

# Focalised Story Sifting: A Matter of Perspective

Sam Drysdale

September 5, 2023

## Abstract

Story sifters are tools for procedural storytelling, designed to recognise and retell the emergent narratives of an underlying simulation. At present, sifting techniques are generally omniscient, each operating over a complete sequence of simulated events. This research introduces *focalised story sifting*: an alternative approach, grounded in narratological theory, that works from only the subsequence of events observed by a given ‘in-world’ character. The hope is to achieve a more subjective framing of emergent narrative, one that allows for unreliable, *Rashomon*-style narration.

To explore the technical and aesthetic implications of the approach in full, the research ‘focalises’ an existing statistical sifting heuristic so as to estimate the moments a character finds most surprising. This sifter surfaces a set of narratives from a custom-built storyworld simulation, for assessment by questionnaire. More than just verify the sifter is implemented correctly, the aim is to find what meanings participants explicitly draw from the narrative artifacts - and, implicitly, what meaning is created by the act of generation itself.

Results are mixed. Participants show an active engagement with the hallmarks of unreliable narration present in the artifacts, but ultimately find them too open-ended to make for a fulfilling reading experience. To the extent that this does not constitute an unambiguously positive result, the conclusion suggests promising design improvements to be made to both the sifter and its accompanying simulation. Further opportunities for future work are identified from the literature, considering where else narratology might intersect with the open problems of emergent narrative. The paper ends by drawing a throughline from the origins of the emergentist approach to its idealised vision of what focalised story sifting could one day be.

# Contents

<b>List of Figures</b>	<b>ii</b>
<b>List of Tables</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Background</b>	<b>4</b>
2.1 Narratology . . . . .	4
2.1.1 Emergent Narrative . . . . .	5
2.1.2 Narrative Games and Instruments . . . . .	6
2.1.3 Retellings . . . . .	6
2.2 Simulation . . . . .	6
2.2.1 Epistemological Frameworks . . . . .	7
2.3 Story Sifting . . . . .	8
2.3.1 Sifting Patterns . . . . .	9
2.3.2 Sifting Heuristics . . . . .	11
<b>3 Methodology</b>	<b>12</b>
3.1 Simulation . . . . .	12
3.1.1 Dialogue . . . . .	12
3.1.2 Characterisation . . . . .	15
3.1.3 Blocking . . . . .	16
3.2 Sifting Patterns . . . . .	19
3.3 ‘Selecting the Unexpected’ . . . . .	22
3.4 Focalisation . . . . .	23
<b>4 Results and Discussion</b>	<b>24</b>
4.1 Quantitative Analysis of Heuristics . . . . .	24
4.2 Qualitative Analysis of Focalisation . . . . .	27
<b>5 Conclusions and Future Work</b>	<b>32</b>
<b>Appendix A: Questionnaire Extracts</b>	<b>38</b>
<b>Appendix B: Questionnaire Data</b>	<b>43</b>
<b>References</b>	<b>50</b>
<b>Bibliography</b>	<b>55</b>

# List of Figures

2.1	Pseudocode for <i>Hide and Sneak</i> 's drama-managed assumptions (ten Brinke, Linssen & Theune 2014). This is called whenever A needs a new search strategy. . . . .	8
2.2	Ryan's canonical 'Arson-Revenge' sifting pattern, as written in <i>Felt</i> (Kreminski, Dickinson, Wardrip-Fruin & Mateas 2022). This describes an <b>arsonist</b> who, having first been <b>harmed</b> , goes on to <b>hatch</b> and then carry out an <b>arson</b> attack that claims a <b>victim</b> - though not necessarily the intended <b>revengeTarget</b> ... . . . .	10
2.3	Leong, Porteous & Thangarajah's 'Icarus' and 'Cinderella' sifting patterns (left and right, respectively), as drawn in ARC SIFT. Both describe two characters that grow close, then fall apart; in the latter, they further reconcile. . . . .	10
3.1	A vertex, as stored in its JSON file. The author specifies its <b>DIEGETIC</b> description, and the <b>EDGES</b> connecting it to the rest of the graph, whilst granting characters agency in their pathfinding. . . . .	13
3.2	A verbal beat: capitalised are two static topics of conversation, dynamically combined into a single segue. While character-specific dialogue can be attached to any topic, this is nonessential. . . . .	13
3.3	An early, emergent interaction; extra edges are added to the script at runtime, allowing juxtapositions of topics unanticipated by the author. Whether <b>Bart</b> 's segue seems sincere, suspicious, or simply uncanny, is left to reader. . . . .	14
3.4	Pseudocode for <i>MiGS</i> ' search assumptions, written for comparison with Figure 2.1. . .	16
3.5	<i>MiGS</i> ' upstairs (upper left), downstairs (upper right), and their graph representations (below). Being unable to represent proxemics in continuous space, the model assumes all characters in the same room are engaged in the same conversation. . . . .	17
3.6	Graph representations of the stage's LOS (above) and ES (below). Notice how only voices carry up and down the stairwell; how characters are seen but not heard through the French windows to the balcony; how conversations behind closed doors are afforded a total privacy. . . . .	18
3.7	Pseudocode finding all of a trace's microstories for a given pattern. . . . .	19
3.8	The simulation's custom 'Rumour' sifting pattern, written in the syntax of <i>Felt</i> . This matches any <b>beat</b> of dialogue, binding its <b>actor</b> and an <b>audience</b> member for consideration by STU. . . . .	20
3.9	The simulation's custom 'He Said, She Said' sifting pattern. The <b>heSaid</b> beat describes a character, <b>himself</b> , who will <b>talk</b> about <b>hisTarget</b> to his audience; <b>sheSaid</b> describes a <b>distinct</b> actor behaving similarly <b>herself</b> . Note that, despite the naming convention, the pattern contains no constraints on either character's gender. . . . .	20

3.10	The simulation’s custom ‘Meet the Candidates’ sifting pattern. This is highly specific, describing a <b>focalCharacter</b> ’s knowledge of two of the three <b>candidates A, B, and C</b> . For both, the <b>focalCharacter</b> recalls a direct <b>interaction</b> with the <b>candidate</b> in question, and an <b>accusation</b> heard about them elsewhere. However, including a loosely defined <b>beat</b> in which extra, unexpected details can emerge renders queries for this pattern slow... . . . . .	21
4.1	Question I.i (Extract <b>a</b> is heuristic-preferred, with success rate 77.78%). . . . .	25
4.2	Question I.viii (Extract <b>a</b> is heuristic-preferred, with success rate 44.44%). . . . .	27
4.3	The extract’s correct timeline of events (top row), and its reconstruction by each participant. Solid outlines group the beats that overlap (or that participants say overlap) in each timeline. However, Question II.i’s wording is vague, with some reading “overlap” as ‘simultaneous’, others ‘simultaneous or consecutive’, so the actual orderings of beats (a.i)-(c.v) are far more instructive. . . . .	29
5.1	UI for an interactive build of <i>MiGS</i> (still in development). The right alignment of the player actions, the emboldening of character names, the italicisation of ‘stage directions’, <i>etc.</i> , all serve as visual story support, presenting the storyworld with more clarity than the original text console. . . . .	34
5.2	A focalised retelling from an NPC, mounted within the interactive build. To relate these retellings to the player’s own experience, avoiding any “random inclusion of a new character” (#6), this perspective comes from one of the three NPCs the player interacts with most over the playthrough. . . . .	34
5.3	A focalised and voiced retelling, generated by GPT-3.5. . . . .	36

# List of Tables

2.1	Prototypical rules for EN, inspired by role-play (Louchart & Aylett 2004 <i>b</i> ). . . . .	5
3.1	Properties in a <b>Microstory</b> 's property signature. . . . .	22
4.1	Summary of success rates in Part I of the questionnaire. . . . .	25
4.2	Exploratory data analysis on success rates, by participant. . . . .	26
4.3	Hypothesis testing for Table 4.2. . . . .	26
4.4	Responses to the debrief, by participant, and their correlations to overall success rates.	26
4.5	Edit distances of the participants' reconstructions from one another, and from the correct timeline. When beats are grouped as overlapping, the table takes the minimum values over all permutations of each group ( <i>e.g.</i> Timeline #1 has $2! \times 2! \times 3! \times 2! = 48$ possible permutations). . . . .	29

# 1. Introduction

*The best way to prepare for improvisation is to experience life.*

– TJ Jagodowski & Dave Pasquesi, *Improvisation at the Speed of Life*

Procedural storytelling demands to be taken seriously. There exists this funny tendency in the games industry to treat procedural content generation (PCG) as only a means of cost-cutting, whether that be the cost of artists labouring over bespoke assets, or the cost of storing those assets within a system’s memory. It’s a tendency that fails to appreciate the aesthetics of PCG, the ways generators convey meaning through not only their individually generated *artifacts*, but also their very procedures. There’s an authenticity and an uncanniness authored into the texts of *National Novel Generation Month* (*NaNoGenMo*); a fleeting beauty felt by audiences of randomly generated media (Grinblat 2017). It has always held its popular niches - from improv comedy to role-playing to tarot - but procedural narrative stands revitalised in the digital form.

In this experimental field, *emergent narrative* (EN) is a methodology more experimental (oft-times, calamitously so) than the rest. EN is a teleological approach to procedural storytelling, where stories form “bottom-up from the interaction of processes in underlying simulations” (Ryan 2018, p. 23). These *storyworld simulations* are also *character-centric*, inasmuch as the characters within all act autonomously. In a well-made simulation, stories should arise from players and NPCs simply exercising agency.

Having previously introduced the term, Aylett (2000) illustrates emergence in relation to football:

Conflicting aims and some constraints on allowable physical behaviour together with a limited time often - though by no means invariably - produces recognisable narrative structure... the new young player scoring on his debut; the player committing a reprehensible foul who injures himself seriously in the process; the talented but petulant player who retaliates when fouled, gets sent off, and loses his team a crucial match.

If you simulate 22 agents kicking a ball, you’ll only have a football game. *Authoring a simulation* means designing a set of models that interact as more than the sum of their parts. As a means of expression, it is not about writing a single story, but writing the code in which an entire possibility space of stories is implicit - a technical challenge with any number of aesthetic effects.

The EN methodology has been commercialised by *The Sims* (Maxis 2000), the franchise’s dramatic play driven by autonomous agents; *Crusader Kings III* (*CK3*; Paradox Development Studio 2020) infuses grand strategy with emergence. In educational contexts, the same storytelling philosophy can provide PSHE to kids (Aylett et al. 2005) as culture training to soldiers (Mateas & Ferguson 2014). However, the definitive work of EN remains *Dwarf Fortress* (Adams & Adams 2006), a staggeringly complex base-builder that has garnered academic praise for its emergent phenomena, spawned its own subgenre of ‘dwarflikes’, and even been mounted in New York’s Museum of Modern Art (Antonelli 2011).

*Retellings* (Eladhari 2018) are yet another measure of the game’s success. *Dwarf Fortress*’ emergentist approach makes no consideration of overarching themes or plot, yet players invest great meaning in its simulated events. That so many will rewrite and republish their experiences as blog posts, ‘Let’s Plays’, or other such media surely speaks to a deep engagement with the game. Retellings can also vary greatly in content and tone, ranging from grounded, transgressive play about homelessness in *The Sims* (Burkinshaw 2009), to ironically detached *CK3* walkthroughs on cannibalising the Pope (Livingston 2020). Even exaggerating an in-game achievement to friends counts, in that it adds embellishment and a rhetorical intent to a computer-generated experience.

These retellings rely on *story sifting*: the underlying (and not necessarily outward-facing) act of recognising emergent narratives. Simulations might produce comprehensive timelines of events, but by their rules-based nature these are little more than clinical sequences of cause-and-effect. To retell their experiences in other mediums, consumers and critics alike must necessarily ‘sift out’ the subsequence(s) worthy of retelling.

Ryan, Mateas & Wardrip-Fruin (2015) ask, can story sifting be proceduralised? The open problem has been approached with sifting patterns, heuristics, and other techniques in recent years, garnering interest for good reason. A retrospective *story sifter*, capable of reviewing a complete sequence of events and surfacing story structures that, even incidentally, emerge, offers immediate benefits. Sifters can better tutorialise a simulation (scoping out its expressive range of interactions), and better curate any narrative artifacts. *Dwarf Fortress*’ Legends mode summarises each civilisation’s history. Meanwhile, its *Dwarf Grandpa* mod (Garbe 2018) aims for specificity, parsing a crude syntax of sifted events into player-readable character biographies.

Though retrospective sifting emphasises the moments that have emerged, it can’t guarantee a storyworld is ‘storyful’ in the first place. As much as this paper argues for the expressive potential of EN, it is not without difficulty. Work in the field must engage with many common design problems,<sup>1</sup> to say nothing of the inherent precarity of the emergentist approach. It remains a truism that “one of the risks of emergent narrative is that it may not emerge” (Aylett 1999, p. 85).

At the other extreme of procedural narrative, then, are the more reliable, ‘top-down’ methodologies. Some are *story-centric* (‘slotting together’ common narratological structures), others *author-centric* (modelling an author’s own writing process); all involve greater degrees of *explicit creation* than EN, guided by static data rather than dynamic procedures. These techniques better accommodate the plot possibilities of story generation, including the possibility of replayable mysteries.<sup>2</sup> In *We Happy Few* (Compulsion Games 2018), for instance, different playthroughs are structured around different omissions, inconsistencies and other sleights of hand to maintain an engaging mystique. The game generates *dirty narratives*: stories in which the reader cannot simply accept the stated facts, but must actively interrogate them (Epstein 2019).

Dirty narrative can also be achieved through *focalisation*, the limiting of a narrative to a certain character’s perspective. By presenting a linear sequence of events through the eyes of different characters, each account riddled with its own tangents and omissions, the story is reframed as a puzzle to be pieced together - and although this device has rather archaic roots in the epistolary, its scope for tension and intrigue is self-evident. It has been embedded in the popular consciousness by whodunnits (Christie 1943), horror (King 1974), and - of course - *Rashomon* (Kurosawa 1950), a film so toweringly influential that the device is best colloquialised as ‘the *Rashomon* effect’.

<sup>1</sup>The narrative paradox, *Tale-Spin* effect, and ‘10,000 Bowls of Oatmeal’ problem will all be covered in due course...

<sup>2</sup>There’s a strong argument to be made for the procedural murder mystery being a ‘holy grail’ of PCG (Maher 2013). There’s also an argument that, requiring the emergence of both a murderer with means, motive, and opportunity, and a detective capable of catching them, character-centric efforts like Sheldon Klein et al.’s murder mystery generator (1973) constitute some of EN’s “most famous failures” (Ryan 2018, p. 3).



Restated: the *Rashomon* effect occurs when an author refuses to impose a single, objective truth on their story, instead constructing each character’s own emotional truth ‘bottom-up’. It seems almost uniquely well-suited to EN. Sifting from the subsequence of events observed by a focal character provides a limited, local perspective on their storyworld, yet together these perspectives form a fragmentary account of the simulation as a whole. This paper envisions this process of *focalised story sifting* as an emergentist approach to dirty narrative, and more generally as a refinement to current sifting practices.

The project therefore explores three key research questions:

**RQ1.** How might existing sifting techniques be focalised?

**RQ2.** What meaning might focalised story sifting create in an individual EN artifact?

**RQ3.** What meaning might focalised story sifting create across a range of EN artifacts?

As the user encounters a range of outputs, they build a ‘mental model’ of the underlying generator. Should an artifact feel improbable or ephemeral or uncanny (Short 2019), these meanings would not be explicit in the artifact itself, but implicit in one’s understanding of how it was generated. This is why the project aims, in implementing focalisation (**RQ1.**), to both explore how focalised sifting enhances individual narratives (**RQ2.**), and extrapolate their cumulative effect (**RQ3.**).

Chapter 2 will review the formal literary theory of EN, and some instructive case studies in simulation and sifter design. Chapter 3 covers the development of *Men in Grey Suits*, this project’s own non-interactive simulation, and its accompanying retrospective, focalised story sifter. Results will be evaluated by questionnaire, detailed in Chapter 4, wherein participants qualitatively analyse the sifter’s focalised narratives, as well as quantify their ‘interestingness’. Finally, Chapter 5 will synthesise this experimental evidence with a broader, more artistic statement on the ‘authenticity’ of story sifting, and explore how future work might integrate focalisation into other nascent techniques (including story support and incremental sifting).

## 2. Background

### 2.1 Narratology

In a tradition tracing back to Aristotle (330 B.C.), humans have long tried to deconstruct and define the very essence of stories. *Narratology* is the formalised study of their underlying frameworks. Now, in upholding that tradition and putting forth a preferred definition of emergence, this section must introduce some key narratological concepts. After all, to argue EN is fundamentally divergent from static literature, the paper will first demand an intuition toward static literature itself.

Particularly important are the terms *fabula* and *syuzhet*. They are, respectively, the sequence of events that constitute a story, and the subsequence in which they are told. The precise taxonomy varies author to author, but the underlying idea of ‘layers of narrative’ is resonant - the emergentist approach jointly generates the low-level chronologies and high-level representations of stories.

Genette (1980) further breaks the *syuzhet* into five *analytical segments*:

*Order* How the sequencing of the *syuzhet* relates to the chronology of the *fabula* (*e.g.* does the author use flashforwards, or flashbacks?).

*Duration* How the reading time of the *syuzhet* relates to the passing of time within the *fabula* (*e.g.* does the author use a descriptive pause to elaborate on a pivotal moment?).

*Frequency* How often the *syuzhet* returns to a recurring event in the *fabula* (*e.g.* does the author describe each instance individually, or all at once?).

*Voice* Where the narrator of the *syuzhet* exists relative to the events of the *fabula*.

*Mood* How the narration of the *syuzhet* regulates information from the *fabula*.

The distinction between these last two terms might seem murky, with other critics combining them into a single “point of view” (Brooks & Warren 1943, p. 146), but Genette distinguishes the former as “who speaks” and the latter as “who sees” (p. 186). One modality of mood, for instance, would be the relationship between the *syuzhet*’s telling of events and each character’s own interpretation. Omniscient narratives are *nonfocalised*; *internal focalisation* gives a narrator all information known to a ‘focalised’ character; *external focalisation* restricts access to some subset thereof.<sup>1</sup> Crucially, these degrees of *focalisation* exist independently of voice - narration in the third-person can nonetheless be internal, narration in the present tense nonetheless omniscient.

While Genette’s work does not fully consider how analytical segments might bleed into one another (Guerlac 1980), an overly simplistic model can lend itself well to procedural generation. However, at its high level of abstraction, there is concern that his formalist-structuralist approach to narrative is “difficult for direct computational application” (Louchart & Aylett 2004a, p. 513). For a working definition of EN, one must therefore look further afield...

---

<sup>1</sup>Borrowing Genette’s own example, consider how Agatha Christie narrates Poirot: though she presents the reader with many of his observations and inner thoughts, his grand theory of ‘whodunnit’ must be externalised through dialogue.

### 2.1.1 Emergent Narrative

How interactive should interactive narrative be? The naïve response is “as much as possible!”, the player free to shape their surroundings in both the short- and long-term. Then again, true freedom includes the freedom to subvert, the freedom to accidentally miss or deliberately derail a key event; an author might have to impose constraints to ensure a satisfying story structure. This inherent tension between agency and authorial intent is Aylett’s *narrative paradox* (2000), to which EN was envisioned as a solution.

More so than classical literature, or even branching narrative, the emergentist approach is rooted in role-playing (Louchart & Aylett 2004b). Both tabletop role-playing games (TTRPGs) and live action role-playing (LARP), for instance, cast the player not as an author or spectator but as a *participant*; both unfold iteratively, guided largely by their characters, unobtrusive in their constraints on agency. Just as the storyworld of a TTRPG is bound by a ‘magic circle’ of social norms and dramatic sensibilities (Kawitzky 2020), the user’s *social presence* and sense of self in interactive narrative regulates subversive behaviour from within (Aylett 2000). Most strikingly of all, LARP’s customary debriefing session, where players retell their own experiences and the motivations behind them, shows how emergent narratives may yet be recomposed in other mediums. In narratological terms, the debrief exists as syuzhet to the fabulae of player experiences, each LARPer curating their most illuminating moments and sharing them as an exercise in multiple internal focalisation.<sup>2</sup>

<b>Rule 1</b>	The story must be perceived as a process, in a dynamic rather than analytical way.
<b>Rule 2</b>	The plot is at best hypothetical. It is composed of interrelated hypothetical plot elements.
<b>Rule 3</b>	Plot events are written before performance, certain types of events outcomes can be controlled.
<b>Rule 4</b>	Narrative authoring is done before rather than during the performance.
<b>Rule 5</b>	The storyline is created as the different characters interact with each other, in real-time.
<b>Rule 6</b>	Environments and props must be complementary to the hypothetical plot or plot elements.
<b>Rule 7</b>	Characters must be written for interaction, implying rich, deep and emotionally engaging traits.
<b>Rule 8</b>	The user must be considered as a participant to the performance rather than a spectator or author.
<b>Rule 9</b>	The user should be encouraged to act in role and should be penalised for not doing so.
<b>Rule 10</b>	User enjoyment and interest are the only reasons for the performance to take place.

Table 2.1: Prototypical rules for EN, inspired by role-play (Louchart & Aylett 2004b).

Similar parallels exist with improvisational theatre (improv). The autonomous agents of a simulation serve a dual purpose: each embodies an *in-character* (IC) role, and the *out-of-character* (OOC) actor playing that part.<sup>3</sup> Swartjes (2010) recognises this duality, arguing agents should inform their IC decision-making with an OOC *actor-awareness* of drama management. Like improv, he sees EN as an unscripted act of co-creation, with players putting their agency towards active collaboration if approaching it as “computer-mediated performance” (p. 105). Inasmuch as others like Ryan (2018) would place EN in the genre of nonfiction, each fabula of events having “actually happened” (p. 6), as it were, in our paper’s view this is only the nonfiction of ‘real’ actors playing imaginary characters.

These comparisons are instructive, but they do not constitute a definition. Louchart & Aylett’s investigation (2004b), an early attempt at formalisation, characterises EN by the rules in Table 2.1. There is arguably a lack of clarity here, with some trying to define the methodology in and of itself (**Rules 1, 5, 7**), but others defining it as experienced by the user (**Rules 8, 9, 10**). Moreover, with the investigation so motivated by the narrative paradox, Table 2.1 fails to consider non-interactive

<sup>2</sup>On top of its degree, Genette (1980) classes focalisation as *fixed* (on one character), *variable* (across many characters, to narrate distinct events), or *multiple* (many characters, overlapping events). *Rashomon*, too, fits this final category.

<sup>3</sup>In improv, IC interactions are termed *intradiegetic*, OOC interactions *extradiegetic* (Landert 2021; see Chapter 5).

EN (where the paradox clearly does not apply). Ryan (2018) proposes a more general definition, but makes no mention of how participants and agents embody both IC and OOC roles. To that end, our definition of *emergent narrative* is a slight elaboration on Ryan’s own, intended to better evoke its origins in improv and role-play:

*The methodology characterised by computational systems in which narrative emerges bottom-up from the interaction of processes in underlying simulations, wherein human participants or autonomous agents embody any characters (or, alternatively, the actual narrative material produced by this method).*

### 2.1.2 Narrative Games and Instruments

Interestingly, the above definition does not require a user feel any social or dramatic presence. Whilst some works of EN (commercial works especially) would still be grouped with other narrative games, there are others that exist not so much to be ‘won’ as to generate stories for stories’ sake. These are instead known as *narrative instruments*.

As imagery, this term speaks less to the methodology of EN, and more to the underlying purpose. Much like their musical analogues, narrative instruments are “to be used for expression first and foremost” (Kreminski & Mateas 2021, p. 501). The appeal of learning such a system is the appeal of practice and experimentation, of honing a mental model for how it works, of turning that intuition to inward- or outward-facing performance. The metaphor even reconciles the narrative paradox: in much the same way that, say, bagpipes are striking in their expressive quality, a narrative instrument is no less freely playable for preserving an author’s desired ‘sound’.

### 2.1.3 Retellings

Note that Eladhari (2018) positions *retellings* - how one recounts any experience with a game or instrument - as an extra narratological layer of EN, one above the syuzhet. Retellings range from off-hand anecdotes to prosaic journalings of quests, can be user- or computer-generated; indeed, while Eladhari restricts her attention to interactive narrative, one could equally retell the surprising behaviours of a non-interactive generator. Crucially, they are understood not as mere metatext, but synecdoche for the storyworld itself.

Chapter 1 tries to quantify *Dwarf Fortress*’ success by the number of retellings (and hence, the user engagement) it generates. This is, however, something of an oversimplification, in that it ignores the role of irony as a confounding variable (Sych 2020). Retellings can just as well be centred around uncanny behaviours or unforeseen bugs, their tones ranging from good-humoured to flatly critical. They are far better suited to qualitative analysis, with researchers using ‘cherry-picked’ successes and failures of EN to evaluate their work overall (Kreminski, Dickinson, Wardrip-Fruin & Mateas 2022).

## 2.2 Simulation

Knowing the ‘what’ and ‘why’ of EN, there remains a question of ‘how’. *Storyworld simulations* are the generators wherein characters act, each run returning a game log (or *trace*) of simulated events (*beats*). Designing a storyworld is an act of *implicit creation* (Spierling 2007), authoring the processes by which interactions emerge without being overly prescriptive about the interactions themselves.<sup>4</sup>

---

<sup>4</sup>It’s left unsaid in Chapter 3, but implicit creation is genuinely challenging. An author will generally have to extrapolate their simulation’s expressive range from the traces it generates, searching for unwanted emergent behaviours largely through trial and error. On finding one, they would either have to *debug* the code, applying constraints until the behaviour becomes impossible, or - in the spirit of improv - say ‘Yes, and...’ and add nuance *until the behaviour instead makes sense* (Swartjes 2010). Both attitudes have informed the development of this paper’s own simulation.

Mateas & Ferguson (2014), drawing from the former’s pioneering work on *Faade* (Mateas 2002; Procedural Arts 2005) and *Prom Week* (McCoy et al. 2011), imagine an ideal simulation built from ‘social LEGOs’. The LEGOs in question would be lightweight modules of code, that are:

*Reusable* Easily migrated into new simulations, easily customised.

*Mixable* Readily combined with other modules, with minimal authorial burden.

*Symmetric* Usable by both player and non-player characters.

*Real-Time* Run over continuous time intervals.

Mateas clarifies that these are not strict criteria. His emphasis on real-time over turn-based simulation, in particular, reads more as personal preference than the be-all and end-all of social simulation design. Nevertheless, from a research perspective, designing for reusability and mixability makes it easy for others to replicate and expand upon one’s own results, with symmetry encouraging future work in interactive and non-interactive contexts alike.

The simplification of social interactions into a gamified *social physics* is also, to some extent, motivated by the *Tale-Spin effect* (Wardrip-Fruin 2007). Named for one of the earliest works of EN (Meehan 1977), the effect occurs when a generator fails to externalise its processes, making it seem unsophisticated to the outside observer. The short stories of *Tale-Spin* are an abject lesson that complex code does not necessarily produce complex artifacts.

### 2.2.1 Epistemological Frameworks

As an extended example in simulation design, consider the problem of simulating character knowledge. Each agent would require their own *epistemological framework*, one that holds their personal, mental model of the storyworld. As with any form of authorship, creating such a module means drawing inspiration - direct or indirect - from prior works.

Gervás (2013) offers a minimalistic approach. In focalising the traces of finished chess games, his research naturally requires a model of what each in-game character has or has not seen. For a piece with discrete position  $\mathbf{x}(t) \in \mathbb{Z}_8 \times \mathbb{Z}_8$ , Gervás defines a fixed *range of perception*  $N = 2n - 1$ ,  $n \in \mathbb{N}$ , such that the piece will have seen an event at time  $t_0$  if and only if it took place at one of the  $N \times N$  coordinates  $\{\mathbf{x}(t_0) + (i, j) : |i|, |j| \in \mathbb{Z}_n\}$ . There is no simulation here per se, but it is nonetheless a valid approach to perspective: given the position and orientation of an agent, one can just calculate whether an event occurred in their field of view, and therefore whether they saw it, entirely post-hoc.

This model is perfectly suitable for chess, in that the game is played from the top down; no piece needs to make observations at runtime, since no piece acts autonomously of its player. An actual simulation might add nuance, with agents informing their actions and strategies from an underlying epistemological framework. Centred around a literal cat-and-mouse chase, *Hide and Sneak* (ten Brinke, Linssen & Theune 2014) is a very physical implementation of this concept. Its choice of ‘visibility rules’ - characters can see all other characters in their current location, characters can see other locations through open doorways, *etc.* - indeed sets the stage for a game of hide and seek, the seeker ruling out visibly empty rooms in a process of elimination.

What happens, though, when a character hides in a room that has already been ruled out? Rather than just stopping indefinitely (an unwanted emergent behaviour), ten Brinke, Linssen & Theune allow the seeker to dynamically update their search strategy via epistemological framework. If the character has an assumption of where an object (*e.g.* another character) is, they will go to that location. When this is ruled out, they then search ‘believable’ locations (*i.e.* those they haven’t entered in the last 10 turns) until they find what they’re looking for - or otherwise form a new assumption.

Moreover, incorporating drama management, agents use an omniscient OOC knowledge of the world state to track their IC enemies. A purist might argue this is an unrealistic level of actor-awareness,<sup>5</sup> but it nonetheless lets agents steer themselves into confrontations that their character would not reasonably know how, or be expected, to avoid (see Figure 2.1).

```

if (agents A and B are enemies && [OOC] A knows B's LOCATION
    && [IC] A does not assume B is at LOCATION
    && [IC] A does not assume OBJECT is anywhere other than LOCATION
    && [IC] A believes OBJECT could be at LOCATION)
    A assumes OBJECT is at LOCATION;

```

Figure 2.1: Pseudocode for *Hide and Sneak*'s drama-managed assumptions (ten Brinke, Linssen & Theune 2014). This is called whenever A needs a new search strategy.

*Talk of the Town* (Ryan, Summerville, Mateas & Wardrip-Fruin 2015) affords agents even more subtlety to how they (mis)use their knowledge. Here, each NPC's mental model encompasses not only the facts of life in their titular town, but a metaknowledge of where those facts were learned. Information propagates and deteriorates; characters might misunderstand an event they've observed, misremember its less salient details in hindsight, or outright mislead one another if they can get away with it. The emergent possibilities are inspiring - but not without their challenges. Aside from the non-negligible computational cost, Ryan & Mateas (2017) reflect that their work relies on many finely-tuned parameters. A research project like ours, simulating epistemological frameworks only as means to an end, might prefer to minimise scope by streamlining some of *Talk of the Town*'s core mechanics.

## 2.3 Story Sifting

In narratological terms, traces are just fabulae; to parse a simulation's output into an actual story requires substantially more effort. Consider how *Talk of the Town* has been adapted into *Bad News* (Ryan, Summerville & Samuel 2016), a computationally assisted piece of theatre set in a simulated American township, generated afresh at the start of every show. For 45 minutes, an actor multi-roles as the town's NPCs, improvising conversations with the audience much as the game master of a TTRPG would. This is guided by a *Wizard-of-Oz technique* (Dahlbäck, Jönsson & Ahrenberg 1993): their performance is reliant on a separate 'man behind the curtain' running a concurrent, real-time analysis on the simulation's trace from backstage. The *wizard* finds prompts for the actor, reading between the lines of the fabula to spot the rumours that emerge - anything from an urban legend about the town undertaker to allegations of nepotism at a local restaurant (Ryan 2018).

These wizard-actor interactions prove an illustrative example of *story sifting* (or *story recognition*), the process of identifying quality EN. With thousands, maybe millions, of narrative events to sift, though, even experienced wizards must settle for whatever they can find, in whatever time they have. Conversely, a computer could systematically search the entire event sequence, but without any human understanding of what to search *for* - as plot considerations are alien to its character-centric approach, a simulation can't understand what makes its own stories 'storyful'. A *story sifter* that operationalises this subjective search for narrative therefore represents an open problem in the field (Ryan, Mateas & Wardrip-Fruin 2015).

---

<sup>5</sup>Improvvisers aren't omniscient in real life; their OOC plot considerations are still informed by subjective experience...

The process is central to *curationism* (or *curationist emergent narrative*): the philosophy that the raw data of EN exists to be sifted and subsequently exhibited. As argued by Ryan (2018), the great successes of EN are those “mounted in full-fledged media experiences” (p. 229), either within the storyworld (*feedback curations* like *Dwarf Fortress*’ Legends mode) or without (*feedforward curations* like *Bad News*). Failures like *Tale-Spin* produce only unceremonious prose for academic papers.<sup>6</sup> Under curationism, the story sifter finds retellings of the fabula, to be fed into the syuzhet of a media experience.<sup>7</sup>

While all sifters, by definition, impose narratological structure, few go so far as to operationalise specific narrative techniques. Bae & Young (2008) experiment with order, sifting out moments for flashback or foreshadowing. Gervás (2014) uses his prior epistemological framework to produce variable focalisation retellings of chess games from the perspectives of pawns. These are formative efforts, certainly, but they understand that the most high-quality retelling of a sequence of events is not always the linear one. It’s a subtlety this paper feels has been somewhat lost as sifting techniques have been refined over the years.

A note on terminology - unless stated otherwise, “sifting” is used as shorthand for *retrospective* story sifting. Analogous to a debrief, this is run post-hoc on a completed trace. *Incremental* sifters that operate in real time to steer live simulations toward more storyful outcomes (Kreminski, Dickinson & Mateas 2021) remain an open problem for future work (see Chapter 5).

### 2.3.1 Sifting Patterns

*Sifting patterns*, the catch-all for queries with narrative structure that might be applied to a trace, are common in story sifting. When a sequence of simulated events is queried, the sifter returns all subsequences satisfying the pattern’s conditions. These matched subsequences are termed *microstories* (Kreminski, Dickinson, Wardrip-Fruin & Mateas 2022).

The patterns of early systems like *Sheldon* (Ryan 2018) and *Felt* (Kreminski, Dickinson & Wardrip-Fruin 2019) are text-based, characterised by the following:

1. *Logic Variables* ‘Slots’ that bind to simulated entities (*e.g.* story beats, characters, locations).
2. *Constraints* Conditions satisfied by the logic variables (*e.g.* ‘beat ?e has attribute a with value ?v’, ‘beat ?e precedes beat ?f’).

In *Felt*’s syntax, the prefix ?- distinguishes logic variables ?x from constant values x. Constraints of the form ‘beat has attribute with value’ are written in square brackets [] (called *bindings*), whereas all other constraints will be denoted by () (called *relations*).

Again, a purist might suggest that sifting patterns impose too much top-down structure on the emergentist framework, that a computer finding stories with human-authored queries is no different from a human just writing those stories directly. As a criticism, it’s at its most acute when the patterns involved are too explicit - but this is of course avoided with careful authorship. *Felt* will allow a writer to (against their better instincts) be overly specific; notice, though, that where Figure 2.2 describes a *victim*, all other traits go unmentioned, still free to emerge over the course of simulation.

---

<sup>6</sup>Though if there’s one criticism to be made of Ryan’s thesis, it’s that his definition of “full-fledged” is rather fuzzy. Personally, I completely agree that narrative artifacts printed in **MONOSPACED BLOCK CAPITALS** are hardly compelling, but the counterargument should amount to more than “This is not really a real media genre, is it?” (p. 202).

<sup>7</sup>Note also that Ryan, viewing EN as nonfiction, prefers a narratological framework separating “the events of a world, the events that are recounted, and those events *as* recounted” (p. 25). While this lends itself to the three-step process of simulation, sifting, then *story support* - how one presents each narrative to the end user (see Chapter 5) - papers focused on only those first two steps are better served by a two-layered model.

```

(eventSequence ?harm ?scheme ?arson)
  [?harm tag harm]
  [?harm actor ?revengeTarget]
  [?harm target ?arsonist]
  [?scheme eventType hatchRevengeScheme]
  [?scheme actor ?arsonist]
  [?scheme target ?revengeTarget]
  [?arson eventType setFire]
  [?arson actor ?arsonist]
  [?arson target ?revengeTarget]
  [?arson victim ?victim]

```

Figure 2.2: Ryan’s canonical ‘Arson-Revenge’ sifting pattern, as written in *Felt* (Kreminski, Dickinson, Wardrip-Fruin & Mateas 2022). This describes an **arsonist** who, having first been **harmed**, goes on to **hatch** and then carry out an **arson** attack that claims a **victim** - though not necessarily the intended **revengeTarget**...

More recent efforts better encourage implicit creation. Kreminski, Wardrip-Fruin & Mateas (2020) propose an example-driven model, where the author provides a set of handwritten microstories and the sifter returns any similar moments from the trace. At an even higher level of abstraction, ARC SIFT (Leong, Porteous & Thangarajah 2022) adopts line drawings as its sifting patterns, using dynamic time warping to fit them to the ebb and flow of character relationships (see Figure 2.3). More than just reducing the authorial burden of writing patterns directly - an accomplishment in and of itself - these sifters are such that a writer might be actually be surprised by their sifted microstories. Their advanced techniques lie outside the scope of this research, but they make the necessary case that pattern-based sifting is not strictly antithetical to EN.

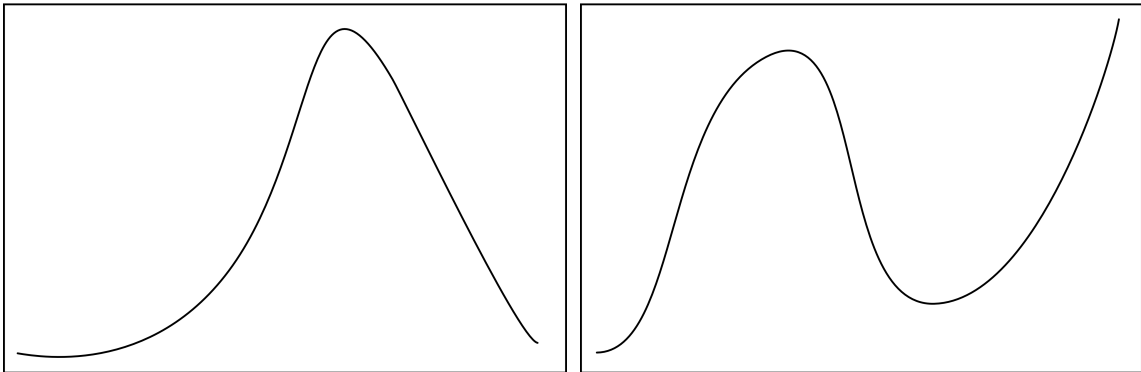


Figure 2.3: Leong, Porteous & Thangarajah’s ‘Icarus’ and ‘Cinderella’ sifting patterns (left and right, respectively), as drawn in ARC SIFT. Both describe two characters that grow close, then fall apart; in the latter, they further reconcile.



### 2.3.2 Sifting Heuristics

Another high-level tool would be the *sifting heuristic*, proposed by Ryan as “a policy for sifting, rather than an explicit pattern” (2018, p. 250). These can be implemented atop sifting patterns in a two-part sift, where matched microstories are surfaced to the user only if they are (by the heuristic’s standards) sufficiently entertaining. Kreminski, Dickinson, Wardrip-Fruin & Mateas (2022) take this approach with SELECT THE UNEXPECTED (STU) - arguably the first major example of a sifting heuristic in action, and one that our research relies on extensively.

As the name suggests, STU’s underlying policy is one of maximising surprise. Given a trace’s set of microstories, the authors first define any contextual *properties* - character traits, relationships, *etc.* - of narrative interest. Each microstory is then analysed in context, the properties it satisfies together forming its *property signature*. STU counts how many microstories from each pattern fulfil each property, uses these counts to make a statistical evaluation of each property’s likelihood, and thereby surfaces the microstories with the least likely (*i.e.* most surprising) signatures.

Crucially, properties not explicitly specified by sifting patterns can still emerge; if a pattern has three roles for a character, their lover, and their rival, then microstories in which one character functions as both lover and rival are tagged with the ‘same-character property’. Kreminski, Dickinson, Wardrip-Fruin & Mateas find that STU will consistently surface microstories with such unexpected character connections, just as it will unexpected interactions.

At the start of this section, the story sifter was set out as a means of searching an entire sequence of simulated events. STU, *Felt*, and other such efforts indeed work from complete knowledge of their storyworlds, and indeed with complete control over what to surface. If sifting represents the transitional phase from fabula to syuzhet, from objective history to retelling, then these sifters are analogous to the omniscient, omnipotent storyteller - and therefore at odds with the emergentist approach.

This literature review has established EN as an inherently collaborative process, grounded in improvisation and role-playing. It has identified a clear research interest in simulating ‘fallible’ characters, the earlier epistemological frameworks allowing authentic misunderstandings to emerge. That contemporary sifters do not afford, or even attempt to afford, that same collaboration or fallibility creates a certain dissonance. In response, this research presents a refinement to story sifting, co-created amongst OOC actors in much the same way a debrief is amongst OOC LARPerS. Like previous projects (Gervás 2014, and also Hoek, Theune & Linssen 2014), ours sifts out focalised stories, but takes the novel step of autonomising that sifting process across the actors themselves.

Put formally, *focalised story sifting* means sifting through only the subsequence of events observed by a character, according to a (non-omniscient) actor-awareness of what makes a compelling story. On a purely technical level, Chapter 3 describes a focalised implementation of STU, where each agent surfaces microstories by their own statistical definition of surprisingness. On the level of the individual artifact, Chapter 4 will explore the explicit narratological qualities of the sifter’s outputs (with a particular authorial interest in whether any dirty narratives emerge). Ultimately, though, the paper is concerned with the technique in and of itself, with Chapter 5 interrogating whether it is somehow ‘more true’ to the established vision of EN.

## 3. Methodology

### 3.1 Simulation

Focalised story sifting, of course, requires something to sift. With an overall aim established, the first research objective here is writing the right simulation, one that generates traces rich with EN. Each focal character will - clearly - need their own epistemological framework. Moreover, to achieve an emergent *Rashomon* effect through multiple internal focalisation, the agents' internal logic must afford different characters different interpretations of the same event. These criteria lend themselves to a *subjective simulation design* (Kim 2018), concerned not with the absolute truth of the storyworld but the emotional truths of the characters within.

The resulting application, *Men in Grey Suits (MiGS)*, is created with this philosophy in mind. Set in Westminster at the very height of the Cold War, the simulation gathers a cast of Conservative MPs at an exclusive soirée, each using the evening to gossip about one another. There is an overt, utility-based consideration to any exchange of information, a model of social physics that should reflect the characters' deeply transactional relationships. The simulation also, quite deliberately, refuses to verify any given rumour, the truth only mattering inasmuch as someone else believes it. There is even a thematic resonance to *MiGS'* inevitable *perceptual collapse*, the cumulative effect that comes when the user completes their mental model of a generator, its artifice of infinite possibilities falling away as they realise “this is all there is” (Kreminski 2023a, p. 2). This collapse should speak to an emptiness at the heart of these characters, this world, this mode of politics; that their jockeying amounts to little more than hollow spectacle, changing nothing.

All of this is to say, *MiGS* has been written with an explicit (and yes, explicitly political) authorial intent. The emergentist approach articulates themes through both its artifacts and the modules that generate them - in this case, modules for dialogue, characterisation, and movement. Throughout the following, mathematical models and code snippets will be kept to a minimum, edited for clarity rather than accuracy to the original application.

#### 3.1.1 Dialogue

Underpinning all of this work is a use of *waypoint narrative* (Short 2016). In her story *Glass* (2006), Short models dialogue as a directed graph  $D$ , with vertices  $v \in V(D)$  as topics of conversation, and edges  $e \in E(D)$  (weighted by some  $w : E \rightarrow \mathbb{R}$ ) the segues in between. Using the A\* algorithm, NPCs manoeuvre discussions from  $v_\alpha$  to their choice of  $v_\beta$  by segueing along the corresponding minimal (weighted) path on  $D$ . Despite being explicitly authored, waypoint narrative is in some sense a form of character-centric storytelling, driven as it is by the pathfinding strategies of autonomous agents.

```

"C-7": {
  "DIEGESIS": "asserting that the only people testifying
    against the cabinet have their own eyes on the leadership",
  "ROLES": {
    "STAR WITNESS": {
      "IDENTIFICATIONS": [ ],
      "IMPLICATIONS": [ "LEADERSHIP BID" ]
    }
  },
  "EDGES": [ "C-4", "C-8", "C-9" ],
  "MODE": "RUMOUR"
},

```

Figure 3.1: A vertex, as stored in its JSON file. The author specifies its DIEGETIC description, and the EDGES connecting it to the rest of the graph, whilst granting characters agency in their pathfinding.

```

Sir Geoffrey is talking with Peter, Harry, Bart, and his own aide.
They are STATING THAT THESE ANTI-COMMUNISM HEARINGS ARE ALL PART
OF A CONSPIRACY TO UNDERMINE THE PARTY.
Sir Geoffrey segues to ASSERTING THAT THE ONLY PEOPLE TESTIFYING
AGAINST THE CABINET HAVE THEIR OWN EYES ON THE LEADERSHIP.

```

Figure 3.2: A verbal beat: capitalised are two static topics of conversation, dynamically combined into a single segue. While character-specific dialogue can be attached to any topic, this is nonessential.

Though it is not discussed by Short, the real elegance of waypoint narrative lies in how it allows for knowledge propagation. Letting  $C$  be the set of all characters in a simulation, and supposing some character  $i \in C$  starts out ‘knowing’ only some vertices  $V_i \subseteq V(D)$ , their mental model of graph  $D$  will be its induced subgraph  $D[V_i]$ . On hearing a segue to any other  $v^* \in V(D) \setminus V_i$ , this model is extended to  $D[V_i \cup \{v^*\}]$ , the character incorporating new information into their epistemological framework.

Dialogue is instantiated in *verbal beats*  $\nu = (v, a, A, R)$ , where a character  $a \in C$ , talking with an audience  $A \subseteq C$ , segues their conversation to topic  $v \in V_a$ . *Rumours*  $\rho = (r, c, I) \in R$ , moreover, match a role  $r$  to an accuser  $c \in C \cup \{\varepsilon\}$  (possibly ‘empty character’  $\varepsilon$ ), and a set of implications  $I$  (possibly  $I = \emptyset$ ). *Congruent* rumours are denoted

$$\rho_1 \simeq \rho_2 \Leftrightarrow \rho_1[0] = \rho_2[0] \text{ and (at least) two elements of } \{\rho_1[1], \rho_2[1], \varepsilon\} \text{ are equal}$$

(the notation  $x[n]$  describing the  $n$ th entry of a tuple  $x$ ; indexing begins at 0).

The vertex described in Figure 3.1, for instance, would suggest that the STAR WITNESS ready to testify in the upcoming McCarthyite hearings is only doing so to advance their own LEADERSHIP BID. This IMPLICATION carries to the beat instanced in Figure 3.2, where

$$R = \{(\text{STAR WITNESS}, \varepsilon, \{\text{LEADERSHIP BID}\})\}$$

- noting that Sir Geoffrey does not accuse any character by name, since this particular vertex does not allow an IDENTIFICATION to be made. The statement is therefore (pairwise) congruent with any others of his about the STAR WITNESS.

With a model of how characters receive information, the next step is deciding when they believe it. For any sequence of beats  $\mathcal{V} = \nu_0 \cdots \nu_N$ , let

$$\begin{aligned} R_{\mathcal{V}}(r) &= \{\rho : \exists n \in \mathbb{N} \text{ such that } \rho \in \nu_n[3], \rho[0] = r\}, \\ n_{\mathcal{V}}(r) &= \min \{n : \exists \rho \in R_{\mathcal{V}}(r) \cap \nu_n[3] \text{ such that } \rho[1] \neq \varepsilon\}, \end{aligned}$$

meaning all rumours about role  $r$  in sequence  $\mathcal{V}$  can be matched to an accusee

$$c_{\mathcal{V}}(r) = \begin{cases} c & \text{if } n_{\mathcal{V}}(r) \text{ exists} \\ \varepsilon & \text{if otherwise} \end{cases} \quad \text{for } \rho \in \nu_{n_{\mathcal{V}}(r)}[3] \text{ such that } \rho[0] = r, \rho[1] = c$$

by the simulation's *belief function*

$$\rho_{\mathcal{V}}(r) = (r, c_{\mathcal{V}}(r), I_{\mathcal{V}}(r)), \text{ where } I_{\mathcal{V}}(r) = \bigcup_{\rho \in R_{\mathcal{V}}(r)} \rho[2].$$

Suppose, now, that a character  $i$  works from a *script*  $\mathcal{D}_i = (D[V_i], \mathcal{V}_i)$ , where  $\mathcal{V}_i$  contains all beats they have personally heard, in the order they heard them. Under their unique belief function  $\rho_{\mathcal{V}_i}(r)$ , the role  $r$  will be associated with  $I_{\mathcal{V}_i}(r)$ , the sum total of implications they know about it thus far.  $c_{\mathcal{V}_i}(r)$ , meanwhile, describes their *anchoring bias*: once  $i$  hears the role matched to a non-trivial character  $c \neq \varepsilon$ , they will treat this as the unambiguous, unchanging truth; indeed, their overall ‘impression’ of  $c$  can be represented as

$$I_{\mathcal{V}_i}(c) = \bigcup_{\rho \in R_{\mathcal{V}_i}(c)} \rho[2], \text{ where } R_{\mathcal{V}_i}(c) = \{\rho_{\mathcal{V}_i}(r) : \rho_{\mathcal{V}_i}(r)[1] = c\}.$$

If character  $i$  can define their own  $\mathcal{V}_i$ , they can approximate any other  $\mathcal{V}_j$  by  $\mathcal{V}_{i;j}$ : the subsequence of beats  $i$  believes  $j$  has heard.<sup>1</sup> It follows that  $i$  registers any rumour  $\rho$  about role  $r$  from character  $j$  as *incorrect* if  $\rho \not\approx \rho_{\mathcal{V}_i}(r)$ , *dishonest* if  $\rho \not\approx \rho_{\mathcal{V}_{i;j}}(r)$ , and *accidentally honest* if  $\rho \not\approx \rho_{\mathcal{V}_{i;j}}(r)$  but  $\rho \simeq \rho_{\mathcal{V}_i}(r)$ . These interpretations can even occur in cases  $\rho \simeq \rho_{\mathcal{V}_j}(r)$ , where the rumour is actually congruent with  $j$ 's own beliefs!

While this paper presents the model as simply as possible,  $R_{\mathcal{V}_i}(r)$  is readily customised to exclude rumours from characters  $a$  that  $i$  mistrusts, rumours accusing  $i$  themselves, *etc.* A more nuanced  $c_{\mathcal{V}}(r)$  would further allow characters to dynamically ‘change their mind’ during the simulation. Borrowing Mateas & Ferguson's (2014) terminology, the belief function is highly reusable in and of itself - and highly mixable with the simulation's other modules...

Sir Geoffrey is talking with Peter, Harry, Bart, and his own aide.  
They are ASSERTING THAT THE ONLY PEOPLE TESTIFYING AGAINST THE  
CABINET HAVE THEIR OWN EYES ON THE LEADERSHIP.  
Bart segues to MAKING SMALL TALK ABOUT THE ASHES.

Figure 3.3: An early, emergent interaction; extra edges are added to the script at runtime, allowing juxtapositions of topics unanticipated by the author. Whether Bart's segue seems sincere, suspicious, or simply uncanny, is left to reader.

<sup>1</sup>This being well-defined in the special case  $i = j$ , where, intuitively,  $\mathcal{V}_{i;i} = \mathcal{V}_i$ .

### 3.1.2 Characterisation

For a game like *Dwarf Fortress*, the novelty of the character-centric storytelling relies, in part, on the new characters generated every playthrough. *MiGS* (maybe wearing influences like *The Thick of It* on its sleeve) has more of a sitcom’s sensibility: it is interested in exploring how the same cast of well-defined, well-developed characters react to different initial conditions, even something as simple as a change of seed. To that end, each agent’s characterisation is stored as a fixed set of parameters, with some supporting logic to reflect their changing motivations.

Character logic is largely *utility-based*, with agents quantifying the ‘usefulness’ of potential actions by a normalised value  $u \in [0, 1]$ . An action  $A$  will only ever be taken if it has non-zero utility, and preferred over action  $B$  if and only if  $u_A > u_B$ . While this paper won’t review the background of utility theory in full, rest assured that it is a tried-and-tested means of simulation design. *The Sims* has long used utility to represent its NPCs’ needs (Graham 2014); *Prom Week* blends the approach with symbolic reasoning for ease of implicit creation (Mateas & Ferguson 2014).

In *MiGS*’ case, ‘authoring motivations’ means authoring *utility curves* (or *surfaces, etc.*): scalar fields that take in a range of inputs, and return a single, normalised utility value. Designing these curves is more art than science, and while this project follows best practices (Lewis 2017), many decisions come down to personal intuition.<sup>2</sup>

Generating an action might also come with several stages of utility considerations. Take the steps that go into producing a single beat of conversation:

1. Each character  $a$  calculates the utility of all vertices in  $V_a$ .
2. Each character  $a$  calculates their highest-utility beat  $\nu = (v, a, A, R)$ .
3. One character, chosen by weighted probability distribution, delivers their optimal beat.<sup>3</sup>

Without going through the equations one-by-one, the ‘vertex utility’ in *Step 1.* is a broad metric, used to get a sense of how well a given segue fits the current conversation. *Step 2.*’s ‘beat utility’ on the other hand, allows characters to much more finely assess how a specific action would align with their own goals. Its corresponding curve is a product of many factors, and highly customisable - amongst other considerations, *MiGS*’ characters will try to spread rumours about others they hold VENDETTAS against, avoid incriminating themselves or those they are LOYAL to, and more generally EXPLORE the topics they feel have interesting implications (these capitalised details all loaded in from `character_traits.json`). The earlier dialogue module can also be mixed in to prioritise beats congruent with  $\mathcal{V}_a$ .<sup>4</sup>

Note that *Step 2.*’s calculations incur a computational cost. Looping through  $v \in V_a$  in descending order of vertex utility, characters run the calculations only until they have matched three distinct vertices to three useful beats. Of these options, the highest-scoring is taken as  $\nu$ .

---

<sup>2</sup>For example, pathfinding returns distances in the range  $[0, \infty)$ , so I’ve found it useful to ‘bound’ these using the family of asymptotic curves  $u(a, b, u_0, u_\infty) = u_0 + (1 - a^b)(u_\infty - u_0)$ , where  $a \in (0, 1)$ ,  $b \in [0, \infty)$ ,  $u_0, u_\infty \in [0, 1]$ . These are not the only functions with  $u(a, 0, u_0, u_\infty) = u_0$ ,  $\lim_{b \rightarrow \infty} u(a, b, u_0, u_\infty) = u_\infty$ , though, and a different author may well have chosen differently.

<sup>3</sup>The weights, too, can be thought of a form of utility - specifically, a function of  $v$ ’s vertex utility, and the time since character  $a$  last spoke, chosen such that conversations ‘flow’ nicely along  $D$  without being dominated by any one character. However, it was also important to include a random component in these interactions, making the dialogue feel that bit more spontaneous.

<sup>4</sup>...In theory. The code included in `AgentTalking.cs`’ `ConsistencyUtility` function would allow for more convincing, consistent lies, but they actually become *too* convincing. If focal characters don’t have an equally advanced model of seeing through lies (and *MiGS*’ characters don’t), the behaviour ends up lost to the *Tale-Spin* effect. This particular utility consideration is therefore ignored to produce more dramatically compelling, if less realistic, narratives.

### 3.1.3 Blocking

Zooming out, *MiGS*' simulation is turn-based, iterated in discretised time. On every tick, exactly one verbal beat is generated (by the steps described above) in each room containing two or more characters. Characters are allowed to move between rooms on the 'off-beats', as it were.

The agents' movements (or, to continue the analogy to theatre, their *blocking*) rely on a similar model to conversation. These are staged on another graph  $S$ , with settings  $v \in V(S)$  connected by edges  $s \in E(S)$  (see Figure 3.5). Likewise, movements are instanced in *spatial beats*  $\varsigma = (v, a, A)$ , and likewise, character  $i$  commits to memory the sequence  $\mathcal{S}_i = \varsigma_0 \cdots \varsigma_N$  of beats they personally see.

As visibility rules, characters will observe all other characters in their current setting, and in all settings linked by line of sight (LOS). They similarly hear all of their current conversation, and have a probability  $p = 0.2$  of overhearing dialogue from rooms within earshot (ES). LOS and ES are also represented as graphs, variations on  $S$  that best reflect the stage's physical reality (see Figure 3.6).

In isolation, turn-based movements along a discretised map may seem a shallow approximation of reality; it is through the interplay of the spatial and verbal components that meaningful narratives emerge. Knowledge propagation and belief functions are, clearly, only non-trivial when agents aren't confined to a single setting. On a more structural level, allowing separate focalised narratives to converge and diverge with their narrators' movements should lay the foundation for an emergent *Rashomon* effect.

Most of the time, blocking is random: when a character is idle, each tick comes with a small chance of them picking a new setting and making a beeline towards it. They do so by following the minimal weighted path on  $S$  (again using pathfinding). However, since *MiGS*' characters can work from their epistemological frameworks to track each other's locations, the simulation takes inspiration from *Hide and Sneak* and includes its own searching behaviour.

By default, any character will assume a target remains at their last known setting - this being read from  $\mathcal{S}_i$ . If they make LOS on that setting and see they are no longer there, the character is then treated as lost. As in *Hide and Sneak*, this prompts a search of believable locations (in our case, those they haven't had LOS on in the last  $|V(S)|$  turns), nearest first, until they find the character or otherwise give up (see Figure 3.4).

```

if (agent A is following agent B && [IC] A has lost B
    && [IC] A believes B could be in LoS of LOCATION
    && [IC] A is not closer to any other such location)
    A searches for B by making a beeline to LOCATION;

```

Figure 3.4: Pseudocode for *MiGS*' search assumptions, written for comparison with Figure 2.1.

This is a nice reimplementaton of ten Brinke, Linssen & Theune's work under new visibility rules, but poorly deployed. For a start, not many characters are actually characterised as **FOLLOWING** anyone in `character_traits.json`. If they are, their behaviour is discreet - *MiGS* doesn't offer the same scope for direct confrontation as *Hide and Sneak*, so they achieve little more than shadowing their assigned target. While that is, theoretically, quite fitting for *MiGS*' paranoid, Cold War setting, the application's text-only interface makes it hard to track individual tendencies as 20 characters move simultaneously. It is a subtlety of the simulation subsumed by *Tale-Spin* effect.

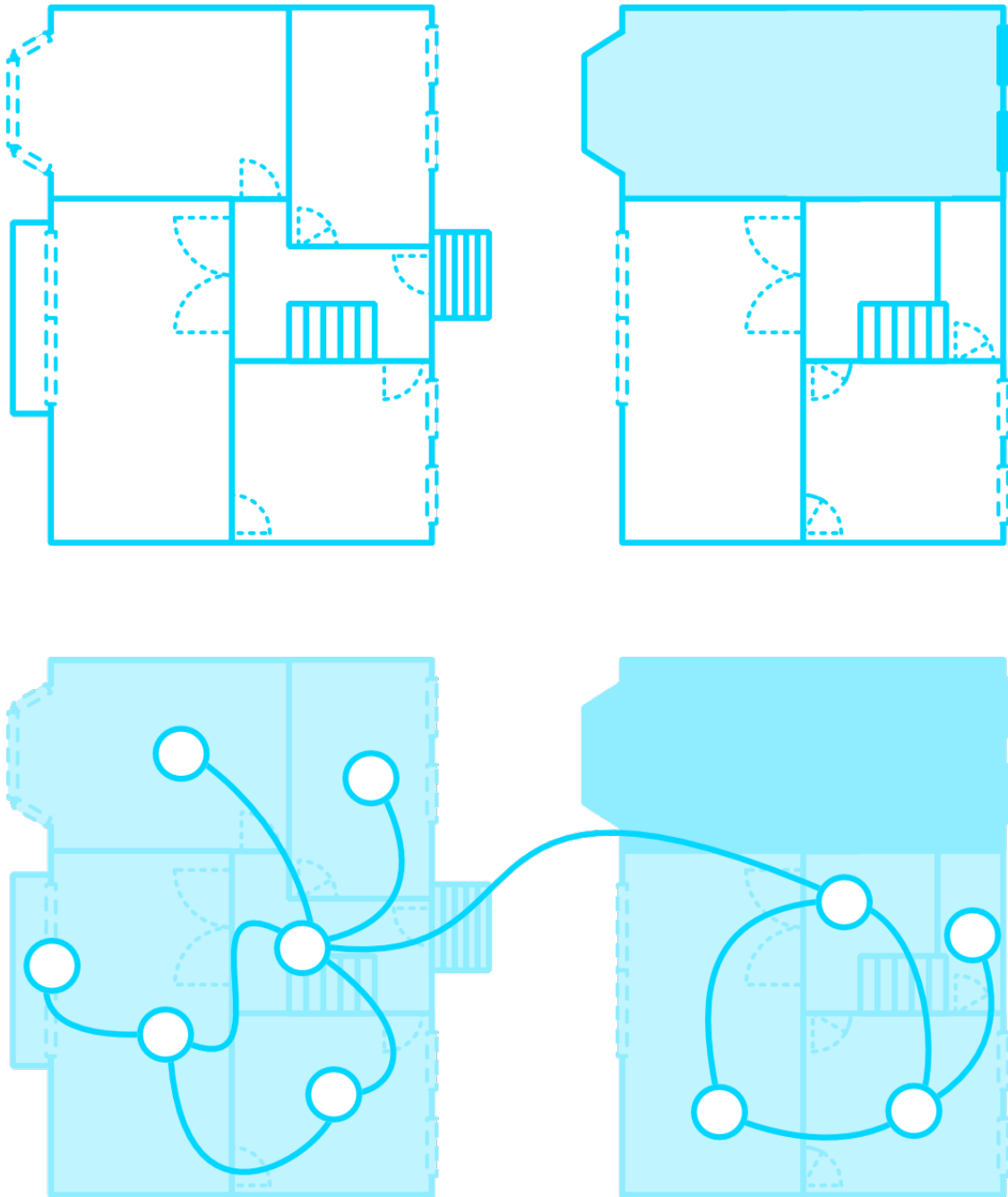


Figure 3.5: *MiGS'* upstairs (upper left), downstairs (upper right), and their graph representations (below). Being unable to represent proxemics in continuous space, the model assumes all characters in the same room are engaged in the same conversation.

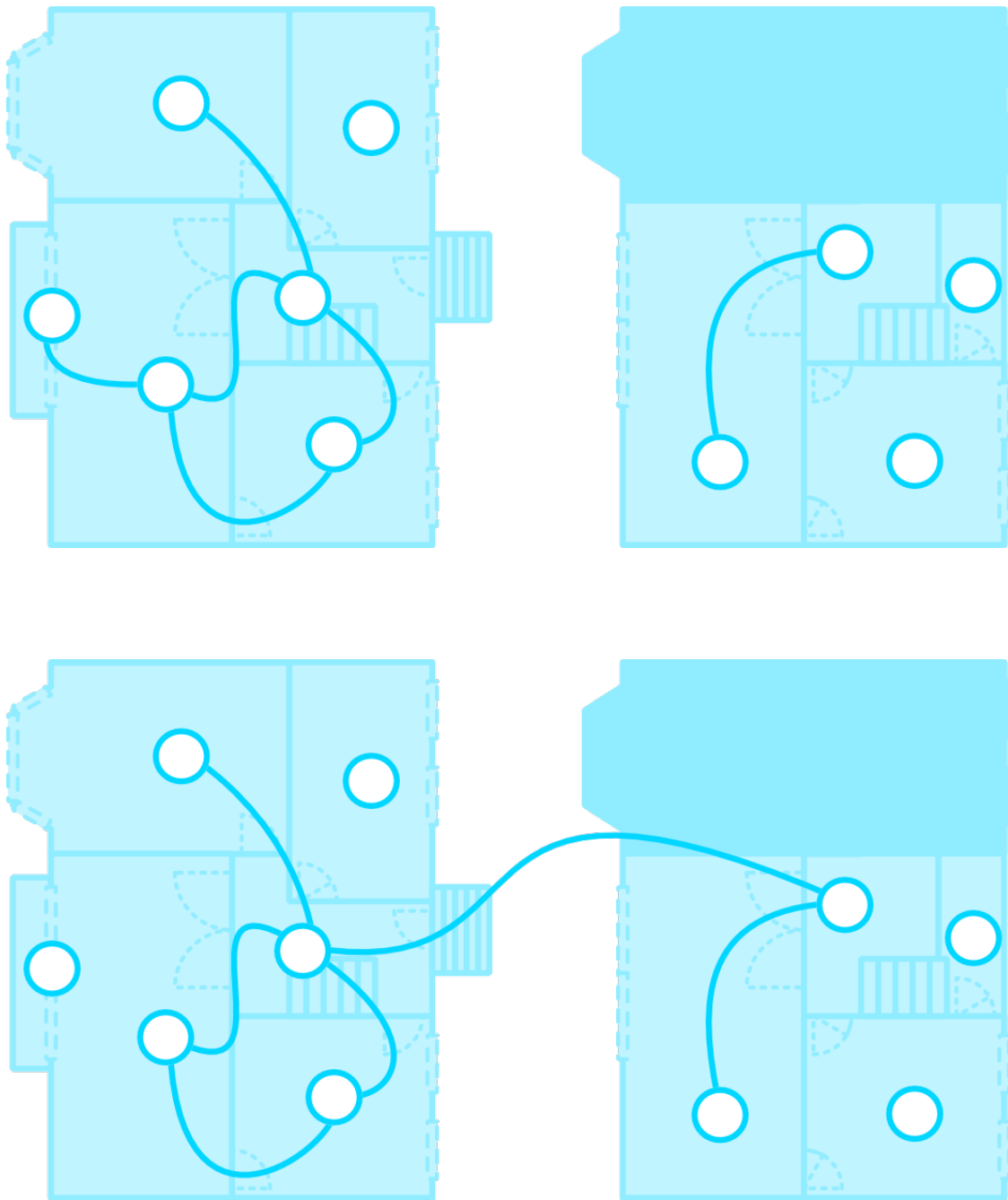


Figure 3.6: Graph representations of the stage's LOS (above) and ES (below). Notice how only voices carry up and down the stairwell; how characters are seen but not heard through the French windows to the balcony; how conversations behind closed doors are afforded a total privacy.



Note also that, unlike in *Hide and Sneak*, Figure 3.4 does not include any OOC drama management. Agents, in their capacity as actors, are not afforded an omniscient view of the stage, even if it would allow them to better steer their characters into conflict. Instead of using drama management to manufacture traces dense with stories, they sift out their storyful moments after the fact.

## 3.2 Sifting Patterns

In the code, agents' actions are stored in a `Trace` struct as `Beats`, all relevant details - where they are, whom they're with, whether they're telling a lie, *etc.* - encoded as strings in a dictionary of `m.attributes`. The class also includes an robust, hard-coded functionality for writing out these attributes in a serviceable prose. While there would be nothing technically wrong with printing the trace in its entirety,<sup>5</sup> this research demands those beats be packaged into microstories.

Given a trace, the `SiftingPattern` class will, with its `SiftMicrostories` function, return a complete list of microstories as follows. The list is initialised as

```
List<Microstory> matches = new List<Microstory>() { pattern };
```

containing only the initial sifting pattern, its logic variables unassigned. The list is then iterated as in Figure 3.7, binding variables beat-by-beat and discarding any partially completed matches as soon as they fail any constraint. Because of how the `for` loops have been organised to iterate over all partial matches simultaneously, it's hard to break the loop when say, 1000 matches are in the current iteration of the list - there's no guarantee they won't all fail some later constraint.<sup>6</sup> For a trace with  $N$  beats and a pattern with  $k$  event slots, being unable to terminate early allows computational cost to blow up with  $\mathcal{O}(N^k)$ . Luckily, authoring patterns for *brevity* (*i.e.*  $k$  small) and/or *specificity* (*i.e.*  $n \ll N$  beats satisfying each slot) means this hasn't proved an issue at the project's current scope.

```
foreach (?beat in pattern)
    List<Microstory> iteration = new List<Microstory>();
    foreach (partial_match in matches)
        foreach (beat in trace)
            bind beat to ?beat slot of partial_match;
            bind values to any empty ?value slots of any bindings [?beat a ?value]
                in partial_match; // This may return multiple partial matches...
            if (partial_match's logic variables satisfy all constraints)
                add partial_match to iteration;
    matches = iteration;
return matches;
```

Figure 3.7: Pseudocode finding all of a trace's microstories for a given pattern.

Checking a binding is satisfied is the same every time: if the constraint is of the form `[?e a v]`, the sifter verifies the beat in slot `?e` has an `m.attribute[a]` containing `v`; if it is of the form `[?e a ?v]`,

<sup>5</sup>In fact, it's very useful to have a complete log of every event that's happened in a simulation when debugging!

<sup>6</sup>This code can, of course, be refactored, but this came up too late in *MiGS*' development; for a non-interactive simulation that is not presented to users directly, the paper can abide a degree of inefficiency.

the sifter either binds a value from `?e` to slot `?v`, or again verifies consistency if a value already exists. Relations, on the other hand, need handled on a case-by-case basis, the following checks all hard-coded into the sifter:

- `(distinct ?v1 ... ?vn)` verifies that  $n$  values are pairwise distinct.
- `(within ?v ?v1 ... ?vn)` verifies that the value of `?v` is equal to at least one of  $n$  others.
- `(simultaneous ?e1 ... ?en)` verifies that  $n$  beats occur on the same tick.
- `(sequence ?e1 ... ?en)` verifies that  $n$  beats occur in non-decreasing order of ticks.
- `(without interruption ?e1 ?e2 a1=?v1 ... an=?vn)` verifies that the trace contains no beat `e` between beats `?e1`, `?e2` satisfying all  $n$  bindings `[e a1 ?v1]`, ..., `[e an ?vn]`.

As well as the earlier efficiency considerations, *MiGS*' patterns are written with best practices in mind. Though recognising the benefits of designing both components simultaneously, Kreminski, Wardrip-Fruin & Mateas advise “alternating between simulation-focused authoring and sifting-focused authoring” (2022, p. 216), to avoid imposing too much of a sifting pattern’s explicit structure onto the storyworld itself. ‘He Said, She Said’, for instance, was written to showcase a particular emergent scenario - two characters spreading the same rumour about each other - but it does so without being overly specific (see Figure 3.9).

```
(distinct ?actor ?audience)
  [?beat action talk]
  [?beat actor ?actor]
  [?beat audience ?audience]
```

Figure 3.8: The simulation’s custom ‘Rumour’ sifting pattern, written in the syntax of *Felt*. This matches any `beat` of dialogue, binding its `actor` and an `audience` member for consideration by STU.

```
(sequence ?heSaid ?sheSaid)
(distinct ?himself ?herself)
  [?heSaid action talk]
  [?heSaid actor ?himself]
  [?heSaid initialLine ?hisInitialLine]
  [?heSaid finalLine ?hisFinalLine]
  [?heSaid target ?hisTarget]
  [?sheSaid action talk]
  [?sheSaid actor ?herself]
  [?sheSaid initialLine ?herInitialLine]
  [?sheSaid finalLine ?herFinalLine]
  [?sheSaid target ?herTarget]
```

Figure 3.9: The simulation’s custom ‘He Said, She Said’ sifting pattern. The `heSaid` beat describes a character, `himself`, who will talk about `hisTarget` to his audience; `sheSaid` describes a `distinct` actor behaving similarly `herself`. Note that, despite the naming convention, the pattern contains no constraints on either character’s gender.

```

(within ?firstCandidate A B C)
(within ?secondCandidate A B C)
(distinct ?focalCharacter ?firstCandidate ?secondCandidate)
(within ?firstActor ?focalCharacter ?firstCandidate)
(within ?firstAudience ?focalCharacter ?firstCandidate)
(distinct ?firstActor ?firstAudience)
(distinct ?firstAccuser ?focalCharacter)
(within ?secondActor ?focalCharacter ?secondCandidate)
(within ?secondAudience ?secondCharacter ?secondCandidate)
(distinct ?secondActor ?secondAudience)
(distinct ?secondAccuser ?focalCharacter)
  [ ?beat focalisation ?focalCharacter ]
  [ ?beat action talk ]
  [ ?beat actor ?actor ]
  [ ?beat allAccused ?accused ]
  [ ?firstInteraction focalisation ?focalCharacter ]
  [ ?firstInteraction action talk ]
  [ ?firstInteraction actor ?firstActor ]
  [ ?firstInteraction audience ?firstAudience ]
  [ ?firstAccusation focalisation ?focalCharacter ]
  [ ?firstAccusation action talk ]
  [ ?firstAccusation actor ?firstAccuser ]
  [ ?firstAccusation allAccused ?firstCandidate ]
  [ ?secondInteraction focalisation ?focalCharacter ]
  [ ?secondInteraction action talk ]
  [ ?secondInteraction actor ?secondActor ]
  [ ?secondInteraction audience ?secondAudience ]
  [ ?secondAccusation focalisation ?focalCharacter ]
  [ ?secondAccusation action talk ]
  [ ?secondAccusation actor ?secondAccuser ]
  [ ?secondAccusation allAccused ?secondCandidate ]

```

Figure 3.10: The simulation’s custom ‘Meet the Candidates’ sifting pattern. This is highly specific, describing a `focalCharacter`’s knowledge of two of the three candidates A, B, and C. For both, the `focalCharacter` recalls a direct interaction with the candidate in question, and an accusation heard about them elsewhere. However, including a loosely defined `beat` in which extra, unexpected details can emerge renders queries for this pattern slow...

### 3.3 ‘Selecting the Unexpected’

Having generated a list of microstories, the **SiftingUnexpectedness** heuristic orders these according to their ‘surprisingness’. Each microstory’s property signature  $\sigma = ((p_1, v_1), \dots, (p_n, v_n))$  is just a dictionary, where properties  $p_i$  from Table 3.1 are paired with the corresponding, contextual values  $v_i$ . In the case where multiple values apply to a single property,  $v_i$  is chosen to minimise  $\mathbb{P}(p_i = v_i)$  (these likelihoods determined empirically, using some sample from the simulation; see below).

Property	Description
interlocution	The character speaking is ignoring the previous line.
rumour learned	Any character has just learned a rumour.
joke stolen	The character speaking is telling a joke someone previously told them.
incorrect	The character speaking is making an incorrect accusation.
dishonest	The character speaking is making an accusation they don’t believe.
accidentally honest	The character speaking is making a (correct) accusation they don’t believe.
confidentiality	The characters have some degree of privacy.
actor vendetta exercised	The character speaking is making an accusation against a supposed nemesis.
actor loyalty tested	The character speaking is making an accusation against a supposed ally.
logic variable == external	Same-value property; a logic variable is bound to an eavesdropping character.
logic variable == accused	Same-value property; a logic variable is bound to an accused character.
logic variable == focalisation	Same-value property; a logic variable is bound to the focal character.

Table 3.1: Properties in a Microstory’s property signature.

The heuristic can hence treat surprisingness as a form of utility, determined by some curve  $u(\sigma)$ . If a property signature has one value that is genuinely impossible (*i.e.*  $\mathbb{P}(p_i = v_i) = 0$ ), then, mathematically, the whole signature must be impossible (*i.e.*  $u(\sigma) = 1$ ), regardless of how probable any other properties might be. However, defining surprisingness as ‘how unlikely it is that  $\sigma$  contains all of the values  $v_i$ ’, with

$$u(\sigma) = \mathbb{P}\left(p_1 \neq v_1 \vee \dots \vee p_n \neq v_n\right) = 1 - \mathbb{P}\left(p_1 = v_1 \wedge \dots \wedge p_n = v_n\right) = 1 - \prod_{i=1}^n \mathbb{P}(p_i = v_i),$$

misses an emotional understanding that the more impossibilities a story has, the more unexpected it will feel. Instead taking surprisingness as ‘how unlikely it is  $\sigma$  contains *any* of the values  $v_i$ ’ gives

$$u(\sigma) = \mathbb{P}\left(p_1 \neq v_1 \wedge \dots \wedge p_n \neq v_n\right) = \prod_{i=1}^n \mathbb{P}(p_i \neq v_i) = \prod_{i=1}^n \left(1 - \mathbb{P}(p_i = v_i)\right),$$

but this, conversely, goes to zero if any single  $\mathbb{P}(p_i = v_i) = 1$ . *MiGS*’ implementation of STU uses

$$u(\sigma) = \prod_{i=1}^n \left(1 - \left(1 - \frac{1}{k}\right) \cdot \mathbb{P}(p_i = v_i)\right), \text{ where } k > 1,$$

such that a signature with some  $\mathbb{P}(p_i = v_i) = 0$  scores  $k$  times higher than if it had  $\mathbb{P}(p_i = v_i) = 1$  (the simulation returning good results with  $k = 4/3$ ).<sup>7</sup>

Even if these edge cases  $\mathbb{P}(p_i = v_i) = 0$  don’t occur in practice, they build an intuition behind our definition of  $u(\sigma)$ . Suppose  $\sigma$  has one ‘impossible’ property, and  $n - 1$  with likelihood 1: here,  $u(\sigma) = (1/k)^{n-1} > 0$ , which satisfies  $\lim_{n \rightarrow \infty} u(\sigma) = 0$ . Letting surprisingness slowly tend to zero like

<sup>7</sup>Note that, to make a fair, heuristic comparison of two microstories, if  $\sigma_1$  has some property  $p^*$  not in  $\sigma_2$ , then  $u(\sigma_2)$  must be scaled down by a factor of  $1/k$ . This is skimmed over in the above model, but handled in the code.

this reflects a human capacity to rationalise even the most overt statistical impossibilities if presented in a familiar enough context, making the model, to me, feel more true to life. As is a theme in this paper, even something as seemingly dispassionate and objective as a statistical sifting heuristic can communicate an author’s view of the world.

One criticism to be made of STU would be its approach to statistical sampling. The heuristic, under Kreminski, Dickinson, Wardrip-Fruin & Mateas’ implementation, uses the same sample of simulated events to evaluate likelihoods as it does to surface microstories. With no other frame of reference, the more disproportionately surprising a given trace actually is, the more emphatically the heuristic will surface its infrequent, boring moments (a paradoxical behaviour felt most acutely over shorter simulations, with smaller sample sizes). As a solution, this project calculates likelihoods from a separate sample of traces, generated by running the simulation with different seeds some  $n$  times, then saves these to its `LIKELIHOODS.json` for future use. Indeed, an even more nuanced approach would be generating these traces under ‘normal’ conditions, then tweaking character parameters for user-facing runs such that uncharacteristic behaviours become especially pronounced!

### 3.4 Focalisation

From a technical perspective, focalisation is just a generalisation of the above. Genette (1980) understands it as a mode - it is not that narratives from *Felt*, STU and other story sifters aren’t presented from a perspective, just that that perspective is (usually) nonfocalised. As such, this methodology has so far implemented its simulation and accompanying sifter as any project would. Focalising the trace, then the sifting patterns, and finally STU, is left to the end of development.

Just as *MiGS*’ complete trace  $\mathcal{T}$  consists of all its verbal and spatial beats, arranged in chronological order, so is character  $i$ ’s trace  $\mathcal{T}_i$  the combination of their  $\mathcal{V}_i$  and  $\mathcal{S}_i$  (defined above). Gervás (2014) and Hoek, Theune & Linssen (2014) treat  $\mathcal{T}_i$  as a subsequence of  $\mathcal{T}$ ; in our case, the two are constructed separately, so that  $i$  does not inherit the objective viewpoint of the complete trace.

Sifting patterns, too, are easily focalised: they are just queried against  $\mathcal{T}_i$  instead of  $\mathcal{T}$ . A more sophisticated approach would be to afford different agents different patterns to best reflect their characterisations - a paranoid or insecure NPC, for instance, might use patterns centred around betrayal. On a methodological level, this could create a very authentic effect of actors trying to rationalise and find structure in the data  $\mathcal{T}_i$  ‘through their character’s eyes’. Considering the authorial burden of writing 20 sets of sifting patterns for *MiGS*’ 20 characters, though, we are content to leave this to future research.

In the exact same way, where STU’s likelihoods were once approximated from a sample of previously generated  $\mathcal{T}$ , each character now calculates their own from the corresponding  $\mathcal{T}_i$ . It’s worth taking a moment to tease out the implications of this model. Here are a set of agents, each working OOC from their ‘past lives’ in prior runs, to identify the most interesting moments they have just experienced IC. The behaviour is analogous to how improvisers inform their performances with the practice of past rehearsals (and, for that matter, role-players with past campaigns); the best way to prepare for improvisation, after all, is to experience life.

When it talks about focalised story sifting, it should be clear by now why this paper treats it as more than just generating focalised narratives. The point-of-view stories that will be explicitly told by the simulated characters are also - implicitly - the ones their actors think best to surface. In *MiGS*’ case, allowing those actors to work from individualised sets of likelihoods affords them autonomy in both their IC *and* OOC roles, in keeping with the roots of EN and the experiences it tries to create.

## 4. Results and Discussion

Being highly subjective, this research relies on a questionnaire. Nine participants have taken part, each providing quantitative and qualitative feedback on a series of extracts sifted from *MiGS* (see Appendices A, B). Their unbiased, external perspectives were necessary to both verify the project’s focalisation of STU replicates the benefits of the original (**RQ1.**), and see how one engages with its generated artifacts without a prior mental model of their generator (**RQ2.**, **RQ3.**).

All extracts in the questionnaire were sifted from a single run of the simulation, initialised with a random seed (the likelihoods used for focalised story sifting calculated prior, over 50 seeded runs). The ‘experimental build’ used has been included as part of the project’s submission, complete with the relevant output logs, likelihoods data, *etc.*

### 4.1 Quantitative Analysis of Heuristics

For ready comparison with past results, the quantitative half of this questionnaire is directly inspired by Kreminski, Dickinson, Wardrip-Fruin & Mateas (2022). To evaluate their original implementation of STU, the team present three participants (or raters) with several pairs of microstories, asking for a binary assessment on which of each is ‘more interesting’. One microstory from each pair scores highly under STU, whilst the other is a match chosen completely at random (*i.e.* as if no heuristic had been used). Even with some slight changes - focalised likelihoods, fewer questions, a notably larger sample size - we hypothesise this research can replicate and reinforce Kreminski, Dickinson, Wardrip-Fruin & Mateas’ findings, that STU surfaces an overall more engaging set of microstories than a random baseline, and that the heuristic is therefore ‘better than nothing’.<sup>1</sup>

First, methodology: considered here are the simulation’s ‘Rumour’ and ‘He Said, She Said’ sifting patterns (see Figures 3.8, 3.9). In both cases, a focal character’s five ‘most interesting’ microstories are each paired with a different, randomly selected match from the same perspective and pattern. The two focal characters are chosen by hand, such that the resulting 20 extracts seem suitably varied. While this selection of focii is technically an act of sifting itself, it doesn’t interfere with any of STU’s processes, and should ultimately minimise survey fatigue.

To reduce bias, the extracts are edited manually - since the objective here is to see how closely STU’s definition of interestingness aligns with the participants’ own, it would unfair to show participants character names, topics of conversation, or other evocative prose the sifter effectively ‘cannot see’.<sup>2</sup> The resulting texts are generic, and not exactly reader-friendly, but there is no way to add clarity without removing any properties considered by the sifting heuristic, or introducing any that are not. It’s a consequence of “a narrative-specific form of the frame problem: the difficulty of

---

<sup>1</sup>Both papers are acutely aware that this is a highly artificial metric for EN. Just because a microstory is ‘more interesting’ does not make it, well, *interesting*...

<sup>2</sup>The focal character is factored into the heuristic’s calculations, so the extracts label them the [PROTAGONIST], whilst anonymising any others as [A], [B], [C], *etc.*

Which extract is more interesting?

- a. [A] is talking with one other person.  
 [A] tells a nasty rumour about [B], [C], [D]  
 (a rumour hitherto unheard by [PROTAGONIST]).  
 [PROTAGONIST] knows [A] is supposedly loyal to [B].  
 The conversation takes place one-on-one.  
 Still, [PROTAGONIST] overhears.
- b. [PROTAGONIST] is talking with [A].  
 [A] tells a nasty rumour about [B].  
 [PROTAGONIST] believes [A] is misinformed  
 (earlier, [PROTAGONIST] heard this same rumour about [C]).  
 The conversation takes place one-on-one.  
 No one overhears.

Figure 4.1: Question I.i (Extract a is heuristic-preferred, with success rate 77.78%).

determining what context is relevant when trying to determine whether a given event is worth mentioning in a story” (Kreminski, Dickinson, Wardrip-Fruin & Mateas 2022, p. 304). The order of extracts within each pairing is also randomised; unfortunately, though, the Google Forms platform would not allow the reordering of the pairings themselves from participant to participant.

Let’s establish some semantics before continuing. As stated in the questionnaire, this analysis is not about whether a participant can identify the heuristic-preferred answer, but whether the heuristic can identify *theirs*. There is no pretense that a statistical definition of interestingness is in any way ‘objectively correct’; where the following refers to a *success rate*, it means the percentage of times a sifter can meet a human’s preferences, and not the other way around. An omniscient, non-focalised sifter might nonetheless carry that pretense, but focalised story sifting better presents itself as one agent’s actor-awareness of what they personally find unlikely - and therefore, what an audience might want to hear. Considering the responses overall, *MiGS*’ focalised implementation of STU shows a mean success rate of 71.11%, and a (rather high) standard deviation of 23.69% (see Tables 4.1, 4.2).

Sifting Pattern	Success Rate (#)	Success Rate (%)
‘Rumour’	37/45	82.22%
‘He Said, She Said’	27/45	60.00%
<b>Overall</b>	<b>64/90</b>	<b>71.11%</b>

Table 4.1: Summary of success rates in Part I of the questionnaire.

If this experiment were run with two random heuristics, each would be expected to agree with the participant exactly half of the time. A heuristic therefore outperforms a random baseline if and only if its success rate is strictly greater than 50%. Taking significance at the 5% level, the Shapiro-Wilk test will first confirm that the sample of success rates does not significantly deviate from a normal distribution ( $p = 0.268 > 0.05$ ). As such, running a (parametric) one-sample  $t$ -test against a reference value of 50% and considering its results ( $t(8) = 2.674$ ,  $p = 0.028 < 0.05$ ) alongside the above analysis, the heuristic does indeed perform significantly better than its baseline (see Table 4.3).

Sifting Pattern	#1	#2	#3	#4	#5	#6	#7	#8	#9	Mean	Std. Dev.	Median	IQR
‘Rumour’	5/5	4/5	5/5	5/5	5/5	4/5	4/5	3/5	2/5	82.22%	21.08%	80.00%	30.00%
‘He Said, She Said’	4/5	3/5	0/5	5/5	5/5	4/5	3/5	1/5	2/5	60.00%	34.64%	60.00%	60.00%
<b>Overall</b>	<b>9/10</b>	<b>7/10</b>	<b>5/10</b>	<b>10/10</b>	<b>10/10</b>	<b>8/10</b>	<b>7/10</b>	<b>4/10</b>	<b>4/10</b>	<b>71.11%</b>	<b>23.69%</b>	<b>70.00%</b>	<b>50.00%</b>

Table 4.2: Exploratory data analysis on success rates, by participant.

Sifting Pattern	Shapiro-Wilk Stat.	Sig.	Decision	Hypothesis Test	Test Stat.	Sig.	Decision
‘Rumour’	$W(9) = 0.825$	0.039	Reject $H_0$	One-sample Wilcoxon signed rank test	$W(9) = 2.521$	0.012	Reject $H_0$
‘He Said, She Said’	$W(9) = 0.933$	0.510	Accept $H_0$	One-sample $t$ -test	$t(8) = 0.866$	0.412	Accept $H_0$
<b>Overall</b>	<b><math>W(9) = 0.903</math></b>	<b>0.268</b>	<b>Accept <math>H_0</math></b>	<b>One-sample <math>t</math>-test</b>	<b><math>t(8) = 2.674</math></b>	<b>0.028</b>	<b>Reject <math>H_0</math></b>

Table 4.3: Hypothesis testing for Table 4.2.

Partitioning by sifting pattern, more details emerge. The median success rate for the ‘Rumour’ pattern is 80.00%, significantly higher than 50% with  $p = 0.012 < 0.05$ . Not only does ‘He Said, She Said’ perform worse (mean success rate 60.00%), but shows no such significance ( $p = 0.412 > 0.05$ ). While it could be argued that the latter pattern was poorly authored, or indeed that the participants had already begun to experience fatigue, there is a no less valuable criticism to be made of the heuristic itself: some properties will simply seem more dramatically compelling than others.

This might explain the response to Question I.viii, the only question with success rate strictly lower than 50% (see Figure 4.2). Here, the heuristic-preferred text satisfies several same-value properties, allowing the two beats to feel interlinked in a way that not all ‘He Said, She Said’ microstories do. However, as Kreminski, Dickinson, Wardrip-Fruin & Mateas recognise, these artifacts are presented without the context of their simulations; the `actor loyalty tested` property in the randomly surfaced extract might have stood out as more overtly interesting, since participants don’t have a mental model of how frequently it actually occurs.

A notable outlier in Table 4.2 is Participant #3, finding all five of the heuristic’s ‘He Said, She Said’ microstories less interesting than their randomly surfaced counterparts. Considered in the context of Part II, however, it’s maybe fair to suggest they didn’t really engage with the survey (#3 is the only person to describe Question II.i as “easy”, despite leaving their answer to it incomplete). This is not to discount their feedback, only to highlight that not ‘getting’ these texts represents a fundamentally different criticism from not finding them interesting.

Finally, the debrief of the questionnaire asks participants to rate how much they agree with three statements, on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). No significant relationship to the results of Part I is found (see Table 4.4).

Statement	#1	#2	#3	#4	#5	#6	#7	#8	#9	Spearman Corr.	Sig.	Decision
‘I am an avid reader’	4	2	2	1	2	4	3	2	4	$r_s = -0.220$	0.570	No significant relationship
‘I enjoy analysing media’	5	5	4	3	3	5	4	2	3	$r_s = 0.223$	0.547	No significant relationship
‘I am familiar with procedural generation’	1	5	5	5	1	5	4	1	3	$r_s = 0.045$	0.909	No significant relationship

Table 4.4: Responses to the debrief, by participant, and their correlations to overall success rates.

This discussion is inexhaustive. The above addresses some limitations of STU (and its accompanying questionnaire), but many more are already identified by Kreminski, Dickinson, Wardrip-Fruin & Mateas themselves. Nevertheless, the results are sufficient for our paper’s purposes. They show the



Which extract is more interesting?

- a. [PROTAGONIST] is talking with [A], and three other people.  
[A] tells a nasty rumour about [B].  
The conversation takes place behind closed doors, with no one else around.  
No one overhears.
- Later, [PROTAGONIST] is talking with [B], [C], and one other person.  
[B] tells a different nasty rumour about [A], [D].  
[PROTAGONIST] knows [B] has longstanding vendettas against [A], [D].  
[PROTAGONIST] believes [B] is misinformed  
(earlier, [PROTAGONIST] heard this same rumour about [C], [D]).  
The conversation takes place behind closed doors, with no one else around.  
No one overhears.
- b. [A] is talking with one other person.  
[A] tells a nasty rumour about [B], [C], [D].  
[PROTAGONIST] knows [A] is supposedly loyal to [B].  
The conversation takes place one-on-one.  
Still, [PROTAGONIST] overhears.
- Later, [PROTAGONIST] is talking with [A], [C], [E], and one other person.  
[E] talks negatively about [F].  
The conversation takes place behind closed doors, with no one else around.  
No one overhears.

Figure 4.2: Question I.viii (Extract a is heuristic-preferred, with success rate 44.44%).

successful focalisation of an existing sifting heuristic, such that it still outperforms a random baseline - with a greater sample size than the original study, no less!

## 4.2 Qualitative Analysis of Focalisation

Part II, by contrast, faces participants with a single, long-form extract. After some handwritten **Context** about *MiGS*' storyworld, the text presents three different microstories from three different focal characters (**Bart Blakely**, **Seb Slater**, and **Nancy Schwarz**). The stories are the respective agents' most interesting matches to the 'Meet the Candidates' pattern, each consisting of five distinct beats (see Figure 3.10). While it is implicit that characters tell their stories in chronological order, all timestamps have been removed so as to present a narrative that needs 'pieced together'.

Of course, the meaning the extract creates through multiple internal focalisation will be hard to disentangle from how it uses voice. Each microstory is told in the first person, in past tense, by a character that exists within the fabula - undoubtedly heightening the effects identified below. To the extent that these factors have no bearing on the computational processes at play, being parsed in 'from the top down' only at the very end,<sup>3</sup> the following analysis is happy to take "focalisation" as shorthand for "focalisation, plus a blanket choice of voice".

---

<sup>3</sup>Parsed in by hand, for the experimental build - but using the same rules-based logic a computer would.

Accompanying the extract are five open-ended questions:

- II.i. Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?* First, the participant is asked to make connections between the three narrators' points-of-view. Whilst there are consciously not enough details for them to arrive at the 'correct' answer, the question should ensure a close reading of the text.
- II.ii. Did you find it easy or difficult to complete Question II.i? Why?* User experience feedback is essential in identifying shortcomings of the questionnaire that could affect its conclusions, or future research. The real purpose of this question, though, is to see if the participant discusses the task in gamified language (*e.g.* describing its difficulty as 'like a jigsaw puzzle'), suggesting an active engagement with the extract as a singular piece of dirty narrative.
- II.iii. Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?* This question draws the participant's attention to the extract's use of focalisation, as broadly as is possible.
- II.iv. Does any one rumour or discrepancy stand out to you? Why?* Last of all, the participant is asked to produce their own short retelling from the extract. Of interest are both what they sift out, and whether they ascribe it any meaning absent in the text itself.
- II.v. (Optional) Do you have any other thoughts on these narratives?* Only 3/9 participants take the opportunity to give more general feedback.

Surely enough, no participant would perfectly reconstruct the timeline; indeed, no two responses to Question II.i would turn out the same! On the numbers alone,<sup>4</sup> the closest answer to the 'correct' timeline comes from Participant #1 (at an edit distance of 8), two of the closest answers to each other from #1 and #2 (edit distance 6; see Table 4.5). While this isn't promising in and of itself, a closer look at the timelines shows which features of the story are preserved from participant to participant (see Figure 4.3). 4/5 respondents that provide complete answers identify beats (a.v) and (c.iv) as two perspectives on the same event (at least to the extent that they remain side-by-side in their own orderings). More surprisingly, a third of all participants pick up on the subtlety that (a.iii) leads directly in to (b.ii), despite the two distinct beats being told from two distinct perspectives. The question might be underdetermined, but at least some responses still leave these key moments of simultaneity and causality intact.

Others, though, are not so successful. Participant #3 has already been discussed, but #8 misreads the question as asking them to reorder the three narratives *separately* (their response being omitted from Figure 4.3), whilst #9's answer is so visibly off that it's hard to decipher how they arrived at it (using Roman numerals to label beats may have caused confusion, an oversight missed in our original pilot study). These answers are, surely enough, at the greatest distances from the original timeline.<sup>5</sup> However, that two further participants - participants more engaged with the rest of the survey - also respond with incomplete timelines suggests a genuine exhaustion with this particular question...

---

<sup>4</sup>Very briefly: the *edit distance* between two strings is the minimum number of operations it takes to rewrite one string into the other. *Longest Common Subsequence (LCS) distance*, for instance, counts insertions and deletions - the operations physically available if a participant were to write the extract's beats (a.i)-(c.v) on Post-It notes and reorder them by hand. On this intuition, the paper defines the edit distance between two timelines as their LCS distance, plus their difference in length (thereby penalising incomplete answers). The code used to calculate these distances is included in the project submission, but only for completeness' sake; this metric is just an informal means of data abstraction, so the paper doesn't place much importance in it.

<sup>5</sup>Not that it's a meaningful result, but discounting these outliers' answers to Part I, the 'He Said, She Said' success rate increases to 80% - a one-sample *t*-test now giving this as significantly different from 50% ( $t(5) = 4.108, p = 0.009 < 0.05$ )!

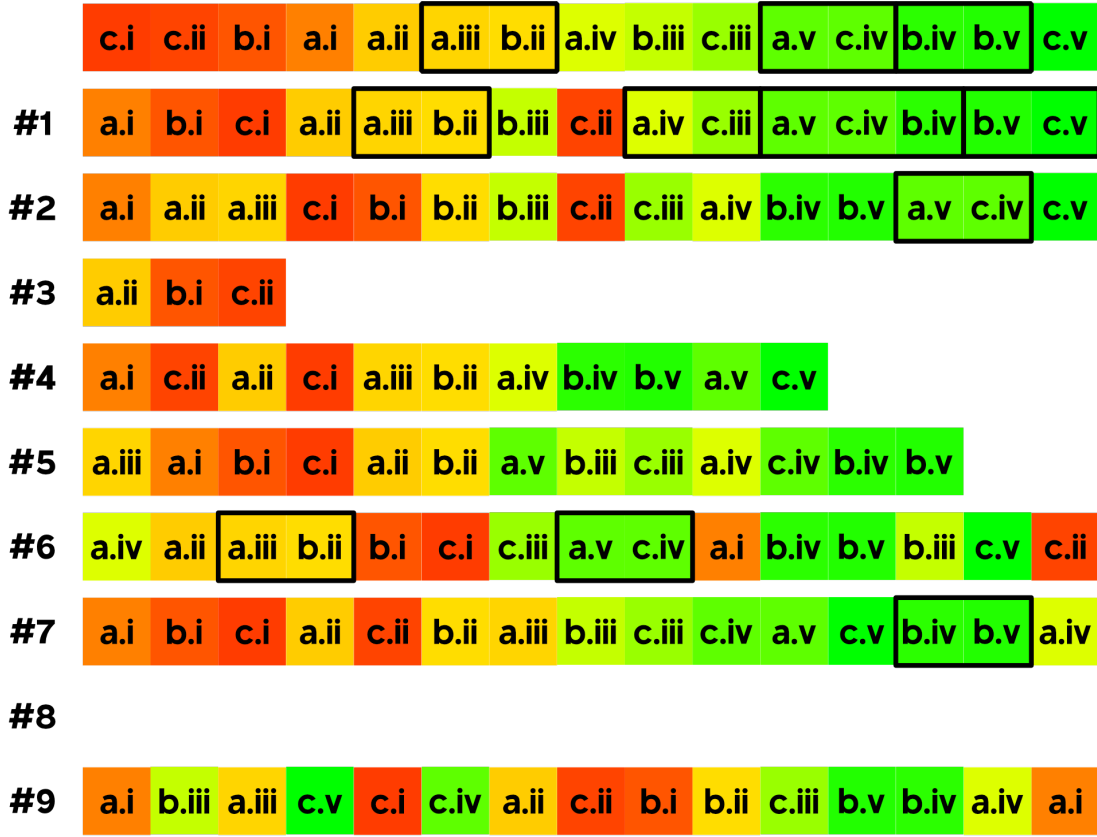


Figure 4.3: The extract’s correct timeline of events (top row), and its reconstruction by each participant. Solid outlines group the beats that overlap (or that participants say overlap) in each timeline. However, Question II.i’s wording is vague, with some reading “overlap” as ‘simultaneous’, others ‘simultaneous or consecutive’, so the actual orderings of beats (a.i)-(c.v) are far more instructive.

	#1	#2	#3	#4	#5	#6	#7	#8	#9	Timeline
Participant #1	0	6	26	14	8	12	10	30	16	8
Participant #2	6	0	24	12	14	16	12	30	14	12
Participant #3	26	24	0	20	24	24	26	6	26	28
Participant #4	14	12	20	0	14	18	18	22	20	14
Participant #5	8	14	24	14	0	16	14	26	16	14
Participant #6	12	16	24	18	16	0	14	30	20	12
Participant #7	10	12	26	18	14	14	0	30	12	10
Participant #8	30	30	6	22	26	30	30	0	30	30
Participant #9	16	14	26	20	16	20	12	30	0	18

Table 4.5: Edit distances of the participants’ reconstructions from one another, and from the correct timeline. When beats are grouped as overlapping, the table takes the minimum values over all permutations of each group (*e.g.* Timeline #1 has  $2! \times 2! \times 3! \times 2! = 48$  possible permutations).

As alluded to in Section 4.1, all but one response describe the task as difficult. Focalised story sifting has clearly recreated the “lack of a completely contiguous narrative” (#6) and overabundance of “small moving parts” (#2) that are staples of the *Rashomon* effect. These narrative features are well-employed in, say, *Five Little Pigs* (Christie 1943), a whodunnit that oversaturates its reader with information until they miss clues hidden in plain sight. Where our own extract instead feels *unfairly* difficult is that - unlike a whodunnit - answers to Question II.i cannot be checked.

This negative response seems to be a consequence of curation. There is, after all, nothing inherently wrong an open-ended narrative; that participants recognise they have “**no way of knowing**” if they are “getting the entire story” (#6) is intentional, implicit in *MiGS*’ subjective simulation design. The unfairness arises not from the extract itself, but how its authorial intent is at odds with the questionnaire. Question II.i is framed as if it really does have a unique solution, as if participants should, on looking carefully enough, reach a correct answer and ‘know when to stop’.<sup>6</sup> Coupled with the generator’s poor story support (see Chapter 5), it makes for an altogether frustrating experience. In the end, there is only one use of gamified language to respond to Question II.ii, and it sums up the sentiment perfectly: “it’s like *Cluedo* and I don’t like *Cluedo*” (#4).

Before moving on to identify what constitutes meaning created by focalised story sifting, it’s worth establishing what doesn’t. To avoid participants misunderstanding the extracts, they are explicitly told that the narrators might be unreliable - which leads to ‘hedged’ statements about how “it is **possible** that the character omits information not favourable to the character” (#3), and how “they **may** selectively remember or convey certain information” (#7). The analysis also ignores feedback like “I enjoyed the scene in (c.i) where Nancy spreads a rumour that Harry Peacock is a Marxist to divert attention from Nancy herself actually being the Marxist” (#2), in which a participant focuses only on the raw events of the simulation. The interest here is specifically whether the narratives curated through focalised story sifting create meaning not present in the fabula alone.

Focalisation explicitly positions readers ‘outside’ of non-focal NPCs. Nevertheless, the questionnaire hears that “Earl Peacock **lies** about his intentions and his own affiliations **to avoid** being seen as a ‘Marxist intellectual’” (#5), and that the rumour in (b.i) “**did not fit** the mental image of Mark that had been created by reading the story” (#6): conclusions drawn about character behaviours whilst, very consciously, missing those characters’ own perspectives. Where Hoek, Theune & Linssen see “focalization has a negative effect on identification with a non-focalized character” (2014, p. 13), our participants are distanced on an emotional *and* an intellectual level. #6, quite understandably, rejects details they find “isolated and conflicting” outright, but #5 shows a willingness to engage with the ambiguity. By first giving the participant access to the focal characters’ thought processes (scoping out the simulation’s expressive range of interactions), they are encouraged to draw their own meaning about non-focal characters’ actions at that same level of detail. To borrow a term from improv, Participant #5 is performing *late commitment*, saying ‘Yes, and...’ to the details they see from the outside and rationalising them after the fact in their own retelling (Swartjes 2010).

Meanwhile, Participants #4, #7, #8 all say the focal character Bart Blakely’s apparent leadership bid stands out to them, with #8 highlighting the rumour specifically because “it was never alluded to from **his own** POV”. There’s a wariness here of how the narrators regulate information, which, positioned not as ethereal voices but figures ‘in story’, in turn informs their characterisation. Bart comes across as untrustworthy not only through his actions, but also his perceived omission of those actions to the reader.<sup>7</sup> Participant #4, however, reasons that if he were running for the leadership,

<sup>6</sup>The question does include a comment that “If there are multiple valid orders, choose one that makes most sense to you”, but this is a gross understatement of its open-endedness.

<sup>7</sup>This isn’t the only such example: though it reads as slightly more hedged, Participant #2 notices that other narrators overhear Seb Slater having conversations in rooms he never admits to being in. They conclude with a value judgement on his character, that his omissions make it “seem like Seb is a more unreliable narrator”.

then “surely he would know which people are losing confidence in the PM... which he says he doesn’t”. That someone has actively engaged with such a subtle context clue - that the access offered by internal focalisation implies Bart truly doesn’t know this rumour - suggests a genuine emergence of dirty narrative.

These two effects are two sides of the same coin, occurring as information is withheld ‘from without’ or ‘from within’ a character’s own perspective. Just as compelling, though, are the moments those external and internal perspectives on a character actively disagree. When Participant #1 says it is “interesting that Blakely thinks Slater is loyal to Bohring **when in fact** he has a vendetta against him”, they are drawing a conclusion on a contradiction in the extract. From the outside, Bart is surprised to overhear Seb spreading a rumour about Mark Bohring, despite being loyal to the man; within Seb’s own narration, though, he reveals a “longstanding vendetta against Mark”.

What makes this moment work (and what makes it, as an author, my favourite example of dirty narrative to come from *MiGS*) is that it rewards active interpretation. Not just a contradiction for contradiction’s sake, Bart’s observation raises a question over Seb’s loyalty, to which Seb himself provides a very definitive answer later on.<sup>8</sup> Participant #1 is, in connecting these two perspectives as more than the sum of their parts, gaining insight into the character’s duplicity. The extract is still too vague for Question II.i to be fully solvable, but here is a reconcilable moment that does emerge - the sifter just needed to get a little lucky.

From this feedback, focalised story sifting can be seen to regulate information, creating omissions from narrators and ambiguities around secondary characters. Equally, it allows a ‘deregulation’ of information in the form of contradictions. These are familiar effects of focalisation, but here they have been achieved by a sifter, to the satisfaction of not just myself but unbiased readers. Inasmuch as their quotes above represent the cherry-picked successes of the study, implicit creation sees value in cherry-picking. It is an iterative process, where only after qualifying the possible effects of a generator can one work to achieve them more reliably. As Mateas & Stern (2005) put it, there is a need to both explore new regions of design space (as this analysis does with focalised sifting) and systematically map those regions out (as is left open for future work).

Recalling our research questions, these findings are also all very explicit: meaning is being drawn from the generator’s individual artifacts, not the user’s understanding of the generator itself. While participants can form a rough mental model from the extracts alone (#6 even recognising how the character-centric simulation does not “guarantee... a clear arc from start to end”), PCG is meant to be experienced as more than just a set of static outputs. Our conclusions must therefore reframe the experimental results within the surrounding literature.

---

<sup>8</sup>There is a fair criticism to be made that this ‘planting and pay-off’ has not been created by the computer itself, since Bart’s narrative was manually placed before Seb’s in curating them for the questionnaire. If the perspectives were shown the other way around, though, the reader would (finding out Seb’s true feelings first) experience a stronger sense of dramatic irony on reading Bart’s narrative; in either order, there’s a rewarding effect for those that pay attention.

## 5. Conclusions and Future Work

### RQ1. How might existing sifting techniques be focalised?

Having now built up a rich foundation of authorial experience, audience feedback, and formal theory, it would seem this first research question is easily answered. The challenge of focalising a story sifter can be simplified to focalising its input data - in STU’s case, its trace, and its sample of likelihoods. Where past projects (Gervás 2014; Hoek, Theune & Linssen 2014) treat focal characters’ traces as purely objective accounts, *MiGS* filters entries through the characters’ eyes themselves, with the understanding that focalised story sifting is enhanced by subjective data. Not only does this paper walk through its methodology of focalising a specific, cutting-edge sifting heuristic, but quantitative analysis (while limited) suggests the methodology is successful. Being run on different simulations does not allow for easy, one-to-one comparison, but both Kreminski, Dickinson, Wardrip-Fruin & Mateas’ implementation of STU and our own outperform their random baselines.<sup>1</sup>

Not that the focalisation of STU is without its issues. It inherits many known limitations of the original heuristic, including the limitation of *correlated unlikelyhoods*. The various definitions of  $u(\sigma)$  in Section 3.3 are all built on an assumption that the properties of  $\sigma$  are independent, and hence that

$$\mathbb{P}\left(p_1 = v_1 \wedge \cdots \wedge p_n = v_n\right) = \prod_{i=1}^n \mathbb{P}(p_i = v_i).$$

However, this misses the possibility that two likely properties might be unlikely to appear in the same signature, or vice versa.

Consider, for example, that characters in *MiGS* often make accusations about one another, and often have loyalties to one another; why should the heuristic register a character making an accusation about someone they are loyal to as surprising? Rather than treating accusations and loyalties as two separate properties, Table 3.1 contains a single **actor loyalty tested** entry that applies to this very specific scenario. Adapting STU to handle correlated unlikelyhoods directly remains open for future work, but - for the benefit of other authors - accounting for interesting correlations by writing new properties has here proven an effective workaround.

Other issues are not so easily resolved. While focalising STU’s likelihoods felt like this big innovation in theory, it has had little effect in practice. The agents’ data sets are unique, but not *perceptually unique*; their distinctions are too subtle for the end user to see different microstories as having been surfaced by different definitions of interestingness. This is the ‘10,000 Bowls of Oatmeal’ problem, a hallmark of not just EN but all of PCG (Compton 2019).

Reframing **RQ1.**, perhaps the real technical challenge of focalised story sifting is not implementing the sifter itself, but creating a storyworld that lends itself to focalisation. Simulation design is a well-recognised *wicked problem*, in that “the problem and proposed solutions mutually define each

---

<sup>1</sup>Similar to our poor curation of the qualitative section, Kreminski, Dickinson, Wardrip-Fruin & Mateas recognise their quantitative questions represent “a worst-case scenario for the presentation of sifted microstories” (p. 300). Achieving a significant, overall success rate of 71.11% in this worst case is a sound result, even with its variance across patterns.

other” (Mateas & Stern 2005, p. 7). Only after creating its own, bespoke social simulation can this paper evaluate what has and hasn’t worked, and with that identify the features that could make each actor’s experience feel meaningfully unique. *MiGS*’ model of characterisation does not lead to significantly different behaviours across agents, for a start. *Gossamer* (Kreminski 2023b), meanwhile, is a refinement of previous gossip simulations like *Talk of the Town*, which - published too late to inform *MiGS* itself - should serve as an instructive comparison for future development.

Personally, though, I find my experimental build is most limited by its interactivity - or lack thereof. Returning to the analogy of the narrative instrument, non-interactive simulations might be criticised as “windchimes: they produce a kind of pretty but uncomplicated ambient background narrativity” (Kreminski & Mateas 2021, p. 505). That ‘background narrativity’ seemed fun to try and capture at first (not to mention lesser in scope!), but both the questionnaire and my own experiences of debugging make clear how difficult it is to parse the resulting screeds of text. To that end, I’ve already started on a new ‘interactive build’ of *MiGS* (see Figure 5.1).

While the work has so far been focused on migrating the existing code into Godot, being able to interact with my characters firsthand lays bare a lot of the crudeness of my models. Equally, though, it allows for easier bug-fixing, and a more tactile understanding of whether or not each module works. As an writer, I’m absolutely satisfied with how my implementation of waypoint narrative allows me to blend explicitly authored text into implicitly created character interactions (even if the new interface shows just how often characters end up ‘talking in circles’). Another success of the simulation design exists outside of any reusability or mixability concerns - the ease with which I was able to integrate an actual player into the storyworld proves the modules really were designed for symmetry, that they really can be ported project to project as social LEGOs.<sup>2</sup>

Indeed, the interactive build showcases a clear use for focalised story sifting in commercial games: the debrief. This paper is by no means the first to suggest retelling ‘key moments’ at the end of a playthrough, or even retelling them through the eyes of NPCs (Hoek, Theune & Linssen 2014), but it’s a use of feedback curation that stands in stark contrast to our questionnaire. If a user first experiences the simulation in full, unfolding at their own pace, they will be anchored to a solid (but nevertheless subjective) timeline of events. Through interaction, there should come a sense they are now “getting the entire story” (#6), hopefully causing less difficulty as they next try and reconcile other characters’ perspectives (see Figure 5.2) with their own, larger piece of the puzzle.

## **RQ2. What meaning might focalised story sifting create in an individual EN artifact?**

As established, *Felt* and STU have very narrow understandings of what makes a story ‘high quality’, recognising the value of event sequences but not the value of how these might be reordered, or the perspective from which they might be retold. Sifting should be, in narratological terms, a transition from fabula to syuzhet, but these efforts consider only the first half of that transition; this paper maintains that sifters must serve a dual purpose, recognising both the quality of extant structures in the trace, *and* the quality at which they can be restructured as part of a media experience.

The results of Section 4.2, though mixed, show promise. Despite the poor curation of the long-form extract, participants with no prior mental model for *MiGS* could identify several key effects of its multiple internal focalisation. Focal characters make omissions; non-focal characters’ actions carry a

---

<sup>2</sup>But, being a wicked problem, there is always more to be done. Two finer points of interactivity I didn’t have time to address in this follow-up work relate to the player’s dialogue options. First, the simulation still calculates the ‘optimal’ rumour  $\rho$  to pair with each line, rather than let them choose whom to accuse by themselves (an unwanted constraint on agency). Furthermore, the simulation logs repetitions of lines that have been said, but not how many times the player is shown the lines as potential options - meaning they might get stuck with the same, undesirable choice over and over...

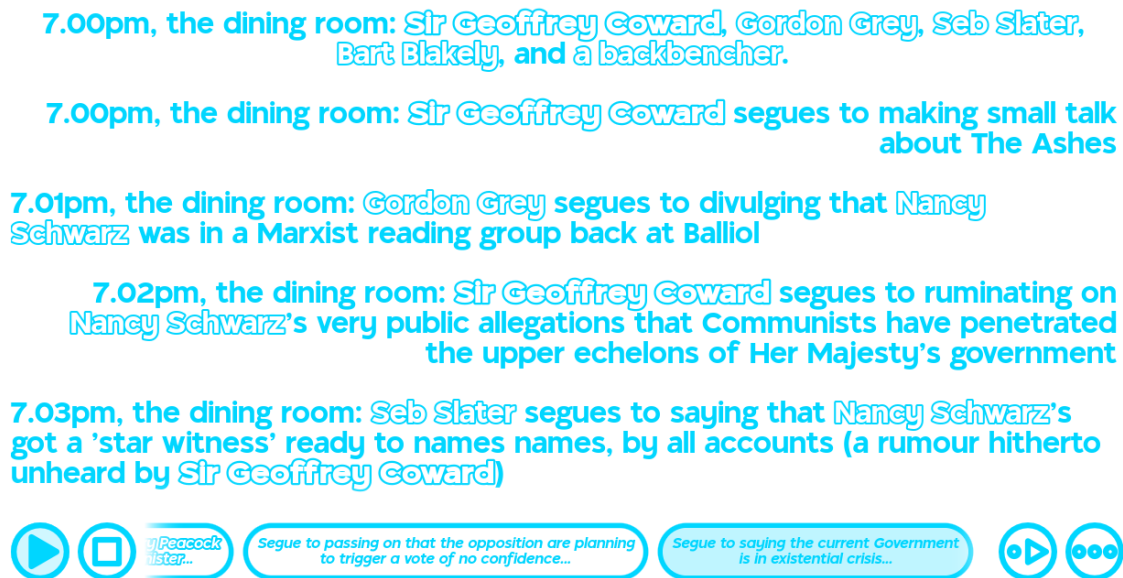


Figure 5.1: UI for an interactive build of *MiGS* (still in development). The right alignment of the player actions, the emboldening of character names, the italicisation of 'stage directions', *etc.*, all serve as visual story support, presenting the storyworld with more clarity than the original text console.

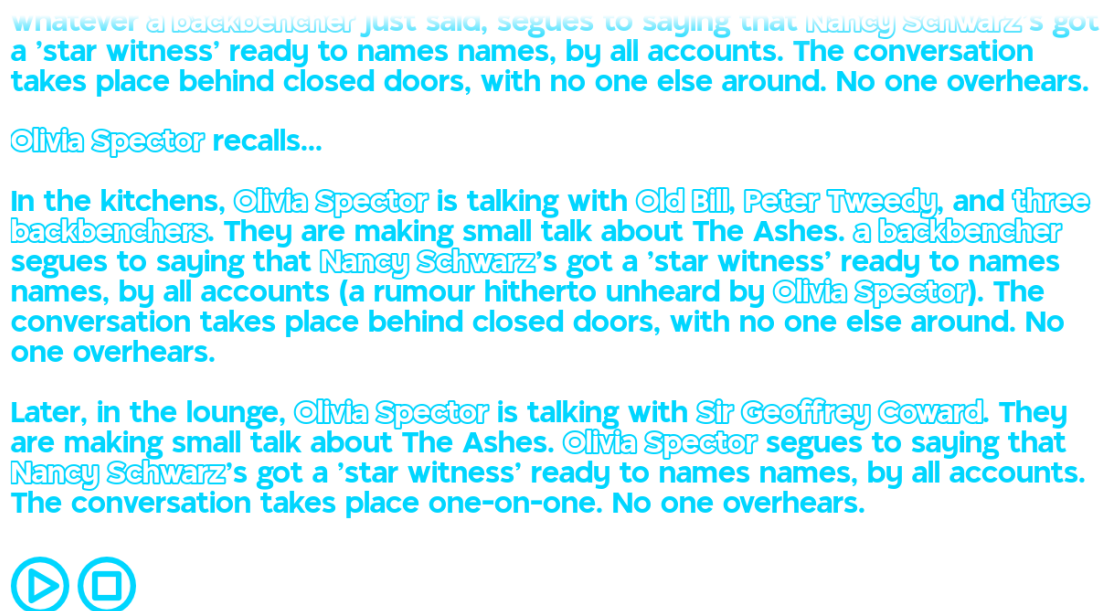


Figure 5.2: A focalised retelling from an NPC, mounted within the interactive build. To relate these retellings to the player's own experience, avoiding any "random inclusion of a new character" (#6), this perspective comes from one of the three NPCs the player interacts with most over the playthrough.



certain ambiguity; internal and external perspectives on the same character will (with luck) lead to solvable contradictions. That’s not to say that omissions, ambiguities and contradictions couldn’t be created by an omniscient narrator - but by positioning them within the story itself, participants read the narrators’ regulation of information through the syuzhet as related to their behaviours in the fabula. Even if it has not achieved a *Rashomon* effect in its totality, focalised story sifting allows such features of dirty narrative to emerge through the story structures of its individual artifacts.

Concerned as this paper is with focalisation, sifters could equally incorporate any of Genette’s analytical segments (see Section 2.1). Feedback to Question II.ii, for instance, highlights the importance of duration in procedural storytelling. “Small moving parts” (#2) may well be a staple of dirty narrative, but authors with a good sense of pacing know when - and when not - to slow down and add detail to with a descriptive pause. A sifter with those same capabilities would go some way toward resolving the earlier ‘narrative frame problem’ (Kreminski, Dickinson, Wardrip-Fruin & Mateas 2022).

Future research might also relate to narrative voice, the joint consideration of “who speaks” and “who sees” (Genette 1980, p. 186) creating richer characterisations of narrators than focalisation alone. While a sifter could weigh up the tense or grammatical person of the narration (beyond the blanket choice made in the questionnaire), voice equally encompasses syntax. Capturing a narrator’s dialect, their malapropisms, their general way of speaking, requires significant *story support* - this being another of Ryan, Mateas & Wardrip-Fruin’s open challenges for EN (2015), the challenge of presenting emergent phenomena in a compelling and user-friendly form.

Just as it would have benefitted from better curation, so too would *MiGS* benefit from better support. Feedback that the “the repetitive phrasing used by **generator**... makes it complicated to separate the narratives mentally” (#4) suggests that, more than just being hard to parse, the questionnaire’s poorly voiced, poorly supported narratives don’t carry enough social presence for a reader to ‘buy in’ and treat them as coming from the focal characters themselves. Data from the simulation’s *Beat* class is printed according to an explicitly authored set of rules, its story support imposed rigidly from top down.

One might be tempted to ask - won’t ChatGPT fix this? Being in the zeitgeist, it’s worth considering the role natural language processing (NLP) could play in EN. Mateas & Ferguson (2014) say generative text would allow for more readily mixable modules; several simulations referenced in Chapter 2 have had their own attempts at generating natural language retellings (Gervás 2012; Antoun et al. 2015). There is nothing that precludes a use of NLP over, say, grammar-based text generation to provide such support, or indeed to reduce a writer’s own explicit authorial burden (Jones 2022).

Consider, however, Figure 5.3. This narrative was generated by feeding the entire long-form extract from Part II of the questionnaire into ChatGPT, to see how an AI might have reworked it into a more readable prose. Having also received extra character details not included in the text, it was asked to rewrite each narrative with a distinct voice, capturing a “sense of the narrator’s personality, key relationships and feelings, and motivations”. The result is less repetitive than before, granted, but doesn’t exactly read as natural speech - it’s not even been parsed from bullet-pointed beats into a single, flowing paragraph! Compounded with ChatGPT’s difficulties remembering and enacting instructions (Méndez & Gervás 2023), it makes for an inept retelling.

Past the technical limitations, though, there exists a more fundamental argument against this use of ChatGPT. For any of its design flaws, there’s an honesty at the heart of *MiGS*: as an author, I’ve tried to distill my own, very human, experience of conversation into a computational model, a model through which a cast of characters I find personally interesting can interact. I do not feel that same

**Nancy Schwarz's Opinionated Eye-Witness Account:**

- As I stood in the downstairs corridor, I found myself engaged in light chatter with Bart Blakely and Sir Geoffrey's butler. We exchanged pleasantries about The Ashes, though truth be told, I wasn't all that interested. I casually let slip a tidbit about Earl Harry Peacock's past dalliance with Marxism. Oh, the irony in that one.
- Later, in the upstairs corridor, I caught sight of Seb Slater, always lurking like a specter, deep in conversation. They were dissecting my claims of Communist infiltration. Seb reluctantly conceded that my faction indeed held some sway, though he likely didn't want to admit it.
- In the billiards room, I found myself in conversation with a couple of backbenchers. We mulled over the possibility of The Standard having a major scoop. One daring backbencher blurted out that Bart Blakely was mulling a challenge against The Prime Minister. The buzz in the air was electric.
- Later still, in the billiards room, Bart and I engaged in conversation with a few others. We pondered the hearings, wondering if they were nothing but a smokescreen. Bart and I exchanged playful quips about making up evidence, a tidbit I'd filched from Sir Geoffrey's butler. Oh, how I savored this little dance.
- And just when I thought the night was winding down, Earl Peacock entered the fray. We dissected rumors from the press, and I seized the moment to declare that Sir Geoffrey Coward was preparing to challenge The Prime Minister. The satisfaction of stirring the pot was truly unmatched.

Figure 5.3: A focalised and voiced retelling, generated by GPT-3.5.

honesty as I chuck *MiGS*' outputs into third-party software for lacklustre story support. Just as writing is self-expression through text, implicit creation is expression through *processes*, and yet this process is not my own. Even if ChatGPT did enhance the meaning of my individual artifacts (which, it cannot be stressed enough, Figure 5.3 does not), those artifacts would speak to an underlying artistic method that is disingenuous at best, plagiaristic at worst.

**RQ3. What meaning might focalised story sifting create across a range of EN artifacts?**

The paper therefore agrees that “the aesthetic effect of a generative system is not solely in the artifacts that are generated, but also in the relationship between them and the emergent operation of the system” (Karth 2023, p. 82). Our own questionnaire presents its artifacts in isolation, but that isn't uncommon for research into PCG - it just requires a more theoretical analysis of how focalised story sifters might be experienced ‘in the long term’. Where focalisation has created meaning in individual outputs, for instance, this will likely lessen upon perceptual collapse. If the user were allowed to generate enough narratives to build an accurate mental model of *MiGS*, they would realise the narrators' omissions, *etc.*, aren't nearly as intentional as they first seem. This doesn't disqualify the previous answer to **RQ2.**, only demonstrate that over a greater range of artifacts, some meanings might be lost, and others gained...

One of the most exciting prospects for the future of EN is the incremental story sifter. Unlike a retrospective sifter, this would run concurrent to its simulation; *Winnow* (Kreminski, Dickinson & Mateas 2021) is an early effort that searches for partial matches to sifting patterns in real-time, and then steers agents towards completing them. Past the immediate drama management capabilities (*i.e.* making sure emergent narratives actually emerge), this technique has been recognised for its clear potential in commercial contexts. Partial matches could be surfaced as ‘quests’ to act out, steering the player without reducing their agency.

Relating this to the inherent “pleasure of emergent narrative” (Ryan 2018, p. 37), consider that even when drama management fails and a simulation derails itself, there can still arise an uncanny humour. It is a pleasure not dissimilar to that of improv, where comedy is contrived not only on the intradiegetic level of characters interacting, but the ‘meta’, extradiegetic level of performers making mistakes (Landert 2021). To focalise an incremental sifter, therefore, is to embed in each agent an actor’s understanding of where a story is going; to see those understandings align and misalign as they are fed back into the simulation; to take pleasure in any authentic, extradiegetic errors. Not only the characters, but their actors, would make real-time decisions from independent epistemological frameworks.

It’s an idealised vision, and one that may still struggle with perceptual uniqueness if implemented poorly. It is also the clearest possible articulation of focalised story sifting as a character-centric approach to recognising - and here, applying drama management in light of - emergent narrative structures. That not even sifting is imposed ‘from the top down’ surely fosters a more meaningful agency, and refines EN to be more meaningfully true to its collaborative, improvisational roots. In creating more and more artifacts, their individual value might diminish, but the sense of authenticity that comes through understanding this generative process can only grow.

As an answer to our third and final research question, the above is grounded in more than just a technical result from *Men in Grey Suits*; in more, even, than any formal finding from the project’s (substantive, but by no means systematic) literature review. These conclusions are all, to some extent, inextricable from my own experiences authoring emergent narrative, and my own philosophy of what a generator should even *be*. Where I personally see value in the identified effects of focalised story sifting, a different project could just as well find them thematically or tonally inappropriate. This paper does not wish to impose an unwanted vision of emergent narrative on other authors - only to present its own, authentic point-of-view.

# Appendix A: Questionnaire Extracts

This appendix contains the text extracts presented within this project’s questionnaire. These were generated using focalised story sifting, with additional curation added by hand (as described in Chapter 4).

## Quantitative Analysis

The following pairings were presented for comparison in Part I of the questionnaire. The first extract in each pairing was the heuristic-preferred option.

**‘Rumour’ Pairing #1** [A] is talking with one other person. [A] tells a nasty rumour about [B], [C], [D] (a rumour hitherto unheard by [PROTAGONIST]). [PROTAGONIST] knows [A] is supposedly loyal to [B]. The conversation takes place one-on-one. Still, [PROTAGONIST] overhears. [PROTAGONIST] is talking with [A]. [A] tells a nasty rumour about [B]. [PROTAGONIST] believes [A] is misinformed (earlier, [PROTAGONIST] heard this same rumour about [C]). The conversation takes place one-on-one. No one overhears.

**‘Rumour’ Pairing #2** [A] is talking with someone out of sight. [A] tells a nasty rumour about [B] (a rumour hitherto unheard by [PROTAGONIST]). [PROTAGONIST] knows [A] has a longstanding vendetta against [B]. The conversation takes place with no one else around. Still, [PROTAGONIST] overhears.

[PROTAGONIST] is talking with two other people. [PROTAGONIST] tells a nasty rumour about [A], [B], [C]. [PROTAGONIST] has a longstanding vendetta against [A]. The conversation takes place behind closed doors, with no one else around. No one overhears.

**‘Rumour’ Pairing #3** [PROTAGONIST] is talking with [A], and one other person. [A], not really paying attention to whatever [PROTAGONIST] just said, tells a nasty rumour about [B]. [PROTAGONIST] believes [A] is lying (earlier, [PROTAGONIST] heard this same rumour about [A] themself!). The conversation takes place openly. Still, no one overhears.

[PROTAGONIST] is talking with two other people. The conversation takes place openly. Still, no one overhears.

**‘Rumour’ Pairing #4** [PROTAGONIST] is talking with three other people. [PROTAGONIST] tells a nasty joke at the expense of [A] (a joke stolen from [B]). [PROTAGONIST] has a longstanding vendetta against [A]. The conversation takes place behind closed doors, with no one else around. No one overhears.

[PROTAGONIST] is talking with two other people. The conversation takes place openly. Still, no one overhears.

**‘Rumour’ Pairing #5** [PROTAGONIST] is talking with [A], and two other people. [A] tells a nasty rumour about [B], [C] (a rumour hitherto unheard by [PROTAGONIST]). [PROTAGONIST] knows [A] has longstanding vendettas against [B], [C]. The conversation takes place behind closed doors, with no one else around. No one overhears.

[A] is talking with three other people. [A] tells a nasty rumour about [B]. The conversation takes place openly. [PROTAGONIST] overhears.

**‘He Said, She Said’ Pairing #1** [PROTAGONIST] is talking with [A], and two other people. [A] tells a nasty rumour about [B], [C] (a rumour hitherto unheard by [PROTAGONIST]). The conversation takes place openly. Still, no one overhears. Later, [PROTAGONIST] is talking with [C], [D], and one other person. [D] tells the same nasty rumour about [B], [E]. [PROTAGONIST] knows [D] has longstanding vendettas against [B], [E]. [PROTAGONIST] believes [D] is misinformed (earlier, [PROTAGONIST] heard this same rumour about [B], [C]). The conversation takes place behind closed doors, with no one else around. No one overhears.

[PROTAGONIST] is talking with [A], and three other people. [A] talks negatively about [B]. The conversation takes place behind closed doors, with no one else around. No one overhears. Later, [PROTAGONIST] is talking with [B], [C], and one other. [B] tells a nasty rumour about [D], [E]. [PROTAGONIST] knows [B] has longstanding vendettas against [D], [E]. [PROTAGONIST] believes [B] is misinformed (earlier, [PROTAGONIST] heard this same rumour about [C], [E]). The conversation takes place behind closed doors, with no one else around. No one overhears.

**‘He Said, She Said’ Pairing #2** [A] is talking with [B]. [A] confronts [B] directly with a nasty rumour about [B] (a rumour hitherto unheard by [PROTAGONIST]). [PROTAGONIST] knows [A] is supposedly loyal to [B]. The conversation takes place one-on-one. Still, [PROTAGONIST] overhears. Later, [PROTAGONIST] is talking with [C]. [PROTAGONIST] tells the same nasty rumour about [D]. [PROTAGONIST] is lying (earlier, [PROTAGONIST] heard this same rumour about [B]). The conversation takes place one-on-one. No one overhears.

[PROTAGONIST] is talking with four other people. [PROTAGONIST] tells a nasty rumour about [A]. The conversation takes place behind closed doors, with no one else around. No one overhears. Later, [PROTAGONIST] is talking with [B], and one other person. [B] tells a different nasty rumour about [C]. The conversation takes place openly. Still, no one overhears.

**‘He Said, She Said’ Pairing #3** [PROTAGONIST] is talking with [A], and three other people. [A] tells a nasty rumour about [B]. The conversation takes place behind closed doors, with no one else around. No one overhears. Later, [PROTAGONIST] is talking with [B], [C], and one other person. [B] tells a different nasty rumour about [A], [D]. [PROTAGONIST] knows [B] has longstanding vendettas against [A], [D]. [PROTAGONIST] believes [B] is misinformed (earlier, [PROTAGONIST] heard this same rumour about [C], [D]). The conversation takes place behind closed doors, with no one else around. No one overhears.

[A] is talking with one other person. [A] tells a nasty rumour about [B], [C], [D]. [PROTAGONIST] knows [A] is supposedly loyal to [B]. The conversation takes place one-on-one. Still, [PROTAGONIST] overhears. Later, [PROTAGONIST] is talking with [A], [C], [E], and one other person. [E] talks negatively about [F]. The conversation takes place behind closed doors, with no one else around. No one overhears.

**‘He Said, She Said’ Pairing #4** [PROTAGONIST] is talking with three other people. [PROTAGONIST] tells a nasty rumour about [A]. [PROTAGONIST] has a longstanding vendetta against [A]. The conversation takes place openly. Still, no one overhears. Later, [PROTAGONIST] is talking with [B], [C], and one other person. [B] tells a different nasty rumour about [A], [D]. [PROTAGONIST] knows [B] has longstanding vendettas against [A], [D]. [PROTAGONIST] believes [B] is misinformed (earlier, [PROTAGONIST] heard this same rumour about [C], [D]). The conversation takes place behind closed doors, with no one else around. No one overhears.

[A] is talking with [B], and one other person. [A] tells a nasty rumour about [C] (a rumour hitherto unheard by [PROTAGONIST]). [PROTAGONIST] knows [A] has a longstanding vendetta against [C]. The conversation takes place openly. [PROTAGONIST] overhears. Later, [PROTAGONIST] is talking with [D]. [D] tells a different nasty rumour about [B], [E], [F]. The conversation takes place one-on-one. No one overhears.

**‘He Said, She Said’ Pairing #5** [A] is talking with one other person. [A] tells a nasty rumour about [B], [C], [D]. [PROTAGONIST] knows [A] is supposedly loyal to [B]. The conversation takes place one-on-one. Still, [PROTAGONIST] overhears. Later, [PROTAGONIST] is talking with [E]. [E] tells the same nasty rumour about [F], [C], [D]. The conversation takes place one-on-one. No one overhears.

[PROTAGONIST] is talking with four other people. [PROTAGONIST] tells a nasty rumour about [A]. The conversation takes place behind closed doors, with no one else around. No one overhears. Later, [PROTAGONIST] is talking with [B], and three other people. [B] tells a different nasty rumour about [A]. [PROTAGONIST] knows [B] has a longstanding vendetta against [A]. The conversation takes place behind closed doors, with no one else around. No one overhears.

## Qualitative Analysis

The following text was presented for long-form analysis in Part II of the questionnaire. Note that the context was written entirely by hand.

**Context** Whitehall, late 1950s. A sclerotic Tory party limps into its seventh year of government, its cabinet marred by scandal, its fragile coalition propped up by an increasingly-emboldened cadre of avowed anti-communists. With the Prime Minister away on diplomatic business, a select group of backbenchers and party grandees has received summons to the Chelsea townhouse of Sir Geoffrey Coward, former PM. Everyone understands their government is in crisis, but no one quite knows why they have been invited here...

**Narrative of Bart Blakely, the militant Defence Secretary** (*a.i*) In the upstairs corridor, Mark Bohring was talking with someone out of sight. They were relaying that the ‘Committee Against Subversion’ is preparing to charge The Foreign Secretary with high treason. Mark segued to joking they’d happily fabricate evidence against Earl Harry Peacock if they didn’t have to work with them anymore. The conversation took place with no one else around. Still, I overheard.

(*a.ii*) Later, in the dining room: Seb Slater was talking with a backbencher. They were pondering why we’ve all been convened here tonight. Seb segued to spreading that several party grandees - Sir Geoffrey Coward and Mark Bohring amongst them - are losing confidence in The Prime Minister (a rumour hitherto unheard by myself). I knew Seb Slater was supposedly loyal to Mark Bohring. The conversation took place one-on-one. Still, I overheard.

(a.iii) Later, in the downstairs corridor, I was talking with Earl Peacock. We were pondering why we've all been convened here tonight. I segued to spreading that several party grandees - Baroness Ashna Bakshi and Sir Geoffrey Coward amongst them - are losing confidence in The Prime Minister. I have a longstanding vendetta against Sir Geoffrey. The conversation took place one-on-one. No one overheard.

(a.iv) Later, in the kitchens, Earl Peacock and I were talking with Sir Geoffrey, his aide, and Peter Tweedy. We were grumbling about the recent weather. Sir Geoffrey segued to fretting about the next round of 'Home Office Hearings'. The conversation took place behind closed doors, with no one else around. No one overheard.

(a.v) Later, in the billiards room: I was talking with Nancy Schwarz, and a couple of backbenchers. We were theorising these hearings are actually Government-sanctioned, a deliberate distraction from real scandals. I segued to joking I'd happily fabricate evidence against Earl Peacock if I didn't have to work with them anymore (a joke stolen from Sir Geoffrey's butler). I have a longstanding vendetta against Earl Peacock. The conversation takes place behind closed doors, with no one else around. No one overheard.

**Narrative of Seb Slater, the diligent Deputy Chief Whip** (b.i) In the downstairs corridor, Bart Blakely was talking with someone out of sight. They were joking they'd happily fabricate evidence against Gordon Grey if they didn't have to work with them anymore. Bart segued to asserting that the only people testifying against the cabinet have their own eyes on the leadership (a rumour hitherto unheard by myself). The conversation took place with no one else around. Still, I overheard.

(b.ii) Later, Earl Harry Peacock was talking with Bart. They were spreading that several party grandees - Baroness Ashna Bakshi and Sir Geoffrey Coward amongst them - are losing confidence in The Prime Minister. Earl Peacock segued to insinuating Baroness Bakshi is a party higher-up with distinctly Soviet leanings. I could tell Earl Peacock was misinformed about Baroness Bakshi being the so-called 'soviet sympathiser' in our midst (I knew this was actually Mark Bohring, the Chief Whip). The conversation took place one-on-one. Still, I overheard.

(b.iii) Later, in the library, I was talking with Mark, a couple of backbenchers, and Sir Geoffrey's butler. We were spreading that Nancy Schwarz is hunting for Communists on someone's explicit instructions. Mark segued to relaying that the 'Committee Against Subversion' is preparing to charge Earl Peacock with high treason. I could tell Mark was lying about Earl Peacock being the so-called 'scapegoat' in our midst (earlier, I told Mark this was actually The Foreign Secretary). The conversation took place with no one else around. No one overheard.

(b.iv) Later, in the drawing room: Mark and I were talking with Earl Peacock, and another backbencher. We were saying that Nancy Schwarz has got a 'star witness' ready to name names, by all accounts. I, confronting Mark directly, segued to joking I'd happily fabricate evidence against him if I didn't have to work with him anymore (a joke stolen from Mark). I have a longstanding vendetta against Mark. The conversation took place openly. Still, no one overheard.

(b.v) Immediately, Earl Peacock left the drawing room. Mark segued to implying that Nancy's 'mystery witness' is none other than Bart Blakely (a rumour hitherto unheard by myself). The conversation took place openly. Still, no one overheard.

**Narrative of Nancy Schwarz, the contentious MP for Torquay** (c.i) In the downstairs corridor, I was talking with Bart Blakely, and Sir Geoffrey's butler. We were making small talk about The Ashes. I, not really paying attention to whatever Bart just said, segued to divulging that Earl Harry Peacock was in a Marxist reading group back at Balliol. I was lying about Earl Peacock

being the so-called ‘Marxist intellectual’ in our midst (this was actually myself!). The conversation took place openly. Still, no one overheard.

(*c.ii*) Later, in the upstairs corridor, Seb Slater was talking with someone out of sight. They are ruminating on my very public allegations that Communists have penetrated the upper echelons of Her Majesty’s government. Seb segued to conceding that I do represent a small but frighteningly influential faction of our coalition (a rumour hitherto unheard by myself). I knew Seb had a longstanding vendetta against me. The conversation took place with no one else around. Still, I overheard.

(*c.iii*) Later, in the billiards room, I was talking with a couple of backbenchers. We were warning that, apparently, The Standard’s got ahold of something major. One of the backbenchers segued to stating straight out that Bart Blakely is prepared to run against The Prime Minister (a rumour hitherto unheard by myself). The conversation took place behind closed doors, with no one else around. No one overheard.

(*c.iv*) Later, Bart had joined us in the billiards room. We were theorising these hearings are actually Government-sanctioned, a deliberate distraction from real scandals. Bart segued to joking they’d happily fabricate evidence against Earl Peacock if I didn’t have to work with them anymore (a joke stolen from Sir Geoffrey’s butler). The conversation took place behind closed doors, with no one else around. No one overheard.

(*c.v*) Later, Earl Peacock had joined us in the billiards room, and one of the backbenchers had left. We were reflecting on recent rumours circulated in the printed press. I segued to stating straight out that Sir Geoffrey Coward is prepared to run against The Prime Minister. I have longstanding vendettas against The Prime Minister, and Sir Geoffrey. I was lying about Sir Geoffrey being the so-called ‘challenger to the leadership’ in our midst (earlier, a backbencher told me this was actually Bart Blakely). The conversation took place behind closed doors, with no one else around. No one overheard.



# Appendix B: Questionnaire Data

This appendix contains the raw, anonymised responses to this project's questionnaire, published with the informed consent of the participants. Any typos are the participants' own.

## Participant #1 Which extract is more interesting?

Pairing	'Rumour' #1	#2	#3	#4	#5	'He Said, She Said' #1	#2	#3	#4	#5
Heuristic Preferred	Y	Y	Y	Y	Y	N	Y	Y	Y	Y

*Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?* A1, B1, C1, A2, A3/B2 (overlap), B3, C2, A4/C3 (overlap), A5/C4/B4 (overlap), B5/C5 (overlap)

*Did you find it easy or difficult to complete the previous question? Why?* Difficult - hard to assert a definitive sequence of events, remember characters' sympathies/locations/statements, determine who made the joke first etc

*Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?* Blakely omits Peacock's assertion that Bakshi is a communist, possibly due to his focus on his vendetta against Sir Geoffrey

*Does any one rumour or discrepancy stand out to you? Why?* It's interesting that Blakely thinks Slater is loyal to Bohring when in fact he has a vendetta against him, and thinks he is a communist despite this actually being true of Schwarz - both Blakely and Slater's reasons for their erroneous beliefs are left ambiguous

*(Optional) Do you have any other thoughts on these narratives?* [No response]

*(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?*

Statement	"I am an avid reader"	"I enjoy analysing media"	"I am familiar with procedural generation"
Agreement	4	5	1

## Participant #2 Which extract is more interesting?

Pairing	'Rumour' #1	#2	#3	#4	#5	'He Said, She Said' #1	#2	#3	#4	#5
Heuristic Preferred	Y	N	Y	Y	Y	Y	Y	N	Y	N

*Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?* (a.i) > (a.ii) > (a.iii) > (c.i) > (b.i) > (b.ii) > (b.iii) > (c.ii) > (c.iii) > (a.iv) > (b.iv)(b.v) > (a.v) = (c.iv) > (c.v)

**Did you find it easy or difficult to complete the previous question? Why?** I found it very difficult for a few reasons. Because all of the information is presented in a linear order, it then becomes harder to retroactively adjust the order to fit chronologically. In addition, for me personally there were too many small moving parts and named characters to keep track of in such a small scenario, which caused me to personally overthink the question and scenario much more - whilst having difficulty parsing the tiny important details.

**Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?** c.ii refers to Seb talking to someone in the upstairs corridor, but none of Seb's paragraphs take place in the upstairs corridor. a.ii refers to Seb talking in the dining room, but none of Seb's paragraphs take place in the dining room. This two omitted scenes makes it seem like Seb is a more unreliable narrator - where the other characters eavesdrop on scenes that are not included in his narrative.

**Does any one rumour or discrepancy stand out to you? Why?** I enjoyed the scene in (b.iv) where Seb directly confronts Mark using his own joke against him out of spite due to their long running vendetta. Similarly, I enjoyed the scene in (c.i) where Nancy spreads a rumour that Harry Peacock is a Marxist to divert attention from Nancy themselves actually being the Marxist. These scenes stand out in a good way because the feature the protagonists directly causing conflict regarding their own personal alignments - which I personally find more engaging than rumours being spread behind people's backs.

**(Optional) Do you have any other thoughts on these narratives?** [No response]

**(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?**

Statement Agreement	"I am an avid reader" 2	"I enjoy analysing media" 5	"I am familiar with procedural generation" 5
---------------------	-------------------------	-----------------------------	--

**Participant #3 Which extract is more interesting?**

Pairing	'Rumour' #1	#2	#3	#4	#5	'He Said, She Said' #1	#2	#3	#4	#5
Heuristic Preferred	Y	Y	Y	Y	Y	N	N	N	N	N

**Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?** aii, bi, cii

**Did you find it easy or difficult to complete the previous question? Why?** Easy, only relevant informations are location and whom they are talking to.

**Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?** All 1st person are in perspective of the character, it is possible that the character omitts information not favourable to the character.

**Does any one rumour or discrepancy stand out to you? Why?** No, too many rumors, all are equally fasle/true.

**(Optional) Do you have any other thoughts on these narratives?** Too much repetition of no one hears, conversation between party memebbers seems in-organic, CRPG/RPG with party memebbers have more organic conversations as party member talk & listen to each other, but also listen and speak when protagonist is speaking to NPC.

*(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?*

Statement Agreement	“I am an avid reader” 2	“I enjoy analysing media” 4	“I am familiar with procedural generation” 5
---------------------	----------------------------	--------------------------------	---

**Participant #4** Which extract is more interesting?

Pairing	‘Rumour’ #1	#2	#3	#4	#5	‘He Said, She Said’ #1	#2	#3	#4	#5
Heuristic Preferred	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

*Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?* Ai,Cii, Aii,ci,aiai,bii—aiv—,biv,bv av,Cv

*Did you find it easy or difficult to complete the previous question? Why?* Difficult without note taking keeping track of details, characters and locations proved harder than anticipated. Essentially I couldn’t be bothered using my laptop so I could move the order easier. The repetitive phrasing used by the generator also makes it complicated to separate the narratives mentally. Also it’s like cluedo and I don’t like cluedo

*Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?* [No response]

*Does any one rumour or discrepancy stand out to you? Why?* Bart Blakey running against pm because it’s only mentioned once and surely he would know which people are losing confidence in the pm then which he says he doesn’t.

*(Optional) Do you have any other thoughts on these narratives?* [No response]

*(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?*

Statement Agreement	“I am an avid reader” 1	“I enjoy analysing media” 3	“I am familiar with procedural generation” 5
---------------------	----------------------------	--------------------------------	---

**Participant #5** Which extract is more interesting?

Pairing	‘Rumour’ #1	#2	#3	#4	#5	‘He Said, She Said’ #1	#2	#3	#4	#5
Heuristic Preferred	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

*Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?* a-iii, a.i, b.i, c.i, a.ii, b.ii, a.v, b.iii, c.iii, a.iv, c.iv, b.iv, b.v

*Did you find it easy or difficult to complete the previous question? Why?* It was difficult due to the various characters and scenarios but mostly because of the fragmented information that was given.it was also hard to figure out the flow of the scenarios since some characters were hearing things they weren’t meant to and there were multiple locations involved.

*Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?* Yes the narratives omit relevant information. These omissions seem to be intentional and driven by both character and literary motivations.

### Character Motivations:

- I would say that each character had their own goals and motivations and these are the reason for what they choose to reveal or not.

An example would be to:

- Manipulate others: In Bart Blakey and Nancy Schwarz's interactions they spread rumours/make false allegations to further their agendas to manipulate others or to have more control over the way the conversation goes.
- Protect themselves: Earl Peacock lies about his intentions and his own affiliations to avoid being seen as a "marxist intellectual" so i would say he omits information to protect himself from a negative consequence.
- Have their political goals advance: It is seen with Seb Slater and Bart Blakely that they position themselves politically by only sharing some information and hiding other parts. This could be because they want to weaken their opposition or to strengthen those on their side.

### Literary Techniques:

- There is Suspense or Mystery as there are gaps in the narrative when I am reading i feel i need to piece the full story together which brings suspense.
- There are also complicated characters and you can see there are some internal conflicts.

***Does any one rumour or discrepancy stand out to you? Why?*** In the (a.iv) narrative we see Bart Blakely say that Sir Geoffrey Coward is worried about the upcoming "home office hearings" in a conversation in the kitchen. but in the (a.i) one which i would assume happens earlier you can see that Bart Blakely heard that Mark B was talking about the "committee against subversion" in preparation to charge the foreign secretly with high treason. however, is it not mentioned of home office hearings. so this makes me question if bart's information is accurate or credible. this however could be intentional to start a rumour or manipulate the information to benefit him. or if this is the simulation generating something inconsistent showing how misunderstandings can happen in real world scenarios.

***(Optional) Do you have any other thoughts on these narratives? [No response]***

***(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?***

Statement	"I am an avid reader"	"I enjoy analysing media"	"I am familiar with procedural generation"
Agreement	2	3	1

### Participant #6 Which extract is more interesting?

Pairing	'Rumour' #1	#2	#3	#4	#5	'He Said, She Said' #1	#2	#3	#4	#5
Heuristic Preferred	Y	Y	Y	Y	N	Y	Y	Y	N	Y

**Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?**

a.iv  
a.ii  
a.iii, b.ii (overlapping)  
b.i  
c.i  
c.iii  
a.v, c.iv (overlapping)  
a.i  
b.iv  
b.v  
b.iii  
c.v  
c.ii

**Did you find it easy or difficult to complete the previous question? Why?** No. The lack of a completely contiguous narrative made it difficult to guess where some of the pieces were located on the timeline. Additionally, many of the hints may have been too subtle for me to notice.

**Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?** It's difficult to say exactly what was omitted across the stories without guessing and making a fool of myself. However, it does seem that omissions occurred based on who the person was talking with and who they had vendettas against. The "he said, she said" is very difficult to trace as an observer. Nevertheless, the accusations tended to be focused towards people that the narrator had an established vendetta against.

The vendetta's seemed to drive the dialogue with accusations being made towards a character's political enemies. Also, Nancy lying to protect her identity as the Marxist intellectual provides a significant incentive to omit key elements in her exchanges with the other characters.

**Does any one rumour or discrepancy stand out to you? Why?** In b.ii: "I could tell Earl Peacock was misinformed about Baroness Bakshi being the so-called 'soviet sympathizer' in our midst (I knew this was actually Mark Bohring, the Chief Whip)". This stood out as being rather isolated and conflicting. From what I read, it seemed like the sympathizer was Nancy. Also, this did not fit the mental image of Mark that had been created by reading the story. The random inclusion of a new character in Bakshi, also made it feel slightly unrelated.

**(Optional) Do you have any other thoughts on these narratives?** With this simulated storyline, I find it really interesting how the relations between characters are displayed.

However, the biggest takeaway for myself is that there is no guarantee that the story is complete. With human created stories, most characters have a pre-determined role and the story has a clear arc from start to end.

The issue I found with reading these segments was (in a similar vein to the Bakshi appearance). I have no way of knowing if I am getting to entire story. Are all of the characters mentioned? Is there an ending that this information is building towards?

Now with that in mind, this seems remarkable for depth of character. This complexity of interconnection seems difficult to manually create. I feel that this style of simulation would be great to partner with human supervision to extract and refine stories from what could be an extremely complex, simulated environment.

*(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?*

Statement Agreement	"I am an avid reader" 4	"I enjoy analysing media" 5	"I am familiar with procedural generation" 5
---------------------	-------------------------	-----------------------------	--

**Participant #7 Which extract is more interesting?**

Pairing	'Rumour' #1	#2	#3	#4	#5	'He Said, She Said' #1	#2	#3	#4	#5
Heuristic Preferred	N	Y	Y	Y	Y	Y	Y	N	Y	N

*Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?* (a.i)(b.i)(c.i)(a.ii)(c.ii)(b.ii)(a.iii)(b.iii)(c.iii)(c.iv)(a.v)(c.v)(b.iv)(b.v)(a.iv):

*Did you find it easy or difficult to complete the previous question? Why?* Not such easy because the timeline isn't clear. And events in the same location are more likely to occur one after the other, as well as the fact that if someone is in two different places in a short period of time, both events are less likely to occur at the same time.

*Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?* Each character has their own position and point of view. They may selectively remember or convey certain information, especially when it conflicts with their interests or motives. and may deliberately omit information about their own misbehaviour in order to avoid responsibility or damage to their image.

*Does any one rumour or discrepancy stand out to you? Why?* In Nancy's account, there is a reference to Bart Blakely's readiness to challenge the Prime Minister, which is not explicitly mentioned in the other accounts. This may suggest Bart's dissatisfaction with the current regime. There are different views on the level of trust in the Prime Minister within the party: Both Bart and Earl Peacock refer to some of the party bigwigs (such as Sir Geoffrey Coward and Baroness Ashna Bakshi) losing confidence in the Prime Minister, but there is a difference in the names conveyed between the two.

*(Optional) Do you have any other thoughts on these narratives?* [No response]

*(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?*

Statement Agreement	"I am an avid reader" 3	"I enjoy analysing media" 4	"I am familiar with procedural generation" 4
---------------------	-------------------------	-----------------------------	--

**Participant #8 Which extract is more interesting?**

Pairing	'Rumour' #1	#2	#3	#4	#5	'He Said, She Said' #1	#2	#3	#4	#5
Heuristic Preferred	Y	Y	N	Y	N	N	N	N	N	Y

*Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?*

A.i a.ii a.iv a.v a.iii  
b.i b.ii b.iii b.iv b.v  
c.i c.iii c.v v.iv c.ii

*Did you find it easy or difficult to complete the previous question? Why?* Extremely difficult as I struggled to follow all of the individual storylines coherently

*Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?* Bart Blakeley doesn't mention his intention to run for PM but I suppose that fits in with the whole theme of rumours and secrecy.

*Does any one rumour or discrepancy stand out to you? Why?* Bart Blakley running against the PM as it was never alluded to from his own POV stories.

*(Optional) Do you have any other thoughts on these narratives?* This whole thing really confused me. I'm also just very bad at inferring if someone is lying or not unless it's made very clear.

*(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?*

Statement Agreement	"I am an avid reader" 2	"I enjoy analysing media" 2	"I am familiar with procedural generation" 1
---------------------	-------------------------	-----------------------------	--

**Participant #9** *Which extract is more interesting?*

Pairing	'Rumour' #1 #2 #3 #4 #5					'He Said, She Said' #1 #2 #3 #4 #5				
Heuristic Preferred	N	Y	N	Y	N	Y	N	N	N	Y

*Considering all three narratives, please sort events (a.i)-(c.v) into chronological order. Do any events overlap?* a.I,b.iii,a.iii,c.v,c.i,c.iv,a.ii,c.ii,b.i,b.ii,c.iii,b.v,b.iv,a.iv,a.i

*Did you find it easy or difficult to complete the previous question? Why?* difficult because there were so many options

*Does any narrative(s) omit any relevant information? If so, why do you think the omission(s) has been made?* no

*Does any one rumour or discrepancy stand out to you? Why?* Earl peacock insinuating bakshni is a party higher up with soviet leanings

*(Optional) Do you have any other thoughts on these narratives?* [No response]

*(Optional) On a scale of 1 (strongly disagree) to 5 (strongly agree), do you agree with the following statements?*

Statement Agreement	"I am an avid reader" 4	"I enjoy analysing media" 3	"I am familiar with procedural generation" 3
---------------------	-------------------------	-----------------------------	--

# References

- Adams, T. & Adams, Z. (2006), *Slaves to Armok: God of Blood Chapter II: Dwarf Fortress*, [Video Game] Bay 12 Games.
- Antonelli, P. (2011), *Talk to Me: Design and the Communication Between People and Objects*, [Exhibition] MoMA, New York, NY, USA: 24 July-7 November.
- Antoun, C., Antoun, M., Ryan, J., Samuel, B., Swanson, R. & Walker, M. A. (2015), Generating Natural Language Retellings from *Prom Week* Play Traces, in ‘6th Workshop on Procedural Content Generation, International Conference on the Foundations of Digital Games’, Monterey, CA, USA: 23 June.
- Aristotle (330 B.C.), *The Poetics of Aristotle*, London, UK: Duckworth Books.
- Aylett, R. (1999), Narrative in Virtual Environments - Towards Emergent Narrative, in ‘AAAI 1999 Fall Symposium on Narrative Intelligence’, North Falmouth, MA, USA: 5-7 November, pp. 83–86.
- Aylett, R. (2000), Emergent Narrative, Social Immersion and “Storification”, in ‘The 1st International Workshop on Narrative and Interactive Learning Environments’, Edinburgh, Scotland: 30 August-1 September, pp. 35–44.
- Aylett, R., Louchart, S., Dias, J., Paiva, A. & Vala, M. (2005), FearNot! - An Experiment in Emergent Narrative, in ‘5th International Workshop on Intelligent Virtual Agents’, Kos, Greece: 12-14 September, pp. 305–316.
- Bae, B.-C. & Young, R. M. (2008), A Use of Flashback and Foreshadowing for Surprise Arousal in Narrative Using a Plan-Based Approach, in ‘First Joint International Conference on Interactive Digital Storytelling’, Erfurt, Germany: 26-29 November, p. 156–167.
- Brooks, C. & Warren, R. P. (1943), *Understanding Fiction*, 2<sup>nd</sup> edn, New York, NY, USA: Appleton-Century-Crofts.
- Burkinshaw, R. (2009), ‘Alice and Kev: The Story of Being Homeless in *The Sims 3*’, Available at: <https://aliceandkev.wordpress.com/>. (Accessed: 23 July 2023).
- Christie, A. (1943), *Five Little Pigs*, London, UK: Harper Collins.
- Compton, K. (2019), ‘Getting Started with Generators’, in T. X. Short & T. Adams, eds, ‘*Procedural Storytelling in Game Design*’, Boca Raton, FL, USA: Taylor & Francis, pp. 3–17.
- Compulsion Games (2018), *We Happy Few*, [Video Game] Gearbox Publishing.
- Dahlbäck, N., Jönsson, A. & Ahrenberg, L. (1993), Wizard of Oz Studies - Why and How, in ‘ACM 1993 International Workshop on Intelligent User Interfaces’, Orlando, FL, USA: 34-7 January, pp. 193–200.



- Eladhari, M. P. (2018), Re-Tellings: The Fourth Layer of Narrative as an Instrument for Critique, in '11th International Conference on Interactive Digital Storytelling', Dublin, Ireland: 5-8 December, pp. 65–78.
- Epstein, A. (2019), 'Dirty Procedural Narrative in *We Happy Few*', in T. X. Short & T. Adams, eds, 'Procedural Storytelling in Game Design', Boca Raton, FL, USA: Taylor & Francis, pp. 227–240.
- Garbe, J. (2018), 'Simulation of History and Recursive Narrative Scaffolding', <http://project.jacobgarbe.com/simulation-of-history-and-recursive-narrative-scaffolding/>. (Accessed: 12 August 2023).
- Genette, G. (1980), *Narrative Discourse: An Essay in Method*, 1<sup>st</sup> edn, Ithaca, NY, USA: Cornell University Press.
- Gervás, P. (2012), From the Fleece of Fact to Narrative Yarns: A Computational Model of Composition, in 'CMN '12: The Third Workshop on Computational Models of Narrative', Istanbul, Turkey: 26-27 May.
- Gervás, P. (2013), Stories from Games: Content and Focalization Selection in Narrative Composition, in 'I Spanish Symposium on Entertainment Computing', Madrid, Spain: 18-19 September.
- Gervás, P. (2014), 'Composing Narrative Discourse for Stories of Many Characters: A Case Study over a Chess Game', *Literary and Linguistic Computing* **29**(4), 511–531.
- Graham, D. (2014), 'An Introduction to Utility Theory', in S. Rabin, ed., 'Game AI Pro: Collected Wisdom of Game AI Professionals', Boca Raton, FL, USA: Taylor & Francis, pp. 113–126.
- Grinblat, J. (2017), 'Emergent Narratives and Story Volumes', in T. X. Short & T. Adams, eds, 'Procedural Generation in Game Design', Boca Raton, FL, USA: Taylor & Francis, pp. 199–207.
- Guerlac, S. (1980), 'Narrative Discourse, an Essay in Method by Gerard Genette and Jane E. Lewin', *Comparative Literature* **95**(5), 1414–1421.
- Hoek, M., Theune, M. & Linssen, J. (2014), Generating Game Narratives with Focalization and Flashbacks, in 'Games and Natural Language Processing Workshop, AIIDE Workshops 2014', Raleigh, NC, USA: 3-7 October, pp. 9–14.
- Jones, J. D. (2022), 'Authorial Burden', in C. Hargood, D. E. Millard, A. Mitchell & U. Spierling, eds, 'The Authoring Problem: Challenges in Supporting Authoring for Interactive Digital Narratives', Human-Computer Interaction Series, Cham, Switzerland: Springer, pp. 47–63.
- Karth, I. (2023), Recomposing Procgen, PhD thesis, University of California, Santa Cruz.
- Kawitzky, F. R. (2020), 'Magic Circles: Tabletop Role-Playing Games as Queer Utopian Method', *Performance Research* **25**(8), 129–136.
- Kim, J. (2018), 'Subjective Simulation Design: Ludonarrative Congruence in *The Shrouded Isle*', (online video) Available at: <https://www.youtube.com/watch?v=oV5PjvcVky8>. (Accessed: 18 June 2023).
- King, S. (1974), *Carrie*, New York, NY, USA: Doubleday.
- Klein, S., Aeschlimann, J. F., Balsiger, D. F., Converse, S. L., Court, C., Foster, M., Lao, R., Oakley, J. D. & Smith, J. (1973), Automatic Novel Writing: A Status Report, Technical report, University of Wisconsin, Madison.

- Kreminski, M. (2023a), “Generator’s Haunted”: A Brief, Spooky Account of Hauntological Effects in the Player Experience of Procedural Generation, in ‘*18th International Conference on the Foundations of Digital Games*’, Lisbon, Portugal: 12-14 April, pp. 1–3.
- Kreminski, M. (2023b), Toward Better Gossip Simulation in Emergent Narrative Systems, in ‘*IEEE Conference of Games 2023*’, Boston, MA, USA: 21-24 August.
- Kreminski, M., Dickinson, M. & Mateas, M. (2021), Winnow: A Domain-Specific Language for Incremental Story Sifting, in ‘*The 17th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*’, Online: 11-15 October, pp. 156–163.
- Kreminski, M., Dickinson, M. & Wardrip-Fruin, N. (2019), Felt: A Simple Story Sifter, in ‘*12th International Conference on Interactive Digital Storytelling*’, Little Cottonwood Canyon, UT, USA: 19-22 November, pp. 267–281.
- Kreminski, M., Dickinson, M., Wardrip-Fruin, N. & Mateas, M. (2022), Select the Unexpected: A Statistical Heuristic for Story Sifting, in ‘*15th International Conference on Interactive Digital Storytelling*’, Santa Cruz, CA, USA: 4-7 December, pp. 292–308.
- Kreminski, M. & Mateas, M. (2021), Toward Narrative Instruments, in ‘*14th International Conference on Interactive Digital Storytelling*’, Tallinn, Estonia: 7-10 December, pp. 499–508.
- Kreminski, M., Wardrip-Fruin, N. & Mateas, M. (2020), Toward Example-Driven Program Synthesis of Story Sifting Patterns, in ‘*The 12th Intelligent Narrative Technologies Workshop, AIIDE Workshops 2020*’, Worcester, MA, USA: 19-20 October.
- Kreminski, M., Wardrip-Fruin, N. & Mateas, M. (2022), ‘Authoring For Story Sifters’, in C. Hargood, D. E. Millard, A. Mitchell & U. Spierling, eds, ‘*The Authoring Problem: Challenges in Supporting Authoring for Interactive Digital Narratives*’, Human-Computer Interaction Series, Cham, Switzerland: Springer, pp. 207–220.
- Kurosawa, A. (1950), *Rashomon*, [Feature Film] Los Angeles, CA, USA: RKO Radio Pictures.
- Landert, D. (2021), ‘The Spontaneous Co-Creation of Comedy: Humour in Improvised Theatrical Fiction’, *Journal of Pragmatics* **173**, 68–87.
- Leong, W., Porteous, J. & Thangarajah, J. (2022), Automated Sifting of Stories From Simulated Storyworlds, in ‘*Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22)*’, Vienna, Austria: 23-29 July, pp. 4950–4956.
- Lewis, M. (2017), ‘Choosing Effective Utility-Based Considerations’, in S. Rabin, ed., ‘*Game AI Pro 3: Collected Wisdom of Game AI Professionals*’, Boca Raton, FL, USA: Taylor & Francis, pp. 167–178.
- Livingston, C. (2020), ‘This *Crusader Kings 3* Player Ate the Pope’, Available at: <https://www.pcgamer.com/this-crusader-kings-3-player-ate-the-pope/>. (Accessed: 18 July 2023).
- Louchart, S. & Aylett, R. (2004a), ‘Narrative Theory and Emergent Interactive Narrative’, *Int. J. Continuing Engineering Education and Lifelong Learning* **14**, 506–518.
- Louchart, S. & Aylett, R. (2004b), The Emergent Narrative Theoretical Investigation, in ‘*The 2004 Conference on Narrative and Interactive Learning Environments*’, Edinburgh, Scotland: 10-13 August, pp. 21–28.

- Maher, J. (2013), ‘Free Fall, Part 2: Murder on the Zinderneuf’, Available at: <https://www.filfre.net/2013/02/free-fall-part-2-murder-on-the-zinderneuf/>. (Accessed: 8 January 2023).
- Mateas, M. (2002), Interactive Drama, Art, and Artificial Intelligence, PhD thesis, Carnegie Mellon University, Santa Cruz.
- Mateas, M. & Ferguson, B. (2014), ‘Modular, Reusable Social Behavior In Video Game AI’, (online video) Available at: <https://www.youtube.com/watch?v=BSOB19aE9RM>. (Accessed: 12 June 2023).
- Mateas, M. & Stern, A. (2005), Build It to Understand It: Ludology Meets Narratology in Game Design Space, in ‘2005 DiGRA International Conference: Changing Views: Worlds in Play’, Vol. 3, Vancouver, B.C., Canada: June 16-20.
- Maxis (2000), *The Sims*, [Video Game] Electronic Arts.
- McCoy, J., Treanor, M., Samuel, B., Mateas, M. & Wardrip-Fruin, N. (2011), *Prom Week: Social Physics as Gameplay*, in ‘6th International Conference on the Foundations of Digital Games’, Bordeaux, France: 29 June-1 July, pp. 319–321.
- Meehan, J. R. (1977), TALE-SPIN, An Interactive Program that Writes Stories, in ‘*Proceedings of the Fifth International Joint Conference on Artificial Intelligence (IJCAI-77)*’, Cambridge, MA, USA: 22-25 August.
- Méndez, G. & Gervás, P. (2023), Using ChatGPT for Story Sifting in Narrative Generation, in ‘*14th International Conference on Computational Creativity*’, Waterloo, ON, Canada: 19-23 June.
- Paradox Development Studio (2020), *Crusader Kings III*, [Video Game] Paradox Interactive.
- Procedural Arts (2005), *Façade*, [Video Game] Procedural Arts.
- Ryan, J. (2018), Curating Simulated Storyworlds, PhD thesis, University of California, Santa Cruz.
- Ryan, J. & Mateas, M. (2017), ‘Simulating Character Knowledge Phenomena in *Talk of the Town*’, in S. Rabin, ed., ‘*Game AI Pro 3: Collected Wisdom of Game AI Professionals*’, Boca Raton, FL, USA: Taylor & Francis, pp. 433–448.
- Ryan, J., Mateas, M. & Wardrip-Fruin, N. (2015), Open Design Challenges for Interactive Emergent Narrative, in ‘*8th International Conference on Interactive Digital Storytelling*’, Copenhagen, Denmark: 30 November-4 December, pp. 14–26.
- Ryan, J., Summerville, A., Mateas, M. & Wardrip-Fruin, N. (2015), Toward Characters Who Observe, Tell, Misremember, and Lie, in ‘*2nd Workshop on Experimental AI in Games*’, Santa Cruz, CA, USA: November.
- Ryan, J., Summerville, A. & Samuel, B. (2016), Bad News: A Game Of Death And Communication, in ‘*2016 CHI Conference on Human Factors in Computing Systems*’, Vol. Extended Abstracts, San Jose, CA, USA: 7-12 May, pp. 160–163.
- Short, E. (2006), *Glass*, [Video Game] self-published.
- Short, E. (2016), ‘Beyond Branching: Quality-Based, Salience-Based, and Waypoint Narrative Structures’, Available at: <https://emshort.blog/2016/04/12/beyond-branching-quality-based-and-salience-based-narrative-structures/>. (Accessed: 1 February 2023).

- Short, E. (2019), ‘Curating Simulated Storyworlds (James Ryan) – Ch 1-3’, Available at: <https://emshort.blog/2019/05/21/curating-simulated-storyworlds-james-ryan/>. (Accessed: 17 June 2023).
- Spierling, U. (2007), Adding Aspects of “Implicit Creation” to the Authoring Process in Interactive Storytelling, in ‘*4th International Conference on Virtual Storytelling*’, Vancouver, B.C., Canada: June 16-20, pp. 13–25.
- Swartjes, I. (2010), Whose Story is it Anyway? How Improv Informs Agency and Authorship of Emergent Narrative, PhD thesis, University of Twente, Enschede.
- Sych, S. (2020), When the Fourth Layer Meets the Fourth Wall: The Case for Critical Game Retellings, in ‘*13th International Conference on Interactive Digital Storytelling*’, Bournemouth, England: 3-6 November, pp. 203–211.
- ten Brinke, H., Linssen, J. & Theune, M. (2014), Hide and Sneak: Story Generation with Characters that Perceive and Assume, in ‘*The 10th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*’, Raleigh, NC, USA: 3-7 October, pp. 174–180.
- Wardrip-Fruin, N. (2007), Three Play Effects: *Eliza*, *Tale-Spin*, and *SimCity*, in ‘*Digital Humanities 2007*’, Urbana, IL, USA: 2-8 June.

# Bibliography

- Carlisle, P. (2014), ‘A Simple and Practical Social Dynamics System’, in S. Rabin, ed., ‘*Game AI Pro: Collected Wisdom of Game AI Professionals*’, Boca Raton, FL, USA: Taylor & Francis, pp. 441–450.
- Compton, K. (2017), ‘Generative Art Toys’, in T. X. Short & T. Adams, eds, ‘*Procedural Generation in Game Design*’, Boca Raton, FL, USA: Taylor & Francis, pp. 161–173.
- Compton, K., Kybartas, Q. & Mateas, M. (2015), Tracery: An Author-Focused Generative Text Tool, in ‘*8th International Conference on Interactive Digital Storytelling*’, Copenhagen, Denmark: 30 November–4 December, pp. 154–161.
- Cook, M. (2019), ‘Ethical Procedural Generation’, in T. X. Short & T. Adams, eds, ‘*Procedural Storytelling in Game Design*’, Boca Raton, FL, USA: Taylor & Francis, pp. 49–62.
- Crawford, C. (2003), *On Game Design*, San Francisco, CA, USA: New Riders Press.
- Freehold Games (2015), *Caves of Qud*, [Video Game] Kitfox Games.
- Grinblat, J. (2019), ‘Generating Histories’, in T. X. Short & T. Adams, eds, ‘*Procedural Storytelling in Game Design*’, Boca Raton, FL, USA: Taylor & Francis, pp. 179–192.
- Johnson-Bey, S. & Mateas, M. (2021), Centrifuge: A Visual Tool for Authoring Sifting Patterns for Character-Based Simulationist Story Worlds, in ‘*Workshop of Programming Languages in Entertainment, AIIDE Workshops 2021*’, Online: 11–12 October.
- Kim, J. (2019), ‘Generating Personalities in *The Shrouded Isle*’, in T. X. Short & T. Adams, eds, ‘*Procedural Storytelling in Game Design*’, Boca Raton, FL, USA: Taylor & Francis, pp. 303–316.
- Kitfox Games (2017), *The Shrouded Isle*, [Video Game] Kitfox Games.
- Kline, D. (2019), ‘Dramatic Play in The Sims’, in T. X. Short & T. Adams, eds, ‘*Procedural Storytelling in Game Design*’, Boca Raton, FL, USA: Taylor & Francis, pp. 113–125.
- Kreminski, M., Wardrip-Fruin, N., Acharya, D., Junius, N., Oliver, E., Compton, K., Dickinson, M., Focht, C., Mason, S. & Mazeika, S. (2019), Cozy Mystery Construction Kit: Prototyping Toward an AI-Assisted Collaborative Storytelling Mystery Game, in ‘*FDG ‘19: The Fourteenth International Conference on the Foundations of Digital Games*’, San Luis Obispo, CA, USA: 26–30 November.
- Kybartas, Q. (2022), ‘Quantitative Analysis of Emergent Narratives’, in C. Hargood, D. E. Millard, A. Mitchell & U. Spierling, eds, ‘*The Authoring Problem: Challenges in Supporting Authoring for Interactive Digital Narratives*’, Human-Computer Interaction Series, Cham, Switzerland: Springer, pp. 321–334.

- Kybartas, Q., Verbrugge, C. & Lessard, J. (2020), ‘Tension Space Analysis for Emergent Narrative’, *IEEE Transactions on Games* **13**(2), 146–159.
- Manning, C. (2019), ‘Tarot as Procedural Storytelling’, in T. X. Short & T. Adams, eds, *Procedural Storytelling in Game Design*, Boca Raton, FL, USA: Taylor & Francis, pp. 339–350.
- Short, E. (2015), *The Annals of the Parrigues*, self-published.
- Short, E. (2019), ‘Curating Simulated Storyworlds (James Ryan) – Ch 4-5’, Available at: <https://emshort.blog/2019/05/23/curating-simulated-storyworlds-james-ryan-2/>. (Accessed: 17 June 2023).
- Short, E. (2019), ‘Curating Simulated Storyworlds (James Ryan) – Ch 6f’, Available at: <https://emshort.blog/2019/05/28/curating-simulated-storyworlds-james-ryan-ch-6f/>. (Accessed: 17 June 2023).
- Short, T. X. (2019), ‘Maximizing the Impact of Generated Personalities’, in T. X. Short & T. Adams, eds, *Procedural Storytelling in Game Design*, Boca Raton, FL, USA: Taylor & Francis, pp. 271–282.
- Slamdance Film Festival (2017), ‘Slamdance DIG Spotlight - BAD NEWS’, (online video) Available at: <https://www.youtube.com/watch?v=jrrpLFS9Zcw>. (Accessed: 10 July 2023).
- Smith, G. (2017), ‘Understanding the Generated’, in T. X. Short & T. Adams, eds, *Procedural Generation in Game Design*, Boca Raton, FL, USA: Taylor & Francis, pp. 231–244.