and the Environmental Coherence Index (ECI)

Because the seven dimensions were measured on different scales, each observed value was normalized to a 0–1 range by dividing it by a predefined theoretical ceiling. These ceilings (Table 1b) were established *a priori* as part of the measurement instrument rather than derived from the observed data, ensuring that the normalization remained independent of the specific three-game sample analyzed.

Each ceiling was defined during construction of the measurement instrument and before any data were collected. The ceilings fall into three groups. For D3 and D6 the measure is a proportion — successful reroutes divided by obstructed-path encounters, and unique responses divided by the number of player actions tested — hence the ceiling is 1.0 by mathematical construction and involves no judgment. For D2, D5 and D7 the ceiling is fixed

by the instrument: the observation checklist tracked six environmental variables, so a title in which all six influenced NPC behavior would reach 6; seven independent NPC-influencing systems were tracked (schedule, dialogue, weather, danger, pathfinding, physics and social interaction), giving a ceiling of 7; and the believability rating scale ran from 1 to 5, giving a ceiling of 5. Only D1 and D4 required a judgment made in advance. For D1, five distinct context-dependent responses to a single approach was taken as the point at which an NPC varies its behavior across all the contextual factors recorded (time of day, weather, nearby threat, location and social situation). For D4, six emergent events per hour was taken as the equivalent of roughly one unscripted event every ten minutes, treated as the threshold for a densely systemic world. Because only these two ceilings involve judgment, and because the sensitivity analysis reported above produced the same ranking of the three titles under raw averaging and under min–max normalization, the reported ordering is unlikely to be an artifact of the ceiling values chosen.

Table 1b. Ceiling derivations for the seven NPC AI dimensions

Dimension.name	Ceiling	How the Ceiling was derived
D1. Decision Architecture	5	A priori judgment: the number of contextual factors recorded (time of day, weather, nearby threat, location, social situation); five distinct responses indicates variation across all of them.
D2. World Simulation Scope	6	Fixed by the instrument: six environmental variables were tracked in the observation checklist.
D3. Movement & Navigation	1	Mathematical bound: the measure is a proportion (successful reroutes / obstructed-path encounters).
D4. Emergent Behavior	6	A priori judgment: six events per hour corresponds to roughly one unscripted event every ten minutes, taken as a densely systemic world.
D5. AI Complexity Emphasis	7	Fixed by the instrument: seven independent NPC-influencing systems were tracked (schedule, dialogue, weather, danger, pathfinding, physics, social interaction).
D6. Player–NPC Interaction	1	Mathematical bound: the measure is a proportion (unique responses / player actions tested).
D7. Believability Mechanism	5	Fixed by the instrument: the appropriateness rating scale ranged from 1 to 5.

The comparative framework was based on five methodological assumptions. First, all seven dimensions were assigned equal weight in the ECI because no empirical evidence justified differential weighting. Second, the theoretical ceilings were specified *a priori* during development of the measurement instrument. Third, the recorded gameplay footage was assumed to provide a representative sample of typical NPC behavior, with the standardized 30-minute observation period helping to reduce the effects of selective sampling. Fourth, the ECI was intended to quantify observable NPC behavioral coherence rather than immersion itself; the latter would require independent validation through player-centered studies. Fifth, the selected observation periods were assumed to capture sufficient in-game variation to support meaningful comparative analysis.

To evaluate the robustness of the proposed normalization procedure, the ECI was also calculated using two alternative methods: (i) simple averaging of the unnormalized dimension values (1.19, 1.44, and 3.42) and (ii) min-max normalization across the three games (0.05, 0.22, and 1.00). In all three approaches, the comparative ranking remained unchanged (*Ocarina of Time* < *The Wind Waker* < *Breath of the Wild*), indicating that the principal findings were not dependent on the choice of normalization scheme or the predefined theoretical ceilings. All gameplay videos analyzed in this study were publicly available and were used solely as observational sources. No video content or footage was reproduced or redistributed.

3. Results

3.1 Case study 1: *Ocarina of Time*

1998 was a significant year for video games, many of those who say that feel that way because that was the year when *The Legend of Zelda: Ocarina of Time* was released for the Nintendo 64. It is widely considered to be a turning point in the evolution of 3D video games. While *Ocarina of Time* is widely recognized for innovations such as camera control and combat mechanics, it also established foundational approaches to NPC behavior, movement, and time-based routines that influenced subsequent open-world game design (9). *Ocarina of Time* was released at a time when hardware limitations of the Nintendo 64, such as limited processor performance and cartridge-limited storage, were a crucial factor considering the potential of a videogame, but it carefully designed AI systems to create illusions of a living and vivid world, even though NPC behavior remained rigid. By the examination of these systems, the analysis indicates that a comprehension of how decisions were made to balance immersion and control was reached.

In *The Legend of Zelda: Ocarina of Time*, NPC movement and behavior were governed primarily by FSMs supplemented by environment-sensitive scripts. NPCs transitioned between predefined behavioral states when specific conditions were satisfied, such as changes in time of day or player interactions. For example, villagers followed daily routines synchronized with the game's day–night cycle, creating the impression of a living world. However, these behaviors remained constrained by predefined scripts rather than adaptive decision-making. NPCs did not reason about their environment or dynamically modify their actions beyond their

programmed state transitions, resulting in predictable behavioral patterns despite their apparent realism (1).

Enemy NPCs further demonstrate how designers utilized FSM to help the player experience. Hostile characters were designed to have a restrained position when fighting. That meant hostile characters were only designed to attack one at a time (8). However, the game's environments were designed to minimize such situations. This illustrates how early use of AI could work with level design to keep immersion.

An advanced feature for its time was the inclusion of horseback combat with pursuit mechanics. The pursuit mechanics required NPCs to track and chase the player while the player was still mounted. Given the hardware constraints and computational resources available at the time, implementing such advanced AI systems represented a substantial technical challenge. Even though these movements were still largely scripted, they represented an early attempt to integrate movement, combat, and spatial awareness into NPC AI. This technology contributed to the video game world because it raised awareness of how NPCs could engage the player dynamically within a three-dimensional space (9).

Despite its innovations at the time, NPC behavior in *Ocarina of Time* remained deterministic and very predictable. The characters often repeated the same routine across the normal playthrough and barely reacted meaningfully to unscripted player actions or even environmental changes. According to Al Shehabi, from a modern AI perspective, these limitations reduced perceived

agency and realism (10). But at its time, it was more than enough to make the player feel like a part of the world.

3.2 Case study 2: *Wind Waker*

Released for the Nintendo GameCube in 2002, *The Legend of Zelda: The Wind Waker* marked a step toward the right direction in open-world immersion within *The Legend of Zelda* franchise and video games in general. Given the fact that this game was also released in an era where complex AI features were lacking, this game nevertheless found a way to enhance the world's "liveliness".

The Wind Waker was really similar to *Ocarina of Time*, it followed a similar script for their NPCs by using FSMs and event-based scripts. Nevertheless, this game expanded on immersion through schedule-based behaviors, which gave NPCs clearer daily patterns. Although these routines did not involve sophisticated real-time decision-making, they introduced a structured behavioral framework that established a strong foundation for the more advanced NPC AI systems developed in later entries of the franchise.

The main appeal of this game was the significant innovation in the expressiveness of all characters, including the protagonist and NPCs (7). Across the sampled windows, characters displayed a uniquely expressive face, compared to the other titles. This made the player perceive these characters as "behaviorally complex" and emotionally deep. This was achieved by overreactions such as exaggerated facial expressions, which communicated moods.

The Wind Waker also developed a more coherent relationship between the world scale and NPC behavior (7). The inclusion of islands instead of a pangea-like terrain presented developers with a defined radius in which NPCs could work. This made them more effective and their lack of pathfinding less noticeable.

3.3 Case study 3: *Breath of the Wild*

Since its release in 2017, *The Legend of Zelda: Breath of the Wild* represented a major change in how open-world games used AI to produce NPC behaviors similar to humans and world immersion. Comparing it to earlier *Zelda* titles that used scripted routines and fixed patterns, the results indicate that *Breath of the Wild* integrated AI systems into a world so vivid that it could sustain extended player engagement.

Advances in computing hardware and AI architectures enabled *The Legend of Zelda: Breath of the Wild* to implement substantially more sophisticated NPC systems than earlier entries in the series. BTs and GOAP provided flexible decision-making frameworks that allowed NPCs to select actions based on changing environmental conditions and contextual information rather than relying solely on predefined scripts. As a result, NPCs could interrupt their routines to respond dynamically to events in the game world. For example, a villager might travel to work in the morning, seek shelter during rainfall, flee from nearby monsters, and return home at night. Such context-sensitive behaviors enhanced the believability of NPCs and contributed to the perception of a responsive, living world, thereby increasing player immersion compared with rigid scripted systems.

A major advancement over earlier titles was the adoption of modern pathfinding techniques based on navigation meshes (navmeshes). *Breath of the Wild* appears to employ navmesh-based navigation in combination with search algorithms such as A*, enabling NPCs to generate smooth, obstacle-aware paths in real time (1). Consequently, NPCs navigated around obstacles and dynamically rerouted when their intended paths became blocked. Within the observation windows analyzed, rerouting proportion and smooth obstacle avoidance each received the maximum normalized score of 1.00, consistent with Millington's description of navmesh-based pathfinding (1).

Wildlife exhibited similar adaptive behavior, responding to nearby threats by fleeing along dynamically generated escape paths that resembled natural animal movement. Such context-sensitive navigation further enhanced the perceived realism of the game world.

Another major contributor to *Breath of the Wild's* immersion was the integration of AI behavior with the game's systemic environmental simulation. NPCs do not operate independently of the environment; instead, their behaviors are influenced by weather, physics, and other world mechanics. For example, rain creates slippery surfaces, wind affects projectile trajectories, and lightning preferentially strikes metal objects (11). NPCs were programmed to recognize and respond to these changing environmental conditions. During thunderstorms, for instance, enemies may discard metal weapons to avoid lightning strikes, producing behavior that appeared context-aware and enhanced the believability of the game world.

The integration of advanced pathfinding, adaptive decision-making, and systemic environmental interactions produced emergent behaviors that were not feasible in earlier *The Legend of Zelda* titles because of the hardware and AI limitations discussed previously.

Despite its substantial advancements, *Breath of the Wild* also illustrates the practical constraints associated with implementing complex AI systems on limited hardware. Released for the Nintendo Switch, whose computational resources were more limited than those of the contemporary PlayStation 4 and Xbox One, the game employed several optimization techniques to maintain performance. One such technique was distance-based AI culling, in which NPC behaviors were suspended or simplified when characters were sufficiently far from the player (11). By updating only NPCs within an active simulation radius, the game reduced computational demands while preserving the illusion of a persistent, living world. Simulating thousands of interacting environmental and AI processes simultaneously would have exceeded the available hardware resources, making such optimizations essential. These design choices demonstrate how sophisticated NPC behavior can be achieved through careful balancing of AI complexity and computational efficiency.

3.4 Overview

The results of measuring the seven dimensions across the three games are summarized in Table 2 and Figure 1. The normalized scores increased from *Ocarina of Time* (OoT) to *The Wind Waker* (WW) to *Breath of the Wild* (BotW) as expected.

The seven dimensions revealed a consistent progression in NPC behavioral sophistication

across the three games, although the magnitude of improvement differed substantially between dimensions (Figure 1).

D1 (context-dependent responses) remained relatively stable across the series, with *Ocarina of Time* (0.47) slightly outperforming *The Wind Waker* (0.40) before increasing to 0.60 in *Breath of the Wild*. Unlike the other dimensions, this measure showed no monotonic improvement, suggesting that context-dependent behavioral branching was already present in early titles but was not substantially expanded until *Breath of the Wild*.

In contrast, D2 (environmental influence on NPC behavior) exhibited the greatest change. Normalized scores increased from 0.11 (*Ocarina of Time*) to 0.17 (*The Wind Waker*), before rising sharply to 0.94 in *Breath of the Wild*. This represents the largest improvement among all seven dimensions, indicating that environmental responsiveness became a defining characteristic of the most recent title.

D3 (dynamic pathfinding) showed a progressive increase across the series (0.33, 0.89, and 1.00), demonstrating steady improvements in NPC navigation and obstacle avoidance rather than an abrupt technological transition.

D4 (unscripted emergent events) clearly differentiated *Breath of the Wild* from the earlier titles. No emergent interactions were observed during the sampling periods for either *Ocarina of Time* or *The Wind Waker*, whereas *Breath of the Wild* achieved a normalized score of 0.44. This was the only dimension in which the earlier games received identical minimum values.

D5 (responses to simultaneous influences) also increased substantially across the series, from 0.38 to 0.43 and finally to 0.81, indicating that later NPCs increasingly responded to multiple concurrent environmental and gameplay factors instead of isolated triggers.

D6 (responses to different player actions) followed a similar trend, increasing from 0.33 in Ocarina of Time to 0.53 in The Wind Waker and 0.93 in Breath of the Wild. This suggests a

progressive expansion in the variety of player interactions capable of altering NPC behavior.

Finally, D7 (behavioral appropriateness) increased from 0.40 to 0.53 and reached the maximum normalized value of 1.00 in Breath of the Wild. Although this dimension relied on subjective assessment, it followed the same overall trajectory observed for the more objective behavioral measures.

Table 2. Raw and normalized dimension scores and resulting Environmental Coherence Index ECI) (normalized 0-1)

Dimension.name	OoT raw	WW raw	BotW raw	OoT norm	WW norm	BotW norm
D1. Decision Architecture	2.33	2.00	3.00	0.47	0.40	0.60
D2. World Simulation Scope	0.67	1.00	5.67	0.11	0.17	0.94
D3. Movement & Navigation	0.33	0.89	1.00	0.33	0.89	1.00
D4. Emergent Behavior	0.00	0.00	2.67	0.00	0.00	0.44
D5. AI Complexity Emphasis	2.67	3.00	5.67	0.38	0.43	0.81
D6. Player–NPC Interaction	0.33	0.53	0.93	0.33	0.53	0.93
D7. Believability Mechanism	2.00	2.67	5.00	0.40	0.53	1.00
Environmental Coherence Index (0–10)				2.89	4.22	8.19

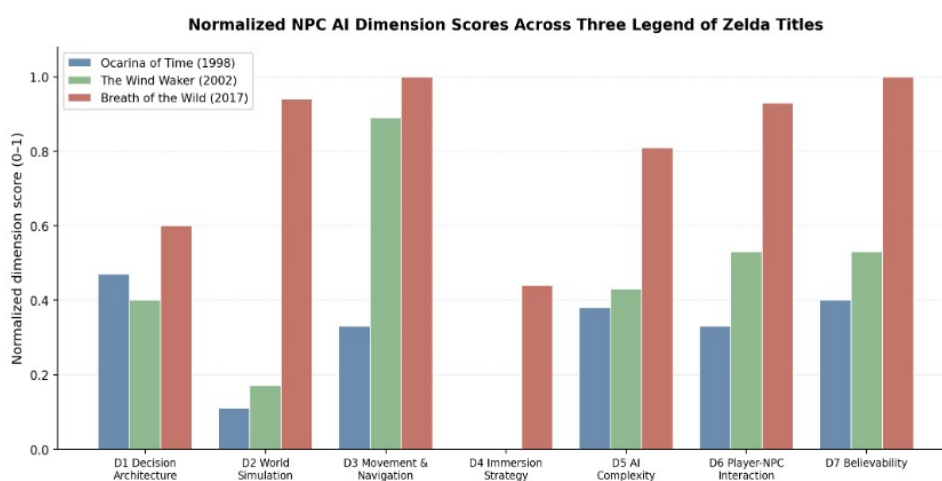


Figure 1. Normalized(0-1) NPC AI dimension scores across the three titles.

The ECI synthesized these seven dimensions into a single comparative metric. The index increased monotonically from 2.89 for *Ocarina of Time* to 4.22 for *The Wind Waker* and 8.19 for *Breath of the Wild* (Figure 2). While *The Wind Waker* demonstrated modest improvements over *Ocarina of Time*, the largest increase occurred between *The Wind Waker* and *Breath of the Wild*. Notably, improvements were not uniform across all dimensions. Behavioral systems associated with environmental responsiveness, emergent interactions, and multi-factor decision-making exhibited substantially larger gains than context-dependent branching, indicating that the evolution of NPC AI was driven primarily by increasing integration between NPC behavior and the surrounding game world rather than by incremental refinement of existing scripted behaviors.

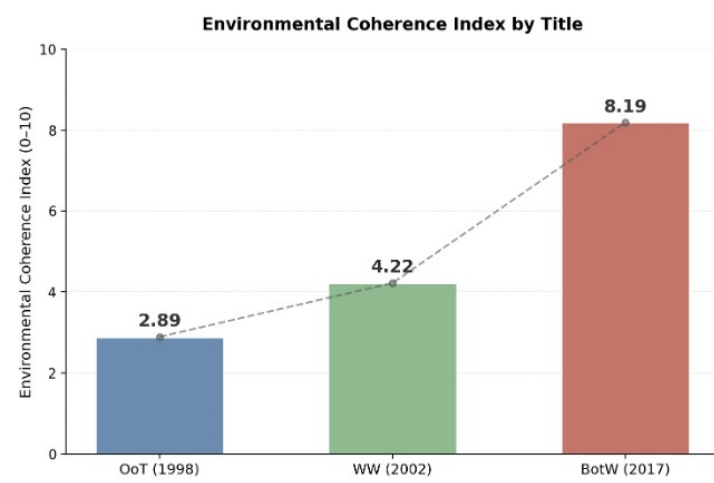


Figure 2. Environmental Coherence Index (0-10) by title.

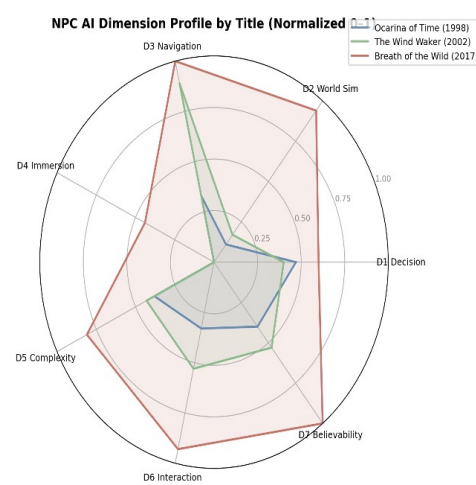


Figure 3. NPC AI dimension profile by title (normalized 0-1).

Figure 3 reveals that the three titles exhibit distinct NPC AI profiles rather than simply differing in overall performance. *Ocarina of Time* displays a relatively balanced but uniformly low-dimensional profile, with no single behavioral capability substantially exceeding the others. Its strongest dimensions are Decision Architecture (D1) and Believability Mechanism (D7), whereas World Simulation Scope (D2), Movement & Navigation (D3), and Emergent Behavior (D4) remain comparatively limited. This suggests that early NPC systems relied primarily on scripted behavioral organization rather than extensive interaction with the surrounding game world.

The Wind Waker retains a profile similar to *Ocarina of Time*, although modest improvements are observed in navigation (D3), player interaction (D6), and believability (D7). Notably, the overall shape of the profile changes only slightly despite an increase in the Environmental Coherence Index, indicating that the title represents an incremental refinement of existing NPC behaviors rather than the introduction of fundamentally new AI capabilities. The complete absence of observed emergent interactions (D4) further distinguishes it from *Breath of the Wild*.

In contrast, *Breath of the Wild* exhibits a markedly different behavioral profile. High scores are observed across nearly every dimension, with particularly large gains in World Simulation Scope (D2), Movement & Navigation (D3), AI Complexity Emphasis (D5), Player–NPC Interaction (D6), and Believability Mechanism (D7). Rather than improving a single aspect of NPC behavior, the

game demonstrates simultaneous advancement across multiple behavioral domains. The only dimension that remains noticeably below the maximum is Immersion Strategy (D4), reflecting that unscripted emergent interactions, although present, occur less frequently than other observable NPC behaviors within the sampled gameplay.

The comparison also shows that NPC AI evolution was not uniform across behavioral domains. Decision Architecture (D1) changed relatively little over the series, whereas World Simulation Scope (D2), AI Complexity Emphasis (D5), and Player–NPC Interaction (D6) experienced substantially larger improvements. Consequently, the principal differences between early and modern NPC systems arise less from changes in basic decision logic than from increasing integration of NPCs with environmental simulation, navigation, and interactive world systems.

4. Discussion

The quantitative results indicate that *Breath of the Wild* represents a substantial advancement in NPC AI compared with the earlier titles examined. Rather than relying primarily on scripted behaviors, the game integrated dynamic decision-making, navigation meshes, environmental responsiveness, and system-level interactions to produce NPC behaviors that were more adaptive and context-sensitive (1,11). This integration was reflected by consistently higher scores across nearly all seven behavioral dimensions, particularly those related to environmental simulation, navigation, and multi-factor decision-making. Together, these systems created a more responsive and coherent game world, supporting the perception of a

living environment in which NPCs interacted not only with the player but also with the broader game ecosystem. The findings therefore suggest that advances in NPC AI have played a significant role in the evolution of open-world game design, with *Breath of the Wild* representing the culmination of a progression from isolated scripted behaviors toward increasingly integrated and systemic AI architectures.

The Legend of Zelda: *The Wind Waker* represents a transitional stage in the evolution of NPC AI. Although its NPCs continued to rely primarily on FSMs, the game enhanced player immersion through expressive animations, consistent daily routines, and coherent visual presentation rather than through substantial advances in decision-making complexity. The quantitative results support this interpretation: while improvements were observed in dimensions such as navigation, player interaction, and believability, environmental responsiveness and emergent interactions remained limited. These findings suggest that perceived intelligence and immersion can be enhanced not only through increasingly sophisticated AI algorithms but also through the effective presentation and consistency of NPC behaviors. At the same time, the restricted adaptability of NPCs highlights the limitations of predominantly scripted systems, reinforcing the need for the more flexible, systemic AI architectures that were later realized in *Breath of the Wild* and subsequently adopted in many modern open-world games.

The Legend of Zelda: *Ocarina of Time* established the foundation for NPC AI within the franchise and significantly influenced

subsequent open-world game design. By combining FSM-based decision-making, scripted movement, and time-based behavioral routines, the game created NPCs that appeared purposeful and contributed to a convincing game world despite the hardware limitations of its era (1,8,9). Although these behaviors remained largely predetermined and exhibited limited adaptability, they represented a substantial advance over earlier approaches to NPC design. The game's influence has been widely recognized, consistently ranking among the most influential video games of all time and receiving critical acclaim upon release. IGN described it as "a game that comes along once a decade" (9,12). Within the context of this study, *Ocarina of Time* serves as the baseline from which the subsequent evolution toward increasingly adaptive and systemic NPC AI can be quantitatively evaluated.

4.1 Comparative analysis

The evolution of the *The Legend of Zelda* franchise provides a quantitative illustration of how NPC AI has progressed over nearly two decades. Comparison of *Ocarina of Time* (1998), *The Wind Waker* (2002), and *Breath of the Wild* (2017) demonstrates a consistent increase in the complexity, responsiveness, and environmental integration of NPC behaviors. The corresponding increase in the Environmental Coherence Index suggests that successive titles progressively created game worlds in which NPC actions became more coherent with the surrounding environment and gameplay systems.

Movement and navigation represent one of the clearest examples of this progression. Both *Ocarina of Time* and *The Wind Waker* rely

primarily on waypoint-based navigation, resulting in predetermined movement patterns with limited adaptability. In contrast, *Breath of the Wild* employs navigation meshes (navmeshes), enabling NPCs to generate dynamic, obstacle-aware paths through complex environments (1). This architectural transition is reflected quantitatively by the Movement and Navigation dimension, which increased from 0.33 in *Ocarina of Time* to 0.89 in *The Wind Waker* and reached the maximum normalized value of 1.00 in *Breath of the Wild* (Table 2). The findings indicate that navmesh-based navigation is particularly well suited to large open-world environments, where flexible movement contributes to the perception of a coherent and believable world.

Table 3. Comparison of NPC behaviors in the legend of Zelda games

Core Dimension	Ocarina of Time (1998)	The Wind Waker (2002)	Breath of the Wild (2017)
NPC decision architecture	Primarily scripted Finite-State Style Behaviors with deterministic transitions tied to triggers and schedules	Expanded scripted logic with contextual decision layers and personality-driven routines	System-oriented decision logic integrated with environmental simulation and adaptive priority systems
World Simulation Scope	Limited simulation; NPC behavior is event-driven and localized	Moderate systemic interaction with time-based and situational NPC reactions	High systemic simulation where NPCs respond to physics, weather hazards, and world state changes
Movement & Navigation Model	Fixed Waypoint navigation and path scripting within constrained spaces	Enhanced spatial routing supporting expressive movement in towns and island	Navigation Meshes and terrain-aware movement integrated with physics-driven environments
Emergent Behavior	Structured routines and spatial storytelling that imply a living world	Character expressiveness and context interactions that reinforce narrative presence	Emergent systems producing unscripted scenarios that stimulate ecological and social responsiveness
AI Complexity Emphasis	Efficiency and reliability under hardware limitations	Behavioral richness balanced with console constraints	Scalability, systemic interaction, and performance optimization to support large-scale simulation
Player-NPC interaction Model	Scripted dialogue and event-triggered interactions	Context-sensitive dialogue and emotional expressiveness	Dynamic reactions to player actions, threats, and environmental stimuli
Performance Optimization Approach	Lightweight scripted AI designed for constrained processing capacity	Balanced scripting depth to maintain responsiveness	Distance-based AI scaling and simulation prioritization for open-world performance management
Believability Mechanism	Predictable routines simulating daily activity patterns	Expressive behavior reinforcing character identity	Emergent behavior systems creating reactive, believable ecosystems

The greatest improvements were observed in World Simulation Scope and Emergent Behavior, indicating that the principal advances in NPC AI were not simply improvements in individual decision-making algorithms but a broader integration of NPC behavior with

environmental systems. Earlier titles relied primarily on scripted interactions, whereas *Breath of the Wild* increasingly embeds NPCs within interconnected gameplay systems that respond dynamically to environmental conditions and player actions.

Taken together, these findings suggest that the evolution of NPC AI within the *The Legend of Zelda* franchise reflects a transition from isolated scripted behaviors toward integrated systemic AI. While narrative design, visual presentation, and artistic direction remain essential contributors to player immersion, the quantitative framework developed in this study indicates that increasingly adaptive and environmentally coherent NPC behaviors constitute an additional mechanism through which modern open-world games create more believable virtual worlds.

4.2 Evolutionary trajectory

The pattern of dimensional changes suggests that NPC AI evolved in two distinct stages rather than through continuous improvement. Between *Ocarina of Time* and *The Wind Waker*, gains were concentrated in Movement & Navigation (D3), AI Complexity Emphasis (D5), Player–NPC Interaction (D6), and Believability (D7) (Figure 3). These dimensions primarily reflect refinements in how NPC behaviors are presented to the player and could be enhanced within a predominantly finite-state-machine framework. The relatively unchanged scores for Decision Architecture (D1), World Simulation Scope (D2), and Emergent Behavior (D4) indicate that the underlying behavioral architecture remained largely unchanged despite these improvements.

The transition to *Breath of the Wild* represents a fundamentally different stage of evolution. Rather than simply refining scripted behaviors, the game introduces substantial improvements in the dimensions that require architectural integration—environmental responsiveness, systemic world simulation, and emergent interactions. These dimensions exhibit the largest gains observed in the study, suggesting that advances in NPC AI were driven less by incremental increases in behavioral complexity and more by embedding NPCs within interconnected environmental systems.

An alternative evolutionary trajectory can be inferred from the data. Had the underlying AI architecture remained predominantly scripted, continued improvements would likely have occurred in dimensions such as movement quality, player interaction, behavioral polish, and perceived believability, as observed between *Ocarina of Time* and *The Wind Waker*. In contrast, dimensions dependent on systemic integration—Decision Architecture, World Simulation Scope, and Emergent Behavior—would have been expected to plateau. The abrupt increases observed in these latter dimensions for *Breath of the Wild* therefore suggest that incremental refinement alone was insufficient; a qualitative shift in AI architecture was required to enable the next stage of NPC evolution.

A useful perspective emerges by comparing the relative increases in the original ECI with those obtained after recalculating the index while excluding Decision Architecture (D1), World Simulation Scope (D2), and Emergent Behavior (D4) (Recalculating Table 2). Using all seven dimensions, the ECI increases by ~ 46% from

Ocarina of Time to *The Wind Waker* and 94% from *The Wind Waker* to *Breath of the Wild*. However, when only Movement & Navigation (D3), AI Complexity Emphasis (D5), Player–NPC Interaction (D6), and Believability (D7) are considered, the pattern reverses: the increase becomes ~ 65% between *Ocarina of Time* and *The Wind Waker*, followed by only 57% between *The Wind Waker* and *Breath of the Wild*. This reversal reveals a biphasic evolutionary regime in NPC AI development. The first phase is characterized primarily by incremental refinement of scripted behaviors—including navigation, player interaction, behavioral richness, and believability—which could be achieved within a predominantly FSM framework. The second phase, represented by *Breath of the Wild*, is distinguished by the introduction of new architectural capabilities—environmentally responsive decision-making, broader world simulation, and emergent behavior—that fundamentally expand the scope of NPC AI. Thus, the markedly larger increase observed in the original ECI between *The Wind Waker* and *Breath of the Wild* reflects not merely continued refinement of existing systems but a transition to a qualitatively different architectural regime of NPC intelligence.

4.3 Alternative explanations

The observed increase in the Environmental Coherence Index values is unlikely to be explained solely by advances in NPC AI. Improvements in hardware capability, graphics, animation quality, audio design, world scale, quest structure, and overall game design also contribute to how players perceive immersion across successive entries in *The Legend of Zelda* franchise. Consequently, later titles may appear

more immersive for reasons extending beyond NPC behavior alone.

The objective of the present study, however, was not to measure overall game quality or immersion directly, but rather to quantify observable aspects of NPC behavior using a standardized framework. The seven analytical dimensions were intentionally selected to capture behavioral characteristics that are largely independent of visual presentation. For example, whether an NPC reroutes around an obstacle, responds to a drawn weapon, interrupts a routine because of changing weather, or reacts differently to multiple player actions can be evaluated regardless of graphical fidelity or artistic style. These behavioral measures therefore isolate observable properties of NPC AI rather than aesthetic presentation.

Nevertheless, interactions between AI systems and other components of game design cannot be completely separated. Improved hardware enables more sophisticated AI architectures, while world design and narrative structure influence how NPC behaviors are perceived by players. Accordingly, the findings should be interpreted as evidence of measurable changes in NPC behavioral complexity rather than proof that NPC AI alone determines player immersion.

4.4 Limitations

This study has several limitations that should be considered when interpreting its findings. First, the analysis was limited to three landmark entries in *The Legend of Zelda* franchise. Although these titles were selected to represent distinct generations of NPC AI development, the results should be interpreted as a

comparative case study rather than a comprehensive evaluation of the entire franchise or of open-world games in general.

Second, the three games were developed for substantially different hardware platforms. Improvements in processor performance, memory capacity, and storage technology undoubtedly enabled increasingly sophisticated AI systems over time. Consequently, part of the observed progression reflects technological advances available to developers rather than design choices alone. The analysis therefore compares the implementation of NPC AI within the technological constraints of each console generation.

Third, some AI architectures implemented in earlier *Zelda* titles have not been publicly documented by Nintendo. As a result, descriptions of certain internal AI mechanisms are inferred from observable NPC behavior, developer documentation where available, and community technical analyses rather than direct examination of source code.

Fourth, data were collected from publicly available gameplay recordings rather than live gameplay. Although this approach provides a reproducible observational record, behaviors occurring outside the selected observation windows were not captured. Similarly, the three sampling points selected for each title were assumed to provide a representative overview of NPC behavior but cannot encompass every possible gameplay scenario.

Fifth, all observations and behavioral ratings were performed by a single observer. While the operational definitions and standardized

observation protocol were designed to improve consistency, inter-rater reliability was not assessed. Independent replication by multiple observers would provide additional evidence of the robustness of the proposed framework.

Sixth, the analysis was intentionally restricted to ordinary town and village NPCs evaluated using a standardized observation protocol. Consequently, the findings may not generalize to other NPC categories, including enemies, companions, merchants, bosses, or scripted quest characters, whose behavioral architectures may differ substantially.

Finally, the Environmental Coherence Index assigns equal weight to each of the seven analytical dimensions. This equal-weight approach provides a transparent and reproducible baseline for comparison, but it does not imply that each dimension contributes equally to perceived behavioral coherence or player immersion. Future studies may investigate alternative weighting strategies based on empirical player studies, expert evaluation, or statistical methods.

Despite these limitations, the proposed framework establishes a transparent, reproducible methodology for quantitatively comparing NPC AI behavior. The limitations primarily affect the generalizability and interpretation of the findings rather than the applicability of the framework itself, which can be refined, validated, and extended in future investigations.

4.5 Future research opportunities

Several opportunities exist to expand and refine the quantitative framework introduced in this

study. First, future work should evaluate its applicability across a broader range of game genres and franchises, including role-playing games, open-world action games, simulation games, and massively multiplayer online games. Such studies would determine whether the seven proposed dimensions are sufficiently general to characterize diverse NPC architectures or whether additional dimensions are required for genre-specific behaviors.

Second, the framework would benefit from formal validation studies. Independent observers could apply the methodology to the same gameplay segments to assess inter-rater reliability, while repeated analyses across multiple playthroughs could evaluate measurement stability. Statistical techniques such as factor analysis or principal component analysis may also help determine whether the proposed dimensions represent distinct constructs or whether some dimensions can be combined without loss of explanatory power.

Third, the ECI could be further developed through alternative weighting strategies. The present study assigns equal importance to each dimension to provide a transparent baseline; however, future research could investigate data-driven weighting methods based on player surveys, expert evaluation, or machine learning approaches. Comparing equal-weight, expert-weighted, and statistically derived indices would provide insight into the robustness and interpretability of the composite metric.

Future studies should also investigate the relationship between objective NPC behavior and subjective player experience. The current framework intentionally quantifies observable

AI behavior rather than immersion itself. Correlating ECI scores with validated immersion, presence, engagement, or player satisfaction questionnaires could determine whether improvements in measurable NPC behavior translate into meaningful differences in perceived gameplay experience.

Another promising direction is the incorporation of automated data collection. Advances in computer vision, gameplay telemetry, and AI-assisted video analysis could substantially reduce observer effort by automatically detecting NPC behaviors, environmental interactions, navigation decisions, and player responses. Automated measurement would improve scalability while reducing observer bias and enabling analysis of substantially larger gameplay datasets.

An interesting direction for future research would be to develop a quantitative surrogate for player immersion and determine whether it exhibits the same biphasic evolutionary regime (section 4.2) observed for the ECI. If a comparable pattern emerges, it would provide evidence that improvements in NPC AI ECI are reflected in the player's immersive experience. Conversely, if the immersion surrogate follows a different developmental trajectory, it would suggest that NPC AI ECI alone is insufficient to explain immersion and that additional latent variable(s), not captured by the present framework, contribute substantially to the perception of immersive virtual worlds. Identifying and characterizing such latent factors could lead to a more comprehensive quantitative model of immersion.

Finally, extending the framework to longitudinal studies across multiple decades of game development may provide a quantitative history of NPC AI evolution. Applying a consistent methodology to titles spanning different hardware generations, development studios, and AI paradigms would enable researchers to identify technological inflection points, compare design philosophies, and evaluate how advances in artificial intelligence have influenced the realism, adaptability, and coherence of virtual worlds over time.

5. Conclusion

This study introduced a structured quantitative framework for comparing NPC artificial intelligence across three landmark titles in *The Legend of Zelda* franchise using seven operationalized behavioral dimensions and a composite Environmental Coherence Index (ECI). Applying this framework to *Ocarina of Time*, *The Wind Waker*, and *Breath of the Wild* demonstrated a clear progression from predominantly scripted finite-state-machine behaviors toward increasingly systemic and environmentally responsive AI architectures. The results indicate that advances in NPC AI encompass not only improvements in decision-making but also greater environmental integration, adaptive navigation, richer player interactions, and increased behavioral coherence.

Beyond documenting this progression, the analysis revealed that NPC AI evolution did not occur through continuous incremental improvement. Sensitivity analysis of the ECI demonstrated a biphasic evolutionary regime. The first phase, represented by the transition from *Ocarina of Time* to *The Wind Waker*, was

dominated by refinements in scripted behaviors, including navigation, player interaction, behavioral richness, and believability. The second phase, represented by *Breath of the Wild*, was characterized by substantial advances in environmentally responsive decision-making, systemic world simulation, and emergent behavior, reflecting a qualitative transition in NPC AI architecture rather than simply further refinement of existing scripted systems. These findings suggest that major advances in believable NPC behavior arise not only through incremental improvements in behavioral complexity but also through architectural innovations that integrate NPCs more deeply with the simulated game world.

Although the proposed framework does not directly quantify player immersion, it provides a transparent and reproducible methodology for measuring observable NPC behavioral coherence. Future research should investigate whether quantitative surrogates of player immersion exhibit the same biphasic evolutionary regime identified for the ECI. If similar trajectories are observed, this would strengthen the relationship between NPC behavioral coherence and perceived immersion. Conversely, divergent trajectories would suggest that additional latent variables, not captured by the present framework, contribute substantially to immersive virtual worlds. Identifying and characterizing such factors would represent an important step toward a more comprehensive quantitative theory of immersion.

Beyond its application to *The Legend of Zelda*, the principal contribution of this work is the introduction of a reusable quantitative

methodology for evaluating NPC AI across game franchises and hardware generations. By transforming qualitative observations of NPC behavior into reproducible quantitative measures, this framework provides a foundation for future comparative studies of game AI and for understanding how advances in artificial intelligence contribute to the creation of increasingly believable virtual worlds.

Acknowledgements

I would like to express my profound gratitude to Rebecca Larkin for her contribution to the completion of this work. An AI assistant (Claude) was used for editorial support, including feedback on structure, language and grammar editing, spreadsheet and formula construction for data organization, and figure generation from author-collected data. All gameplay observations and data were collected by the author, who reviewed all content and takes full responsibility for it.

6. References

1. Millington, I. (2006). *Artificial intelligence for games*. Morgan Kaufmann.
2. Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9), 641–661. <https://doi.org/10.1016/j.ijhcs.2008.04.004>
3. Warpefelt, H., & Verhagen, H. (2017). A model of non-player character believability. *Journal of Gaming & Virtual Worlds*, 9(1), 39–53. https://doi.org/10.1386/jgvw.9.1.39_1
4. Feng, S., & Tan, A.-H. (2016). Towards autonomous behavior learning of non-player characters in games. *Expert Systems with Applications*, 56, 89–99. <https://doi.org/10.1016/j.eswa.2016.02.043>
5. Orkin, J. (2004). Applying goal-oriented action planning to games. In S. Rabin (Ed.), *AI game programming wisdom 2* (pp. 217–228). Charles River Media.
6. BeardBear. (2023). *Zelda: Breath of the Wild – Full game 100% walkthrough* [Video]. YouTube. https://www.youtube.com/watch?v=0b0TNce_9tc
7. Game Master TV. (2022, February 12). *The Legend of Zelda: The Wind Waker – Full game walkthrough / longplay (NGC) 1080p 60fps* [Video]. YouTube. https://www.youtube.com/watch?v=my7foys_pqk
8. SourceSpy91. (2020, May 3). *The Legend of Zelda: Ocarina of Time – Full game – No commentary* [Video]. YouTube. <https://www.youtube.com/watch?v=0ajivln8e68>

9. Ray, J. (2018, November 22). *Ocarina of Time's inescapable influence on modern gaming*. Fandom. <https://www.fandom.com/articles/ocarina-of-time-influence-gaming>
10. Al Shehabi, A. (2022). *Decision-making AI in digital games* (Bachelor's thesis, Uppsala University). DiVA Portal. <https://www.diva-portal.org/smash/get/diva2:1673140/FULLTEXT01.pdf>
11. Fujibayashi, H., Takizawa, S., & Dohta, T. (2017). Change and constant: Breaking conventions with “the legend of Zelda: Breath of the wild.” In *Gdcvault.com*. <https://www.gdcvault.com/play/1024562/Change-and-Constant-Breaking-Conventions>
12. Casamassina, M. (2000, August 15). *Legend of Zelda: Ocarina of Time review*. IGN64. <https://web.archive.org/web/20000815080728/http://ign64.ign.com/reviews/437.html>