Available Online: 9 October 2026 DOI: 10.54254/2753-7102/2026.37575

The obvious clue is not always the truth: evidence chains and misdirection in a marine pollution investigation game

Ziqi Ma

Burr and Burton Academy, Manchester, USA

ziqi0613@126.com

Abstract. In *Last 3 Days*, players investigate unusual marine animal deaths by examining samples and comparing test results with collection locations and reports. This paper analyses how the game's evidence chain develops from these comparisons, with particular attention to clues that suggest an incomplete explanation. Plastic inside a fish and an injury associated with a fishing net indicate harm to individual animals. Repeated radiation abnormalities across different samples raise a wider question about the pollution source. The company's testing report also requires careful interpretation because a normal result may not represent the affected area. These examples show how the design uses real but limited information to create misdirection. The final task asks players to select both a pollution source and supporting evidence, making the connection between their answer and its justification part of the game. The analysis concerns the prototype's design; its effect on players' reasoning remains to be tested.

Keywords: investigation games, evidence chain, misdirection, evidence-based reasoning, marine pollution

1. Introduction

A factory near an area of dead fish may be an obvious suspect in a marine pollution investigation. A nearby ship or a patch of floating rubbish could attract similar attention. Their presence gives investigators possible explanations to examine, but does not establish what caused the deaths. In ecological research, an observed association alone is insufficient to establish causation [1]. Sample testing, comparisons between locations, and relevant records are needed to assess these initial suspicions.

This distinction matters when designing an investigation game. A player may choose the correct source because it looks suspicious, even before examining the evidence. If the game accepts that choice without asking for a reason, it becomes difficult to tell whether the player has understood the case. The clues therefore need to give players a basis for checking their first explanation and recognising when it no longer fits.

Last 3 Days explores this problem through an investigation into unusual marine animal deaths. Players have three days to examine samples, use testing tools, and compare their findings with collection locations and reports. The design question is how these activities can help players develop and justify a conclusion about the pollution source.

The analysis follows the relationships between the clues in the prototype. It considers how evidence from individual animals contributes to an explanation of the wider event, how other possible causes can be assessed,

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

and how a clue can be accurate yet misleading. The final evidence selection and the planned evidence board are also discussed as ways of making the player's reasoning more visible.

2. Empirical analysis

2.1. Investigation and guessing

Take a simple case. A player comes across many dead fish and notices that there is a factory nearby. The player then immediately selects "factory pollution." This is still a guess, because the judgment is based only on the first thing observed. "There is a factory nearby" may count as a clue, but on its own it does not support a conclusion.

A single clue usually provides only part of the picture. For example, a test may detect pollutant X in a fish. That shows X was present in that sample, but it does not show that X caused the fish to die. Nor does it identify a particular factory as the source of X. Correlation does not equal causation [1]. In scientific reasoning, people have to link information to scientific knowledge and explain how that information supports a conclusion [2].

An investigation usually moves through several stages: identifying a problem, forming a hypothesis, gathering evidence, comparing findings, ruling out alternatives, and reaching a conclusion [3]. An initial guess is not necessarily an issue. The issue starts when that first guess is accepted as the final answer. A player's first idea can instead be treated as a hypothesis. From there, the investigation gives the player a chance to test it against evidence and see whether it holds up or turns out to be wrong.

2.2. Building an evidence chain

The investigation may start with a single sample. The player can use different tools to examine it, but the first test result is still only one piece of evidence. Suppose Sample A contains an unusually high concentration of substance X. Even then, the player only knows that A contains X. The source of X remains unknown.

Location information makes the result more useful because it lets the samples be compared by place. Suppose Sample A was found at Location 1 and has a high concentration of X. Sample B was found at Location 2 and has a lower concentration, while Sample C from Location 3 has almost none. With that added context, the player can compare the three samples and check whether the amount of X changes from one location to another. Without location data, each test result stands on its own.

More information can then be added. A map can show which area is most affected. Records on possible pollution sources help the player judge whether the sample results are consistent with what those sources might produce. Submitted testing reports can also be compared with the player's own results to check for inaccuracies. At this stage, these separate pieces of information start to form an evidence chain. They still need to be linked through reasoning before they can support a conclusion [2].

Comparison matters, but so does elimination. A reasoning game does not need to ask the player to spot the right answer at once. It can also allow the player to rule out options that no longer match the evidence [4]. Take three possible pollution sources: A, B, and C. Evidence 1 supports either A or B. Evidence 2 does not support A, and Evidence 3 does not support C. At that point, the player's conclusion is no longer "B looks the most suspicious." It becomes "B is the possibility that is best supported by the evidence I have." The answer may stay the same. The reasoning behind it changes.

2.3. Evidence chain design in *Last 3 Days*

In *Last 3 Days*, the fish containing plastic and the injured sea turtle suggest two possible explanations for the deaths. Plastic pollution is relevant to the fish, while an abandoned fishing net may explain the turtle's wound. At this stage, either clue could attract the player's attention. Examining the remaining samples is necessary to see how much of the case each explanation can account for.

The three testing tools give the player different information about an animal. With the magnifying glass, the player can inspect visible wounds on its surface. The X-ray reveals objects or abnormalities inside the body. The radiation detector checks for radiation abnormalities that would otherwise remain unnoticed. Alongside these results, the sample information and reports show where the animal was found and what has already been recorded.

These findings still require interpretation. Seeing a wound confirms that the animal was injured, but identifying its cause requires other information about the sample.

The player must also compare samples: a finding that explains one injury may have little relevance to abnormalities in another species. This is where examining the samples develops into reasoning about the case.

The repeated radiation abnormalities provide a possible link between animals with otherwise different conditions. If several species show similar readings and their collection locations are associated with the northern factory, industrial wastewater becomes a stronger explanation within the game's scenario. Plastic waste and fishing nets still account for some of the harm, but they leave this shared pattern unexplained. The case against the factory therefore depends on the combination of repeated readings and location information, rather than on a single abnormal result.

The company's testing report complicates this interpretation. Its normal result may initially reduce suspicion of industrial wastewater. Before accepting that result as evidence about the affected area, however, the player needs to check where the company's sample was collected.

If the sample came from another area, its normal reading would not necessarily conflict with abnormal readings in the north. If the report does concern the affected area but differs from the player's findings, that discrepancy requires further examination. In either case, the report alone is insufficient to dismiss the abnormalities found during the investigation.

The player therefore has to decide whether the report provides enough information to rule out industrial wastewater. A normal reading from the reported sample may be accurate while leaving the findings from other locations unexplained.

The final task asks the player to make two choices: select the main pollution source, then select the evidence supporting it. This gives the game a way to check the reason behind the answer. Choosing a source alone would leave that reason unclear.

A player could select industrial wastewater simply because the factory seemed suspicious at first. If the same player then chooses the plastic found in a fish as supporting evidence, the connection is missing: the plastic does not explain why wastewater is responsible for the repeated abnormalities. Evidence about the fishing-net injury would create a similar problem. These choices suggest that the player may have recognised the intended answer without understanding which findings support it.

The evidence selection returns to a question raised during the sample examinations: which findings concern one animal, and which help explain the wider event? To justify the main pollution source, the player needs evidence relevant to the pattern across samples. The final choices can show whether that connection has been made, although they cannot reveal the player's full reasoning process.

A related use of prompts to ask learners for reasons appears in research on students' engineering design decisions. Siverling et al. found that students frequently justified their ideas and decisions when responding to

adults or prompts on worksheets [5]. In *Last 3 Days*, the evidence selection serves a similar purpose by asking players to support their choice explicitly. Whether this prompt leads to more careful reasoning in the game still needs to be examined through player testing.

An evidence board was also planned for the prototype. Players would have been able to save clues and organise connections between samples, locations, test results, and possible pollution sources. Limited development time meant that the board was not included. In the current version, comparisons depend mainly on the sample order, the information pages, and the final evidence selection.

An evidence board in a later version could make their reasoning easier to follow, rather than leaving most of that information to memory.

2.4. Cross-checking information from multiple sources

Cross-checking information from different sources can turn an early guess into a judgment with stronger support. If one test shows a high concentration of substance X, the player only knows that one sample is unusual. That alone is limited. The picture changes when location data shows that the abnormality is clustered in one area and several other samples from the same area show the same problem. Taken together, these pieces of evidence make one possible pollution source seem more likely than it did at first.

Cross-checking is not just a matter of adding more clues. A game may include many clues and still ask for very little reasoning when each one points clearly to the same answer. What matters more is how the clues relate to each other, not how many there are. Information only becomes evidence when the player works out which pieces support the same explanation and which do not fit it [2].

In *Last 3 Days*, each type of information does a different job. A marine animal sample shows the condition of one specific animal, while testing tools give more detail about the kind of abnormality present. Location data helps show whether similar problems are clustered in one area. Testing reports can then be checked against the results collected by the player.One source on its own is not enough to identify the final pollution source. The investigation becomes more focused when several sources point to the same conclusion.

Evidence can rule out some explanations, not just support the right one. For example, an abandoned fishing net may account for a wound on a sea turtle. It does not explain why other animals show the same internal abnormality. The net still matters, and the problem is real. Even so, it is not a convincing explanation for the wider event.

This changes how the player approaches the answer. Rather than asking, "Which option seems most correct?", the player asks, "Which possibilities no longer explain what I have found?" The final answer comes from collecting supporting clues and ruling out explanations that do not fit the overall pattern.

2.5. The design of misdirection

When every clue points to the same pollution source, a player can reach the answer too quickly, without really weighing the evidence. Some misdirection helps avoid that. It pushes players to question an early assumption and keep looking for more information.

Misdirection does not necessarily mean giving the player false information. In research on crime fiction, Gregoriou discusses clue hiding and misdirection, including how information can shape the way readers interpret a case [6]. The study deals with crime fiction rather than games, but the idea still helps explain how clues can steer an investigation without directly giving away the answer.

In *Last 3 Days*, plastic waste and abandoned fishing nets work this way. They are not false clues. Abandoned fishing gear can entangle marine animals and cause injury or death[1]. Both are real environmentalproblems, and both can harm marine animals, so at the start the player has a fair reason to suspect them. Later, additional samples show the limits of these explanations. They account for some injuries, but they do not explain the abnormality that appears repeatedly across different animals.

The company's testing report creates a different kind of misdirection. A normal result can briefly make industrial wastewater look less likely. But the report does not have to be entirely false for that effect to occur. A normal sample only shows what was found in that specific sample. It does not prove that every part of the sea is safe. The player therefore needs to check where and when the sample was taken, and whether the result fits with the other evidence.

Misdirection works best when the wrong path seems plausible at first. The player should be able to see why an explanation made sense early on, then later find the point where it stops being enough. If a game only gives a false test result or withholds information the player needs, arriving at the wrong answer has little to do with reasoning. It isn't fair. The player was never given a real chance to examine the case.

A balance is needed here. If the misdirection is too weak, the answer becomes obvious. If it is too strong, the whole case can start to feel unreliable. In *Last 3 Days*, the aim is to let a false lead account for part of the case while still leaving enough evidence for the player to see where it stops making sense. Later evidence should give the player a clear reason to reconsider the initial explanation.

2.6. Limitations and future improvements

The unfinished evidence board is one limitation of the current prototype. It was intended to let players organise samples, collection locations, test results, and possible pollution sources in one place, but could not be completed within the development time. Players currently compare their findings through the sample sequence, information pages, and final evidence selection screen. These provide access to the clues, but do not record how a player connects them during the investigation.

This makes some difficulties hard to identify. A player may overlook a connection because they misunderstand a test result or have forgotten an earlier sample. The planned evidence board would let players keep the relevant clues together [7]. For example, they could link an animal sample to its collection location and connect its test result to a suspected pollution source [8]. Later findings might give them a reason to change those links. Player testing would be needed to find out whether this feature helps them compare evidence.

The prototype also has relatively few samples and possible pollution sources. A player may therefore choose the correct option without being able to explain it fully. Ruling out a source because it conflicts with the findings is a valid part of the investigation, but guessing among the remaining options tells us less about the player's understanding. The final evidence selection helps check the reason for the answer, though it does not show the earlier decisions.

A future version could include a wider range of samples and evidence for player testing. The test could record each player's first suspected source and ask which findings, if any, changed their view. Their explanations could then be compared with the evidence they select at the end. This would help show how players use the clues and where the investigation needs improvement.

[1] National Oceanic and Atmospheric Administration (NOAA). (2021). What is marine debris? NOAA National Ocean Service.

3. Conclusion

This paper examined how evidence chains and misdirection work in a marine pollution investigation game, shifting players from an initial guess to a conclusion grounded in evidence.

In *Last 3 Days*, plastic inside a fish and a wound on a sea turtle can account for harm to those animals, but leave the repeated radiation abnormalities unexplained. Comparing results across species and collection locations gives the player a basis for assessing industrial wastewater as the main source within the game's scenario.

The company's report adds a further comparison. Its normal result needs to be read in relation to the location sampled and the player's own findings. A result from another area cannot, by itself, resolve the abnormalities found in the north. The evidence chain develops through these relationships between findings and the claims they can support.

An explanation that accounts for one injury but leaves the shared abnormalities unexplained is therefore insufficient as the main explanation of the crisis. The player needs to reconsider its role while retaining the evidence about the individual harm.

References

[1] Correia, H. E., Dee, L. E., Byrnes, J. E. K., Fieberg, J. R., Fortin, M. J., Glymour, C., Runge, J., Shipley, B., Shpitser, I., Siegel, K. J., Sugihara, G., von Holle, B., & Ferraro, P. J. (2026). Best practices for moving from correlation to causation in ecological research. *Nature Communications, 17*, 1981. https://www.nature.com/articles/s41467-026-69878-z

[2] Novak, A. M., & Treagust, D. F. (2026). Developing deep knowledge by connecting scientific ideas with data through reasoning to support claims. *Disciplinary and Interdisciplinary Science Education Research, 8*, 20. https://link.springer.com/article/10.1186/s43031-026-00164-0

[3] Bónus, L., Antal, E., & Korom, E. (2024). Digital game-based inquiry learning to improve eighth graders' inquiry skills in biology. *Journal of Science Education and Technology, 33*(4), 462–478. https://link.springer.com/article/10.1007/s10956-024-10096-x

[4] Pilditch, T. D., Fenton, N., & Lagnado, D. (2019). The zero-sum fallacy in evidence evaluation. *Psychological Science, 30*(2), 250–260. https://journals.sagepub.com/doi/10.1177/0956797618818484

[5] Siverling, E. A., Moore, T. J., Suazo-Flores, E., Mathis, C. A., & Guzey, S. S. (2021). What initiates evidence-based reasoning?: Situations that prompt students to support their design ideas and decisions. *Journal of Engineering Education, 110*(2), 294–317. https://onlinelibrary.wiley.com/doi/10.1002/jee.20384

[6] Gregoriou, C. (2022). On the making of Robinson's stylistic "fast ones" through the Banks series' early years. *English Studies, 103*(3), 407–427. https://www.tandfonline.com/doi/full/10.1080/0013838X.2022.2043002

[7] Saleh, A., Yuxin, C., Hmelo-Silver, C. E., Glazewski, K. D., Mott, B. W., & Lester, J. C. (2020). Coordinating scaffolds for collaborative inquiry in a game-based learning environment. *Journal of Research in Science Teaching, 57*(9), 1490–1518. https://onlinelibrary.wiley.com/doi/10.1002/tea.21656

[8] Liang, H. Y., Hsu, T. Y., & Hwang, G. J. (2021). Promoting children's inquiry performances in alternate reality games: A mobile concept mapping-based questioning approach. *British Journal of Educational Technology, 52*(5), 2000–2019. https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13095