Evaluating AI-Based Games through Retellings

Max Kreminski,[1] **Ben Samuel,**[2] **Edward Melcer,**[1] **Noah Wardrip-Fruin**[1]

[1]University of California, Santa Cruz
[2]University of New Orleans
{mkremins, eddie.melcer, nwardrip}@ucsc.edu, bsamuel@cs.uno.edu

Abstract

We propose a new approach to the human-centered evaluation of AI-based games, grounded in the analysis of player retellings of their play experiences. Retellings offer unique insight into dimensions of player experience that can be hard to get at through existing evaluation methods, such as the typical narrative structures that tend to emerge in the player's mind when they play a particular game; the variety of subjectively experienced narratives that are possible and probable within a particular game; and the ways in which a game supports, or fails to support, the player's process of narrativization. We used a grounded theory methodology to analyze retellings of play experiences in *Civilization VI*, *Stellaris*, and two distinct versions of the research game *Prom Week*. We also interviewed the creators of several retellings to gain insight into the subjective experience of story construction in collaboration with these games.

Introduction and Related Work

One enduring argument for the introduction of artificial intelligence and procedural content generation into games has been that they will enable a wider variety of more meaningfully distinct experiences from player to player. Another, closely related argument holds that AI and PCG will enable games to respond more deeply and meaningfully to player input. However, existing methods of evaluation do little to surface this dimension of player experience. As a result, this dimension of player experience tends to be left out of our analyses of AI-based (Treanor et al. 2015) and PCG-based (Smith et al. 2011) games.

Ryan's work on what he terms *curationist interactive narrative* (Ryan 2018) has recently called for the development of AI systems that assist players in sifting through deep simulations to find storyful content. As a step in this direction, we believe it is valuable to look at how players narrativize their experiences in existing AI-based games. More broadly, we are interested in games as sites of player creativity, and—through a narrative-focused lens—particularly as sites of story construction by players. Therefore, we are interested in developing an evaluation method that addresses the extent to which games support narrativization by players of their play experiences.

Meanwhile, Eladhari (Eladhari 2018) has called for a focus on *retellings*—the stories players tell about their play experiences—as an instrument for the critique of interactive narrative systems. Framing retellings as narrative artifacts co-created by players and the games that they play, she suggests that studying these narrative artifacts may enable a deeper understanding of player experience. From this perspective, players are already making use of games as storytelling partners—and, in some cases, seeking out games with the specific intent of using them to support story construction. Consider, for instance, the case of *Alice and Kev*, a lengthy *Sims 3* retelling that Eladhari highlights as an example of a compelling story co-created by a human and an (AI-based) game.

Retellings are in some ways similar to and in some ways different from *playtraces*, which aim to capture, objectively, what actually happened in the course of play. We propose to study retellings as a kind of subjectively-grounded playtrace. We contend that, in the course of play, players often imagine or create stories beyond those that are represented literally in the game. They may place unexpected weight on events or details, including apparently cosmetic ones, that seem inconsequential from a gameplay perspective; extrapolate the ramifications of events or details in ways that were not intended by the game's developers; and generally bring their own creativity and subjectivity to the process of narrativizing their play experiences. When players reflect on their play experiences in order to relate or retell them to others, the retellings they produce may then capture dimensions of experience that are absent from even the most exhaustive playtrace. Moreover, whether and how a game is narrativizable by players may tell us things about whether and how it makes sense to people more generally.

Simultaneously, it is very difficult for players to convey rich, emotional experiences encountered during gameplay (Laaksolahti, Isbister, and Höök 2009), and asking players about these experiences outright can lead to misleading results (King 2004; Leech 2002). We hope our approach will allow the player to more deeply describe a narrative play experience uninhibited, and in turn allow us to glean much more insightful information on how certain mechanics enable or inhibit players in the construction of stories.

In this paper, we describe several possible ways that player-created retellings of their play experiences may be

used to evaluate AI-based and PCG-based games. First, we explore how written retellings of *Civilization VI* and *Stellaris* play experiences found in online player communities can be used to develop a codebook capturing relevant dimensions of variation between retellings. Next, we analyze a small number of *Civilization VI* and *Stellaris* retellings, in conjunction with interviews with the authors of these retellings, to develop a better understanding of the subjective experience of storytelling in collaboration with these games. Finally, we apply a similar methodology in conjunction with an ablation study approach to investigate the impact of social reasoning on player narrativization in the research game *Prom Week*, for which few if any "naturally occuring" retellings exist.

Collecting and Analyzing "Wild" Retellings

Civilization VI and *Stellaris* are strategy games that make integral use of both AI (to populate simulated worlds with rival factions for players to interact with) and PCG (to generate world maps on which civilizational struggles may play out in interesting and unexpected ways from one playthrough to the next). Both games have attracted sizable player communities, including some players who engage with the game by constructing retellings and posting them to public forums.

We hypothesized, based on our own experience in both games, that *Stellaris* facilitates player storytelling better than *Civilization VI*, and therefore that *Stellaris* retellings would be both more common (relative to the size of the player community) and more narratively rich than *Civilization VI* retellings. Our notion of "narrative richness" at this point was still vague and undefined, so we turned to analysis of existing retellings to determine what a narratively rich retelling might look like.

We used a grounded theory methodology (Salisbury and Cole 2016) to analyze retellings found "in the wild," in online player communities for both *Stellaris* and *Civilization VI*. GTM enables a researcher to simultaneously analyze a body of artifacts (in this case retellings) and develop a theory about what elements of these artifacts are salient, typically embodied as a codebook which evolves over the course of the analysis and is used to note down the features of specific artifacts as they are analyzed. In this particular study, we adopted the Constructivist Grounded Theory flavor of GTM (Charmaz 2006), which frames the researcher as "co-creating meaning within the domain they are studying" (Salisbury and Cole 2016) and focuses on providing lenses for analysis rather than presenting a single objectively correct model of what matters about the domain.

In applying this methodology to retellings, we first located the largest Reddit subreddit for each of the games we wanted to study—/r/Civ and /r/Stellaris respectively—and gathered 50 retellings from each subreddit. Retellings were located by searching these subreddits for the keywords "stories" and "AAR" (an abbreviation of the phrase "after-action report," commonly used to tag retellings in strategy game communities). Retellings from /r/Civ were also manually filtered to ensure that they were *Civilization VI* retellings specifically. We then began analyzing these retellings, making note

as we went of recurring themes that appeared across multiple retellings, and developing a theory which we continually updated of what elements of retellings could reflect design features of the games that were employed to create them. Similar methods have previously been applied to the qualitative coding and analysis of other player-generated texts around games, such as game reviews (Bopp, Mekler, and Opwis 2016).

The codes we chose to include in the codebook were based largely on our own intuitive understanding of what constituted narrative interest in the context of retellings, which grew more sophisticated as we analyzed more of the "naturally occurring" *Civilization VI* and *Stellaris* retellings available to us. Eventually, we reached a point of saturation at which the study of additional retellings ceased to yield new codes for the codebook. The set of codes included in our final codebook is the set that we had developed by this point in our analysis.

Ultimately, we analyzed 20 *Civilization VI* retellings and 20 *Stellaris* retellings (chosen at random from the 50 retellings we had gathered for each game) before reaching saturation. The final codebook (described in Table 1) contained 10 codes in total, plus a qualitative "story type" classification describing the focus of the retelling in terms of subject matter. We encountered four broad story types in *Civilization VI* and *Stellaris* retellings:

- An unstructured chronicle of game events, typically comprising everything that had taken place in a particular playthrough so far. (16 *Civ VI* stories, 9 *Stellaris* stories)

- The story of a demarcated period in history, with a clear start and end—such as a particular war, or the formation of a federation. (3 *Civ VI* stories, 6 *Stellaris* stories)

- The biography of a particular named character, generally some sort of leader. (1 *Civ VI* story, 4 *Stellaris* stories)

- A day in the life of an individual member of a larger faction or civilization. (0 *Civ VI* stories, 1 *Stellaris* story)

Stellaris retellings are substantially more diverse in subject matter than *Civilization VI* retellings. For both games, unstructured chronicles of game events are the most common story type, but *Stellaris* stories are far more evenly distributed across the four story types, and also include the only example we identified of a "day in the life of a citizen" story—wherein the point-of-view character was not the player, not an outside narrator, and not modeled directly as an individual character within the game's systems.

In both *Civilization VI* and *Stellaris*, mixing of perspectives within a single retelling is common. Of the *Civilization VI* retellings we studied, a majority are written partly or wholly in player perspective. Instances of narrator perspective are uncommon, and we identified only a handful of instances of in-world perspective. Perspective in *Stellaris* retellings is more evenly mixed, with in-world perspectives appearing much more frequently.

Cast size varies, but *Stellaris* retellings tend to have more named individual characters than *Civilization VI* retellings overall. 13 *Civilization VI* stories and 7 *Stellaris* stories do not mention any named individual characters, instead

Code name	Code description
Player perspective	A section of text written from the perspective of the *player*, typically using first-person voice (e.g., "I started in Malaysia, about the exact center of the map.")
Narrator perspective	A section of text written from the perspective of an *outside narrator*, typically using third-person voice (e.g., "The Zulu declared a formal war on the Kongolese.")
In-world perspective	A section of text written from the perspective of a *character in the fictional world*. Includes sections written from the perspective of an *in-universe* chronicler. (e.g., "I rushed to the bridge, explosions resounding around me.")
Specific fictional character	Includes a specific fictional character. The number of these identified within a single retelling corresponds to the "cast size" of the retelling as a whole. Applies exclusively to individuals, not to groups or factions. (e.g., "This is the story of Yesenia Qasim, and his truely amazing contributions to the human empire.")
Extrapolated flavor element	An instance of the player extrapolating story details not directly modeled in the game from in-game occurrences (e.g., "As the colony descended into chaos, Dalurkot turned to drink.")
In-world dialogue	An occurrence of in-world dialogue spoken by a particular character (e.g., "Kristina, Bandar Brunei has fallen to Montezuma!")
Mechanical reason for decision	Mentions a mechanical or player-centric reason behind a decision taken by the player (e.g., "I decided a religious victory didn't sound too feasible, and so built St Michel for the relics.")
Non-mechanical reason for decision	Mentions a non-mechanical or character-centric reason behind a decision taken by the player (e.g., "Traak elected to stay and fight. He couldn't bear to leave the home he had known for so long.")
Reversal	An established trend is reversed: perhaps the power relationship between two characters or factions is inverted, or a character changes their mind about one of their goals (e.g., "Doug was starting to doubt his attitude, and he started to look at Oswald in a different light.")
Character development	A character's current characterization is contrasted with their earlier characterization to signal that the character has changed over time (e.g., "His military conquests have twisted his mind and Kupe has begun to militarize on an immense scale.")

Table 1: The codes used in the codebook we developed to note relevant dimensions of variation between *Civilization VI* and *Stellaris* retellings. The same codebook was later applied to the analysis of retellings created in *Prom Week* as well.

referring exclusively to factions. *Stellaris* retellings, however, have a higher ceiling on cast size than *Civilization VI* retellings: of the *Stellaris* stories we studied, the one with the largest cast had 14 named individual characters, while the *Civilization VI* story with the largest cast had only 8. Only one *Civilization VI* retelling we analyzed had any named characters that were not national heads of state. Typically in *Civilization VI* retellings, national heads of state are used metonymously to refer to the factions they lead. *Stellaris*, on the other hand, explicitly models named leader characters (including governors, scientists, admirals, and generals) who are not national heads of state, and these characters appear frequently in *Stellaris* stories.

Anecdotally, the frequency of reversals within a retelling seems to align well with an intuitive understanding of which retellings are more narratively compelling. Retellings with no reversals typically follow a story pattern chronicling the steady rise of the player faction, with few significant obstacles along the way. Retellings with a small number of reversals (up to three) tend to feel more compelling, while retellings with more reversals begin to feel incoherent.

Within *Stellaris* retellings, certain specific game elements make especially frequent appearances. These include scripted narrative events, especially those involving named leader characters, which may occur randomly when certain event-specific preconditions are met; instances in which a named leader character gains a trait, which provide an opportunity for players to infer or extrapolate character devel-

opment; periods of building tension between rival factions, often leading up to a war; and large decisive battles within wars. These events seem to be perceived by players as particularly storyful.

Members of the *Stellaris* community seem to share retellings with other community members more frequently than members of the *Civilization VI* community, especially relative to overall community size. As of May 2019, /r/Civ had an estimated 283,000 members, but only approximately 6,050 search results for the "stories" and "AAR" keywords combined. /r/Stellaris, on the other hand, had only an estimated 140,000 members, but approximately 10,700 search results for both keywords combined. The discrepancy is further compounded by the fact that /r/Civ is a combined community for the entire *Civilization* series, and not all retellings posted there are *Civilization VI* retellings specifically. This may indicate that players find *Stellaris* to be a better storytelling partner, or that the stories they experience within it seem more unique, personally significant, or otherwise worth sharing. There may also be a degree of "social contagion" involved: members of the *Stellaris* community may first see others sharing stories and only thereafter decide to share stories of their own.

Interviewing Retelling Creators

Analysis of a sizable corpus of preexisting retellings enables some forms of insight into the variety of subjective play experiences that are both possible and likely within a particu-

lar game. However, we hypothesized that additional insight would be enabled by analyzing retellings not just as isolated artifacts, but in conjunction with interviews with the players who created these retellings. This poses some initial difficulties: it may not be feasible to get in touch with the creators of retellings that were posted online, and the creator's memory of the experience of working with the game to create the retelling may have faded with time since the retelling was initially created.

As such, we decided that the best approach would be to recruit a small number of players to create retellings of their own, and then interview them about their experience. We could then study the retellings they created in conjunction with their interview responses, potentially giving insight into how the structural features of a retelling are related to dimensions of the creator's subjective experience.

We recruited two undergraduate game design students to play both *Stellaris* and *Civilization VI*, and to create two retellings each—one based on their experiences in each game—for a total of four retellings overall. Neither student had played *Stellaris* before; one had played *Civilization VI* before, but had never previously approached the game with the explicit intent of telling a story.

We coded the retellings created by the students according to the codebook we developed through analysis of wild retellings. In addition, we interviewed both of the students about their experiences creating retellings in both games.

All four student-created *Civilization VI* and *Stellaris* retellings fell into the "unstructured chronicle of game events" story type, which seems to be the natural default for retellings in these games. In *Civilization VI* wild retellings, deviations from this pattern are the exception, and even in *Stellaris* wild retellings this pattern is the most common. In addition, both *Civilization VI* retellings and one of the two *Stellaris* retellings recounted game events more or less directly, with little extrapolation or introduction of new creative elements by players.

However, one of the two student-created *Stellaris* retellings did introduce fictionalized and extrapolated characterization to some extent. For instance, when a character who had been appointed as the governor of a colony gained a new trait representing a substance abuse problem, the player framed this in their retelling as a response to the stress of ruling during a particularly tumultuous time for that colony. Both the trait and the troubled period of the colony's history were modeled in game to some extent, but they were not directly mechanically linked. Instead, the player chose to extrapolate deeper characterization from the juxtaposition of these two mechanical elements.

The relative shallowness of the majority of student-created *Civilization VI* and *Stellaris* retellings may be due partly to inexperience with the games being used. *Stellaris* in particular is a complicated game with lots of menus that takes some time to get used to, and neither of the two players had played *Stellaris* before.

One player found it substantially more difficult to create stories with *Stellaris* than with *Civilization VI*, and did not feel especially strongly about the difference between *Stellaris* and *Civilization VI* with regard to storytelling support

or coherence: "The game went very quickly, and there was a lot of things happening at once of which it didn't always tell me that it was happening. I sometimes missed important things while I was trying to read about other important things, which made my story feel like it was jumping around from point to point and not very linear." The other player found that *Stellaris* was less difficult to tell stories with, strongly supportive of their storytelling process, and substantially better than *Civilization VI* at promoting story coherence. This, too, may be due in part to disparity in previous experience: neither had played *Stellaris* before, so both were confronted by a significant learning curve, while one had played *Civilization VI* before, reducing the impact of the learning curve.

Due to the role of experience in mitigating confusion, interviews with experienced creators of *Civilization VI* and *Stellaris* retellings would likely yield different insights. For this study, however, we wanted to establish a baseline for comparison with *Prom Week*. To the best of our knowledge, there are no experienced creators of *Prom Week* retellings in the wild, so we instead elected to use players with little experience in any of the games being studied.

In future studies taking a retellings-based approach to evaluation, it may be desirable to give players more time to get acclimated to the game before having them try to tell stories with it. Creators of retellings in the wild are often not first-time players, and sometimes mention when reporting play experiences that this is "the most interesting game they've had so far" out of several previous playthroughs. It may simply be the case that it takes time to learn the tools needed to tell stories effectively with any sufficiently complicated game.

Players in both *Civilization VI* and *Stellaris* reported feeling consistently surprised by the stories they ended up telling in collaboration with the game. We address this further in the following section.

Evaluating *Prom Week* Through Retellings

Prom Week is an AI-based research game that makes central use of the *Comme il Faut* (McCoy et al. 2014) "social physics engine" to determine the set of possible actions and responses of a cast of eighteen virtual high school students. Each turn, players select pairs of characters and have them engage in a socially-charged action (such as "Ask Out" their crush on a date, or "Backstab" their trusting friend), the options of which are determined based on thousands of social considerations that take into account the current social state and the characters' past history with one another. Each player-selected action further evolves the social state, which further affects character considerations for subsequent actions, resulting in playtraces that quickly become entirely unique for each player.

To evaluate the impact of *Prom Week*'s social reasoning on player narrativization of their play experiences within the game, we adopted an approach inspired by *ablation studies*, in which a specific system is selectively disabled and comparisons made between scenarios in which the system is active and scenarios in which it is not. We created an alternative version of *Prom Week* that had been modified to disable

the social reasoning system, replacing "intelligent" character decisions about which action they should take next with random selection from the entire pool of possible actions. We refer to the publicly available version of *Prom Week* as *PW-A* and the modified version with social reasoning disabled as *PW-B*.

We hypothesized that *Prom Week*'s social reasoning and explicit modeling of character intentionality would support player storytelling well, leading players to feel that storytelling with the game was easier; that the stories they told were more coherent; and that their storytelling process was better supported in *PW-A* than in *PW-B*. We also hypothesized that players would find the stories they told with *PW-A* more surprising (due to the social reasoning giving the game a kind of coherent creative agency, which might contradict or react in unpredictable ways to the player's own creative intent); that they would enjoy telling stories more with *PW-A*; and that they would experience writing block more frequently when telling stories with *PW-B*.

The same two undergraduate students we had recruited to create *Civilization VI* and *Stellaris* retellings were again instructed to create two retellings each: one in *PW-A*, and one in *PW-B*. They were not informed of the nature of the difference between the two versions of the game, and neither had encountered *Prom Week* before. We exposed them to the two versions of the game in opposite orders, with the goal of minimizing possible ordering effects.

Interview responses showed three of our hypotheses to be correct: *PW-A* was consistently experienced as being less difficult to tell stories with, better for story coherence, and more supportive of player storytelling than *PW-B*. *PW-A* in particular prompted a variety of interview responses from both players about its success at supporting storytelling:

- "This game was the perfect mixture of detailed and open-ended. This made it easier to have a good outline for a story, and personally fill-in the specifics with my imagination. It also wasn't completely up to you, because the game would sometimes end before you managed to complete all your goals so sometimes you didn't get the ending you were hoping for, which allows for stranger details in your story and more realistic life-like endings."

- "I started to understand certain characters' personality and therefore I started considering things from their perspective or why they had made those decisions"

- "events occurred that I felt I could exaggerate since the experience were very broad but the feelings that came out could be added onto. For example, when I wrote about the protagonist going to talk to a crying Naomi after she was heartbroken, I tried to write a convincing talk between the two rather than what game actually had."

- "sometimes it felt out of your control which made it more of a compelling story to write"

Analysis of the retellings created with *PW-A* reveals that they both contain six or fewer reversals, while both of the retellings created with *PW-B* contain eight or more. All four of the *Prom Week* retellings are of similar length (between 1500 and 2000 words), so the *PW-B* retellings are substantially more reversal-dense than the *PW-A* retellings. This corresponds well to both our anecdotal analysis of coherence in relation to reversal count in *Civilization VI* and *Stellaris* retellings, and to player interview responses regarding the coherence of retellings constructed with *PW-A* and *PW-B*.

Contradicting our hypothesis regarding surprise, interview responses showed players to be approximately equally surprised by all the stories they told in both *PW-A* and *PW-B*, with no meaningful difference between the two versions. Similarly high levels of surprise were reported across the board when interviewing the same players about the stories they told using *Civilization VI* and *Stellaris*. This may be an indicator that players will usually be surprised by the stories they end up telling in collaboration with games as long as the game is given meaningful creative input into the storytelling process. Alternatively, it may simply mean that we framed the question poorly, or that asking about the extent to which players were surprised by the stories they ended up telling is unlikely to yield any meaningful responses in this context.

Of the two players, one reported a consistently low-to-moderate level of enjoyment of the storytelling process in both *PW-A* and *PW-B*. The other enjoyed creating stories with *PW-A* but reported greater frustration in *PW-B* due to the greater unpredictability of character behavior: "At first I found it fun, but it started to get more and more annoying when the players didn't respond how I wanted them to, or how I had expected them to after playing the game for a bit." This player had been introduced to *PW-B* first, so their initial expectations had not been shaped by the greater coherence of character behavior in *PW-A*. This seems to suggest that their frustration stemmed directly from the difficulty of accurately predicting character behavior in the absence of social reasoning.

Player interview responses on the subject of writer's block diverged substantially. One player frequently experienced writer's block in *PW-A* but not in *PW-B*, because the constant failure of characters to act in predictable ways propelled the story forward despite decreasing coherence: "I didn't really feel blocked, I just kept writing about my failed attempts to secure the goals." The other frequently experienced writer's block in *PW-B* but not in *PW-A*, due to unpredictable character behaviors forcing constant re-planning of the direction the story might go. This may indicate the existence of two different attitudes toward story construction—one more improvisational, one more rigidly planned—which require different kinds of creativity support.

Discussion

For all of the games examined here, many of the more narratively resonant retellings seem to be driven by a process that we term *extrapolative narrativization*. Rather than stopping at literally reporting the events of a particular play experience as they are modeled within the game (even if in a heavily filtered form), players who create retellings are prone to seizing on particular details or combinations of details that appear within the game—often apparent coincidences, or elements of game events that are largely inconsequential from

a gameplay perspective—and extrapolating from these details to include additional story elements that are not directly modeled within the game's systems.

Creators of *Stellaris* retellings in particular have a tendency to engage in *extrapolative worldbuilding*: inventing details about the nature of the fictional world in which gameplay takes place, for instance by describing elements of day-to-day life in an alien civilization at a far greater level of granularity than is directly modeled by the game's systems.

We argue that the high frequency of extrapolative narrativization in *Stellaris* retellings, especially in contrast to the relative infrequency of extrapolative narrativization in *Civilization VI* retellings, supports our hypothesis that *Stellaris* more effectively supports player storytelling than *Civilization VI*, which seems less liable to prompt its players to engage in imaginative worldbuilding processes. We suspect that *Stellaris*'s use of procedural content generation to create non-player characters and rival factions is partly responsible for this difference. Due to their procedurally generated nature, *Stellaris* characters and factions are underdetermined, providing sketchy outlines of people and civilizations that players aren't familiar with from elsewhere. Randomly assigned character traits, species traits and portraits, civics, and other aspects of government structure provide lots of evocative details from which players are free to hang story bits, without worrying that they may be contradicting actual history as is sometimes the case in *Civilization VI*.

Similarly, *Stellaris*'s direct modeling of a cast of persistent, named leader characters who participate directly in gameplay events—such as researching technologies, exploring star systems, fighting battles, and ruling particular planets—provides players with still more opportunities to extrapolate story details from things that happen in-game. *Prom Week* takes this a step further by directly modeling not only individual characters but also their relationships and social motivations, and then exposing this model directly to the player. This facilitates storytelling even more strongly than *Stellaris*'s modeling of characters, albeit only in the presence of social reasoning, which imposes a kind of predictability on the characters that aids coherent storytelling.

This analysis suggests several possible things you can do as a designer to facilitate player storytelling: provide evocative "hooks," such as character traits, that can be interpreted in a few different ways, especially in juxtaposition with other hooks; leave some details of worldbuilding and characterization strategically underdetermined, to let players fill them in; explicitly model individual characters in game systems; enable these individual characters to participate in a variety of game events; and do what you can to either model character intentionality, give players hooks with which to infer intentionality, or some combination of both.

Potential Drawbacks of Retelling-Based Evaluation

Compared to other approaches to the evaluation of play experiences, retellings-based approaches can be time-consuming. The GTM-inspired qualitative coding process we describe here may take many hours, especially if attempting to develop one's own codebook rather than using an existing one (which we encourage for future studies, especially those not focusing on strategy games.)

Additionally, if studying a game for which retellings don't already exist in the wild, or if you want to have local players create retellings so that you can readily interview retelling creators, it takes more time for players to play the game and write a story with it than it does to go through a traditional playtest or take a simple survey. If the game is at all complicated, you will likely also want to allow players some time to get used to playing the game before they begin creating their stories. In this study, we did not allow players enough time to get acclimated prior to retelling construction.

Conclusions

Echoing Eladhari, we suggest that the existence of a sizable corpus of "naturally occurring" retellings for a particular game may serve as an indicator that the game is in some sense a success. We also conclude that studying the content of retellings can provide deeper insight into what game elements lend support to player storytelling; what overall trajectories and themes are possible and probable in player stories emerging from a particular game; and what dimensions of play experiences stand out to players as especially salient, compelling, and worth sharing with others.

Retellings in the wild capture the bits of subjective player experience that players feel are salient enough to be worth telling others about. Considering a number of retellings for the same game in aggregate can also give a sense of the overall diversity of compelling stories as subjectively experienced by players, something which is difficult to access through asking players directly about their experience. Moreover, many retellings we examined made direct or indirect reference to particular game elements, systems, or emergent play patterns. Considering which of these elements tend to appear more frequently in retellings can give researchers a sense of which parts of games tend to play a role in creating compelling player experiences and helping players narrativize these experiences.

Simultaneously, soliciting players to create retellings as a research method allows you to interview retelling creators about their subjective experience. Studying retellings and interviews in conjunction gives you more insight than you would be able to get from studying either alone.

Future Work

Going forward, we hope to apply the approach developed here to the evaluation of other AI-based and PCG-based games. The ablation study-inspired approach we took to the evaluation of *Prom Week* seems especially likely to be useful for evaluating research games, which we can readily modify to create variants with certain systems disabled.

We also intend to conduct interviews with the creators of several high-profile or especially well-crafted retellings—such as *Alice and Kev*, a lengthy episodic *Sims 3* retelling about a homeless family mentioned by Eladhari in her earlier work on the subject of retellings (Eladhari 2018). By talking to the creators of these especially elaborate retellings, we hope to learn more about why players seek out certain games as storytelling partners and what features can assist players in constructing their own narratives more effectively.

References

Bopp, J. A.; Mekler, E. D.; and Opwis, K. 2016. Negative emotion, positive experience?: Emotionally moving moments in digital games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2996–3006. ACM.

Charmaz, K. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.

Eladhari, M. P. 2018. Re-tellings: the fourth layer of narrative as an instrument for critique. In *International Conference on Interactive Digital Storytelling*, 65–78. Springer.

King, N. 2004. Using interviews in qualitative research. *Essential guide to qualitative methods in organizational research* 2:11–22.

Laaksolahti, J.; Isbister, K.; and Höök, K. 2009. Using the sensual evaluation instrument. *Digital creativity* 20(3):165–175.

Leech, B. L. 2002. Asking questions: Techniques for semistructured interviews. *PS: Political Science & Politics* 35(4):665–668.

McCoy, J.; Treanor, M.; Samuel, B.; Reed, A. A.; Mateas, M.; and Wardrip-Fruin, N. 2014. Social story worlds with Comme il Faut. *IEEE Transactions on Computational intelligence and AI in Games* 6(2):97–112.

Ryan, J. 2018. *Curating Simulated Storyworlds*. Ph.D. Dissertation, University of California, Santa Cruz.

Salisbury, J. H., and Cole, T. 2016. Grounded theory in games research: Making the case and exploring the options. In *Proceedings of the 1st International Joint Conference of DiGRA and FDG*.

Smith, G.; Gan, E.; Othenin-Girard, A.; and Whitehead, J. 2011. PCG-based game design: enabling new play experiences through procedural content generation. In *Proceedings of the 2nd International Workshop on Procedural Content Generation in Games*, 7–10. ACM.

Treanor, M.; Zook, A.; Eladhari, M. P.; Togelius, J.; Smith, G.; Cook, M.; Thompson, T.; Magerko, B.; Levine, J.; and Smith, A. 2015. AI-based game design patterns. In *Proceedings of the 10th International Conference on the Foundations of Digital Games (FDG 2015)*. ACM.