# University of Southampton

Faculty of Engineering and Physical Sciences

School of Electronics and Computer Science

**The Impact of Paratext on Readers of Generative Literature: Human evaluation of generated text**

by

**Lesia Tkacz (Alexandra Tkacz)**

0000-0002-5570-406X

Thesis for the degree of Doctor of Philosophy

June 2025

# University of Southampton

## Abstract

Faculty of Engineering and Physical Sciences

Electronics and Computer Science

Doctor of Philosophy

The Impact of Paratext on Readers of Generative Literature: Human evaluation of generated text

by

Lesia Tkacz

Generative literature has been emerging as a creative form since before the digital era, with one of its defining characteristics being the particular circumstances of its production. Since 2013, the National Novel Generation Month (NaNoGenMo) challenge has been one of the most prolific organized events and archives of generative literature in the form of computer generated novels. Authors present their generated novel projects as a collection of digital files and components which have different functions: the generated text, the code used to generate it, the author's notes and creative intentions, the input data used by the code, and so forth. While generated novels and other forms of generative literature have been researched to a limited extent, little has been investigated about readers and their experience, especially beyond an implied reader perspective as theorized through an academic lens. With the presence of generative qualities and processes becoming more apparent and increasingly present in today's tools, in (creative) media, and in the sociotechnical fabric more generally, it is becoming increasingly relevant to research actual readers' understanding and reception of generated text. However, as both an emerging creative and technical form with many different components, it's unclear which components are central to impacting readers' understanding and reception of a generated novel.

The impact of paratext on the interpretation and reception of media works has long been accepted and theorized in literary studies. Paratext refers to the framing, supplementary, and periphery elements which can accompany a work. Paratext's perceived ability to steer readers' interpretations about a piece was first detailed by structural literary theorist Gérard Genette's established but highly contested conceptualization of paratext. Genette theorizes how paratext functions in the context of books and their consumption, and his conceptualization has since been frequently used to describe new media works and their audiences, although not without challenge by new media forms. However, within this literary and media space there is little in the way of participant-based research which tests assumptions about how paratext functions to impact different readers' interpretation and reception of a work. Nevertheless, this research project demonstrates that Genette's paratextual conceptualization lends itself to being easily operationalized such that the impact of different elements on reading reception can be measured and unpacked using quantitative and qualitative methods.

Focusing on two distinct NaNoGenMo generated novels, this research project conceptualizes the components of a generated novel as paratextual elements in order to test the ways in which paratext impacts potential readers' interpretation and reception of a generated novel. Reception is decomposed into different value dimensions such as literary, creative, and

technical value, as well as understanding, interest, and enjoyment. The research project further investigates how reader's computer programming skills or literary reading experience might interplay with the impact of the paratextual elements, as well as how cultural characteristics of a generated novel might shape interpretation and reception.

The research project is carried out as an explanatory sequential mixed-method design beginning with an online reading experiment where the generated novels are presented in one of three different paratextual conditions. Participant responses are recorded through a survey and a written review where participants give their personal opinions about the work they read. The survey results suggest that while the presence of the paratextual elements in the generated novels have a measurable impact on readers' perceptions across virtually all the value dimensions, the presence of the generated text itself doesn't significantly impact reader valuations. Further, the survey results also suggest that the generated novels' perceived literary value may be largely resistant to being impacted by the paratext.

In addition to further studying paratext's impact on reader interpretation and reception, this unexpected literary value result is further explored in the written review data using Qualitative Content Analysis, and in a second in-person reading group workshop study where a semi-structured group interview is run to discuss reader's impressions of the same generated novels read in printed paperback book form – a physical paratext. The workshop study data is analyzed using Deductive Qualitative Analysis and develops a working conceptual model of a printed generated novel's reading and interpretation process. Here, latent links to a generated novel's technical aspects and a reader's own creative framing are highlighted as primary themes. The analysis develops further and crystalizes into a novel contribution in the form of a minimalist theoretical refinement of Genette's paratextual conceptualization. The findings also suggest that the readers' choice to engage with the generated text ergodically and to interpret it through a creative framing appear to be factors which contribute to the readers perceiving their reading experience to be more enjoyable. Nevertheless, despite workshop discussion about the physical paratext, references to authorial paratext, and the generated novels' cultural value, the literary value of generated novels appears to remain challenging for readers to fully accept.

Keywords: Computer Generated Novel, Paratext, NaNoGenMo, Readers, Generated Text, AI

# Table of Contents

# Table of Tables

# Table of Figures

# Research Thesis: Declaration of Authorship

Print name: Lesia Tkacz

Title of thesis: The Impact of Paratext on Readers of Generative Literature: Human evaluation of generated text

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before

# Acknowledgements

I am deeply grateful to my wonderful joint-supervision team Professor David Millard and Associate Professor Dr. Megen de Bruin-Molé for their support and guidance at every stage of this research project. I have finished every supervision meeting feeling happy and ultra motivated – thank you!

A big thank you to Dr. Vasyl Pihur for help with getting additional results from *RankAggreg, an R package for weighted rank aggregation* (2009), and to Professor Somnath Datta.

I thank my examination team, Associate Professor Dr. Lyle Skains and Professor Daniel Ashton for their time, patience, and valuable feedback. I would also like to thank everybody at the Web Science CDT, the Department of Art and Media Technology and PGR group at the Winchester School of Art, and the University of Southampton for their solid support and opportunities.

# Chapter 1 Introducing the research

This dissertation reports on my PhD research project which started in September 2019. The main argument of my thesis is that paratext can significantly impact how readers perceive the creative and technical value of NaNoGenMo generated novels, as well as perceptions of enjoyment, interest, and understanding. My findings also identify which paratextual elements are judged to be the most impactful for enjoying, understanding, and interpreting these works. More broadly, my research shows that participant-based reader studies contribute insights into actual reader reception and interpretation of generative literature, especially where perceptions of literary value appear to be virtually unimpacted by paratext.

My professional motivation for the project is a gap in research which understands how computer generated novels are read by a wider audience outside of academia and outside of the creative communities which produce them. At present, generated novels are rather loosely defined but definitions can include being a creative text generation work of 50,000 words or more. Works are typically comprised of the generated text, the code used to generate it, the author's notes and creative intentions, the input data used by the code, and so forth. While generated novels and other types of generative literature have been researched as creative forms to a limited extent, little has been investigated about groups of readers and their reading experience, especially beyond an implied reader perspective. This research project contributes to filling this gap.

My personal motivation for the project's focus on generated novels and readers is my interest in electronic literature which began in the final year of my Bachelor of English Literature and Linguistics in 2017. Later, I become more focused on a subtype or related genre - generative literature. This coincided with a technical interest and study of natural language processing technology (an interdisciplinary field where linguistics and Artificial Intelligence meet). These two interests naturally lead me to find the National Novel Generation Month[1] (NaNoGenMo) challenge, a space where creators playfully experiment with data and 'Lo-Fi' and state-of-the-art language processing methods to create generated novels. I fell in love with generated novels and became motivated to contribute to the form's development.

Generated novels and other forms of generative literature such as creative Twitterbots and generated poems have proliferated and are part of web culture. More broadly, the presence of generative qualities and algorithmic processes are becoming more apparent in ordinary life and increasingly present in digital tools, (creative) media, and the sociotechnical fabric more

[1] https://nanogenmo.github.io/

generally. Thus, it is becoming increasingly relevant to research real readers' understanding and reception of generated text both from the perspective of studying emerging creative digital media forms, but also relevant to better understanding the sociotechnical interplay between reader interpretation and beliefs about generated text technology. The vast majority of this research project is focused on the former; studying generated novels in relation to reader interpretation and reception. Some of this research project's outcomes however make valuable contributions to sociotechnical research around user-facing artificial intelligence processes and tools.

The impact of paratext on readers' interpretation and reception of media works has long been accepted and theorized in literary studies. Paratext refers to the framing, supplementary, and periphery elements which can accompany a work. With regards to web published work, (which many generated novels are) paratext can refer to the elements which frame, link, and accompany a work. The concept of paratext is frequently drawn on in literary studies and media studies to discuss works and new media forms, where it is also debated and reformulations are proposed. However, within this literary and media space there is little in the way of empirical research which tests assumptions about how paratext functions to impact different readers' interpretation and reception of a work.

This research project uses the concept of paratext as a means of studying generated novels and reader reception and includes a critical review and refinement of paratextual theory. Reception is decomposed into different value dimensions such as literary, creative, and technical value, as well as understanding, interest, and enjoyment. The research project further investigates how reader's computer coding skills and literary reading experience might interplay with the impact of the paratextual elements, as well as how the technical and cultural characteristics of a generated novel might shape interpretation and reception.

The research project is constructivist in epistemological orientation and follows an explanatory sequential mixed-method design where each component of the research is shaped by the results of the component preceding it. Two studies are conducted and analyzed. The first study is an online reading experiment where the generated novels which are used as study items are presented in one of three different paratextual conditions. Participants' responses to the generated novels are recorded through a survey and through written reviews where participants give their personal opinions about the works. The second study is an in-person reading group discussion workshop where a semi-structured group interview is run to discuss readers' impressions of the same generated novels. But this time, each generated novel is presented and read in printed paperback book form – a physical paratext.

## 1.1. Research questions

The general research question asks: in what ways does paratext influence potential readers' interpretation and reception of a generated novel? This is broken down into three specific research questions which direct the main lines along which the research data from the two studies is analyzed:

Which paratextual elements have a central role in influencing the understanding and reception of a project?

Which reader skills and experience affect the influence of these paratextual elements?

Which technical and cultural characteristics of a generated novel affect the influence of these paratextual elements?

### 1.1.1. Research question motivations

Throughout the research project I specifically focus on generated novels which have been entered into the annual National Novel Generation Month (NaNoGenMo) challenge, because it is the largest documented source of generated novel projects in English. Within this space, I focus on identifying and understanding the impact and interplay of the paratextual elements which are used and which emerge with the creation and online publishing of NaNoGenMo projects. This research project thus aims to answer questions about which NaNoGenMo paratextual elements have a central role in influencing the understanding and reception of a project, if this changes depending on potential readers' skills and previous experience, and depending on the project's technical and cultural characteristics. These characteristics include the cultural value of classic novels whose text might be used as input in the creation of a generated novel. They can also include the perceived status of a software tool or method which might be state-of-the-art or hyped in technology news media.

Generated novels can be approached through both a technical lens and a literary or creative lens (or indeed an amalgamation of both). Therefore readers' literary reading experience and computer coding skills were assumed to play a factor in understanding and interpreting generated novels, and also assumed to play a factor in which paratextual elements might have a greater impact on shaping reader interpretation. For example, it may be the case that programmers value a generated novel's code more than people don't have coding skills. Therefore both studies recruited a mix of participants who had coding skills or literary reading experience. The influence of reader skills is investigated in *Chapter 4 Reading experiment and survey study*, and expanded upon in *Chapter 5 Workshop study* where the results suggest that actually, a reader's ability to develop their own creative framing with which to interpret the

generated text is an impactful reader skill that enables an enjoyable reading experience and therefore positive reception. To the best of my knowledge, this insight has not been made in the research literature and is therefore a novel contribution about generated novel and generative literature reading strategies.

## 1.2. The readers

This research project studies how computer generated novels are read by potential new readers. By new, I mean readers who may be encountering a generated novel or generative literature for the first time. By potential, I mean to capture the fact that the readers are study participants who have been invited to participate in reading studies, rather than somebody who has chosen to read a generated novel of their own accord. I have chosen to focus the research project on new readers because collecting data about initial reactions and first impressions is a valuable contribution because existing research about reading and interpreting generated novels is focused on the implied reader or a single expert academic reader. Therefore, by focusing on potential new readers there is a chance to understand how they might approach the unfamiliar form without explicit expert direction, and a chance to understand how paratext works to impact this. A second opportunity afforded by focusing on potential new readers is that it is far more practical to recruit a large sample of study participants to whom the generated novel or generative literature is an unencountered or unfamiliar form. Because quantitative methods can require larger sample sizes than qualitative methods, the focus on potential new readers is an excellent fit and opportunity for statistical study designs which can compare and measure differences between reader groups.

## 1.3. The generated novel

To the best of my knowledge, the generated novel is yet to be defined in detail. Most scholarship defaults to taking the NaNoGenMo entry rules as a defacto definition, or simply describes the form as a long-form creative generated text. But I find this to be too cursory. For example, Van Stegeren and Theune (2019) conduct a survey of narrative generation methods used in NaNoGenMo novel generation projects, but despite using the term 'generated novel' the authors don't offer a detailed definition. Similarly, NaNoGenMo metadata and aggregate statistics online catalogue NaNoGenMoCat[2] does not offer a definition. To the best of my knowledge, my definition of a generated novel (below) is a novel contribution in that it extends beyond the NaNoGenMo entry or a description of creative generated text more generally.

[2]http://nngm.botstudies.org/s/n/page/home

As I have previously detailed in my master's dissertation *Creative Text Generation: A NaNoGenMo 2018 Study* (Tkacz, 2019), the generated novel is an emerging form of generative literature. It can typically be thought of as a part of web-based digital culture along with other forms such as twitterbots, travesty generators, the creative use of predictive keyboards, poetry, and humorous or bizarre neural network generated content. This is not only because generated novels are frequently published, accessed, and shared online, but because of the input data many of them use as part of the text generation process. For example, this data can originate from open-source and proprietary web sources (such as text from social media platforms) or made available on the web as part of community, volunteer, or public domain libraries and archives (most notably Project Gutenberg[3]). Of course, although it's an overwhelmingly common component of a generated novel, input data isn't strictly necessary to generate text. There are only two necessary components of a generated novel: the code that an author writes or adapts, and the generated text that is generated by the code. There is an except to this which is detailed at the end of this subsection.

While there is currently no established definition for what a generated novel is, the National Novel Generation Month (NaNoGenMo) challenge's short entry requirements are frequently used as a description because a large portion of generated novels in English originate from the online challenge. The NaNoGenMo challenge only asks that entrants write and share computer code which generates a work of 50,000 words or more of generated text. Indeed, this word count is a helpful indication of how large the generated text component of a generated novel should because it's considered to represent the size of a typical print novel (or to be more exact, the size of a National Novel Writing Month entry[4]). Often new readers might implicitly assume that the generated text is presented more or less as it has been output from the code, and that editing has not taken place nor handwritten sections been added. However, some generated novel are edited for spelling and punctuation corrections, and handwritten text templates or framing introductions which have been hard-coded and included in the generated output is very common. It should also generally be assumed that authors have curated their work by cherry-picking the best of several generated text versions.

In terms of presentation, many generated novels are presented digitally but have paratextual elements which are recognized from print books where the title, author name, a table of contents, chapter titles, and numbered pages are included to contain the generated text. Book covers may also be included especially if the generated novel is presented as a PDF digital file. HTML and plain text formats are also used to present the final versions of generated novels. As

[3] https://www.gutenberg.org/
[4] https://nanowrimo.org/

will be expanded on in the second half of *Chapter 2 Background and Literature Review*, when a generate novel is presented online in a digital form it's typical for the code, the input data, and the author's creative or technical intentions behind the project to also be present or accessible. A handful of generated novels which originated as NaNoGenMo entries or which were created outside of the annual challenge have been printed as physical books, although these do not always follow NaNoGenMo requirements or conventions. For example, the code is not typically included in the print book, and it might not be publicly available online for readers to find if it was created outside of NaNoGenMo.

While an adjacent generative literature form, computer generated poetry, can be expected to feature rhyming, meter, simile, metaphor, and so forth, this is not expected in a generated novel. Indeed, while Van Stegeren and Theune (2019) conduct a survey which in part judges NaNoGenMo novels on their level of coherent narrative, they explain that a consistent narrative or character development (or even clearly defined characters for that matter) on the level of regular novels is not expected, because language generation technology is currently not advanced enough. This is still the case in 2024. However, I add that literary devices such as imagery, mood, setting, diction, style, and genre can certainly be expected in a generated novel, as well as imitative forms such as parody and pastiche (Tkacz, 2019). But it is important to stress that works do not have to conform to any of the aforementioned characteristics or cases in order to be considered a generated novel. Indeed, to reiterate NaNoGenMo's requirements, literally 'anything goes' as long as 50,000 words are reached and the code is shared. Eschewing a hard requirement for sharing the code, I build on this by proposing that the most useful guideline for defining a generated novel is that it is a work of long-form creative text generation that falls within a range of about 50,000 words. This word count is relevant not necessarily because of NaNoGenMo's influence on the form, but because producing a coherent creative generated text of this size (as opposed to a functional text which needs to prioritize fact, such as textbook) is currently unachievable in the computer science and Natural Language Generation domains. Indeed, this encourages creative experimentation in the generated novel form space because no author is expecting to be able to produce a generated text that is remotely comparable in coherence and narrative quality to regular novels which are written in the ordinary way; so this leaves authors freer to pursue other aims. For example, generating a piece with a distinct mood or imagery. This is expanded upon further in Tkacz (2019), and has also been noted by Van Stegeren and Theune who agree that "What makes NaNoGenMo extra interesting is that it focuses on the generation of texts with a much longer length than addressed in most scientific research" (2019, p. 65). I would add that this partly what makes generated novels, not just NaNoGenMo, extra interesting.

Returning to my promised exception:

A few years ago I would have confidently written that there are only 2 absolutely necessary components of a generated novel - the code that an author writes, and the generated text that is generated by the code. This has perhaps become complicated in recent years with the availability of user-facing consumer Large Language Model (LLM) text generation tools where users' prompts (rather than computer code) are written in an attempt to induce a model to eventually produce the intended text output. However, it seems to have become common practice in the general area of generative tool use for users to report and detail the prompts and the prompt writing strategies which they use to arrive at the intended version of a generated piece. So in cases where 'vanilla' LLM consumer tools might be used to create a generated novel, even with the absence of model fine-tuning as directed by authorial creative intention, I judge LLM prompts to function as a substitute for code (for better or for worse) not only because they can be used to understand how a particular piece was produced, but the prompt choices and strategies may also be discussed and analyzed through the lens of authorial intention and design.

## 1.4. Novel contributions to the field

### 1.4.1. Contributions to theory

In the beginning of the research project and this dissertation document, I review and critically discuss literature which introduces and challenges the paratextual conceptualization. In the middle part of the project where the studies are conducted I reflect on how ideas about paratext and how it functions fit with the study data, and how other influences appear to be at play. Towards the end of the project and dissertation document I conclude that paratext is a useful concept for studying emerging media forms when it is kept theoretically simple. When conceptualized as a generalizable model, I propose that paratext can aid analysis by conceptually structuring and easily operationalizing specific works or forms into paratextual pieces. By segmenting a work or form into pieces, this allows for each piece's impact on reception to be critically considered or empirically measured and unpacked using quantitative and qualitative methods. This is a novel contribution to theory because rather than advocating for more complexity as other paratextual theory reformulations suggest, I argue that the concept of paratext is most useful to researchers when it is approached as a lean, generalizable conceptual tool which enables it to be used to structure and study emerging media forms. This expanded upon in section 6.*2.4. My minimalist paratextual conceptualization* where I also give examples from previous research.

### 1.4.2. Contributions to methodology

To the best of my knowledge this research project is the first to conduct actual reader multi-reader studies of generated novels more generally, and the first to use mixed methods to empirically test and measure the impact of paratext on the interpretation and reception of generative literature works in the literary studies and media studies disciplinary space. This therefore has a methodological contribution aspect to it.

### 1.4.3. Other contributions

Additional novel research contributions yielded from this project include the collected data from both studies, insights about ergodic reading and creative framing interpretation strategies which are seen with an enjoyable reading experience and positive reception, a conceptual model which describes an abstracted processes of reading and interpreting a generated novel based on links with paratextual elements and other influences, and a research-based recommendation about which paratextual pieces are the most impactful to develop and include in print books versions of generated novels. I also offer and describe the term folk theories of AI and contextualize it with related concepts. The generated novels used as study items which I edited and arranged to be printed in paperback book form are also a practice-based contribution.

## 1.5. Research project scope

While carrying out the research activities, I edited and arranged for the printing of paperback versions of the two generated novels which were used as study items. This brings a practice-based element to the research project in *Chapter 5 Workshop Study*. Although the process itself could have been expanded into a chapter of its own, I did not do this because it falls outside the current reader-focused scope of this research.

This research project was designed and the data collected before the public launch of OpenAI's ChatGPT[5] and the subsequent seemingly ever-growing public exposure and awareness of user-facing generative text tools, and the general hype associated with generative Artificial Intelligence. Therefore, in addition to contributing to filling the research gap on generated novels and reader interpretation and reception, and contributing to research on paratext, the data collected as part of this project can also function as metaphorical litmus test which indicates some of the reasoning and range of opinions that the study participants reported about

[5] https://chatgpt.com/

algorithmic works and technologies. Crucially, this 'litmus test' indicates opinions expressed before the greater public sensationalization of generative Artificial Intelligence (which could indeed warrant a paratextual study of its own). However, I don't foreground this in the research project because it falls outside my current scope. Nevertheless, interested readers can trace this in *Chapter 4 Reading Experiment and Survey Study* and *Chapter 5 Workshop Study* where it appears in qualitative analysis codes, concepts, and themes.

Finally, in this research project I alternate between naming paratext a theory, concept, conceptualization, and an idea. I have not settled on which of these is the best term to use because researching the nuanced meaning behind terminology and terminological development is outside the scope of this project. I have instead chosen to use my research time to study paratext in action and to better understand how the concept might apply to real-world data and the research questions.

## 1.6. The structure of the dissertation

*Chapter 2 Background and literature review* begins with a review of the concept of paratext, which includes a critical discussion of how it is approached and challenged in new media scholarship. Extensions and reworkings of Gérard Genette's seminal paratextual conceptualization are also reviewed here. Links to my own refining of paratextual theory are made in relation to the scholarship, and they are fully developed and expanded on later in *Chapter 5 Workshop study* as a result of both studies' analyses. The second half of *Chapter 2* resumes from section *1.3 The generated novel* to continue introducing the form and the NaNoGenMo web-based community and platform, and details the two NaNoGenMo generated novel projects that are used as study items in this research project. Drawing from the concept of the paratext as outlined earlier in *Chapter 2*, it concludes with operationalizing and discussing the pieces that the two generated novels are comprised of.

Where *Chapter 2* outlines the concept of paratext and how it relates to the generated novel, *Chapter 3 Research Framework* sets out the research design rationale and plan. This begins with discussing how the paratextual conceptualization theoretically underpins the research questions, and then advances to outline an overview of the research steps which make up this explanatory sequential mixed-methods research project. Finally, reader theories are briefly outlined in order motivated the reader-focused research design choices, and the research's constructivist orientation is thoughtfully motivated.

After reviewing related studies, *Chapter 4 Reading experiment and survey* study details an online reading experiment where the generated novels are presented in one of three different paratextual conditions. Here, the operationalization of a generated novel's elements into

paratext pieces as detailed in section *6.2.4. My minimalist paratextual conceptualization* is demonstrated. Participant responses are recorded through a survey and written reviews where participants give their personal opinions about the work they read. The analysis tests for statistically significant differences between paratextual conditions and reader groups across six value dimensions. The results suggest that, while the presence of the paratextual elements in the generated novels have a measurable impact on readers' perceptions across virtually all the value dimensions, the presence of the generated text itself doesn't significantly impact reader valuations. Further, the results show that literary value is barely impacted by the presence of the paratext, which is unexpected. This is further investigated quantitatively through participants' familiarity with the classic novels that have a hypotextual relationship with the generated novels. The perceived ranked importance of each paratextual piece by participants is also analyzed. The second part of the chapter focuses on the written reviews and uses Qualitative Content Analysis to further investigate the unexpected literary value result, Here, I develop categories to compare differences in the review content between paratextual conditions and reader groups. I conclude analysis by advancing to identify a pattern and developing theme which describes how some reviews reject literary value and relocated it to technical value. I answer the research questions based on this first study.

*Chapter 5 Workshop study* presents the second study of this research project which is comprised of two components. The first is a workshop inspired by reading group practices. For this I designed and arranged the printing of physical paperback books which contain the generated text of the two NaNoGenMo projects. These were independently read and annotated by the workshop participants in their own time, who then met together in the workshop to discuss their reading and impressions. I ran this workshop component as a semi-structured group interview. For the second component, I transcribed the workshop audio recording and performed Deductive Qualitative Analysis on the data to develop themes and a conceptual model which describes the different potential ways that a generated novel can be interpreted in relation to paratextual pieces. I answer the research questions based on this second study.

*Chapter 6 Discussion* integrates the results from both studies and discusses the answers to the research questions. I take a firm stance to answer the primary research question by drawing from the studies' data. Here, I develop a critical discussion based on the unexpected results relating to literary value in my studies by synthesizing relevant literary theory from reader response theory and narratological theory of postmodern narratives in relation to generated novels, and I arrive at a theoretical explanation for the unexpected results. In this way, my arguments are contextualized within relevant theory areas and with my data. Finally, I build on my data and reflect on my research process to arrive at and propose my own minimalist paratextual conceptualization section *6.2.4. My minimalist paratextual conceptualization*. The

chapter reengaging with the critical discussion of paratext that I developed in *Chapter 2*, and reflects on Genette's paratextual conceptualization based on my data. I conclude by discussing the implications of my data for using the paratextual model.

Finally, *Chapter 7 Conclusion* draws the PhD research project to a close by summarizing the ground covered and the data analysis and insights developed from *Chapter 2 Background and literature review* to *Chapter 5 Workshop study*. It begins with a brief summary of the research project's analysis narrative and reminds of the research design steps which underpin the analysis, and my research contributions are taken stock of. These contributions include my minimalist paratextual conceptualization. Finally, the research project's limitations are explained and I chart my possible future research plans and offer my thoughts on future research more generally.

# Chapter 2 Background and Literature Review

This chapter begins with a literature review of the concept of paratext, which includes a critical discussion of how it's approached and challenged in new media scholarship. Extensions and reworkings of Gérard Genette's seminal paratextual conceptualization are also reviewed here. Links to my own refining of paratextual theory are made in relation to the scholarship, and they are fully developed and expanded on in section *6.2.4. My minimalist paratextual conceptualization*. The second half of this chapter continues from section *1.3 The generated novel* to continue introducing the form and the NaNoGenMo web-based community and platform, and details the two NaNoGenMo generated novel projects that are used as study items in this research project. Drawing from the concept of the paratext as outlined earlier in this chapter, the chapter concludes with operationalizing and discussing the pieces that the two generated novels are comprised of.

## 2.1. Paratext

### 2.1.1. Genette's concept of paratext

Structural literary theorist Gérard Genette's foundational theory of transtextuality describes how a text is able to transcend its own limits and bear relationships with other texts and elements, where "the textual transcendence of the text…[means] all that sets the text in a relationship, whether obvious or concealed, with other texts" (1997a, p. 1). In *Palimpsests: Literature in the Second Degree*, Genette coins or borrows concepts from previous scholarship to build a foundational "…general poetics[6] of transtextuality" (p. xviii), which is presented as a schema of five transtexstual relationships: intertextuality, paratextuality, metatextuality, hypertextuality (which includes hypotexts), and architextuality (pp. 1-5). *Table A1: Transtextual Relationship Terms and Their Definitions* seen in *Appendix A* shows their respective meanings. For Genette, transtextuality accounts for powerful relationships into and out of the text, as well as its surroundings. These have strong influences on readers, and so transtextuality plays a role in enabling and directing the text's consumption and the interpretations drawn from it. Transtextuality can describe any medium, although Genette illustrates its relationships by

[6] Genette explains that transtextuality is the subject of poetics because the text is considered to be in relation to others. This is contrasted with criticism which considers the text in its singularity (1997a, p. 1)

drawing upon pre-digital print media and restricts examples to literary works. Although Genette makes clear distinctions between the transtextual relationships, these nuances are generally not made in more current Media Studies research.

Genette focuses on paratext in *Paratexts: Thresholds of Interpretation* (1997b). Genette's paratext-as-threshold metaphor is used to describe a means of entry into a text, where the text's centrality situates the paratext in a liminal, off-center position as a textual vestibule that the world must access the text through. "It is an 'undefined zone'…without any hard and fast boundary…as Philippe Lejeune put it, 'a fringe of the printed text which in reality controls one's whole reading of the text'" (p. 2). For Genette, a paratext must be intentional and it must be directed by the author or publisher. They see the paratext as the legitimate conveyor of authorial commentary, which makes this metaphoric threshold a transitional zone and a transactional zone as well (p. 2). In these zones, paratext is:

> "…a privileged place of a pragmatics and a strategy, of an influence on the public, an influence that - whether well or poorly understood and achieved - is at the service of a better reception for the text and a more pertinent reading of it (more pertinent, of course, in the eyes of the author and his allies)" (p. 2).

Genette therefore argues that paratext plays an important role in influencing readers' interpretation and reception of a work. And, as media and cultural studies researcher Martin Barker points out, "Genette is utterly devoted to notions of authorial intent" (2017, p. 240).

Genette generally restricts paratexts to mean the elements which the authors and producers of print novels are thought to create in order to manage and direct the reader towards the intended interpretations of a work: "something is not a paratext unless the author or one of his associates accepts responsibility for it, although the degree of responsibility may vary" (p. 9). As will be discussed below, this has proven to be problematic because of who is unaccounted for in meaning making, and for oversimplifying the production and publishing process.

### 2.1.2. Criticism of paratext

Genette's ideas about transtextuality, especially the later expansion of paratext, have remained highly influential and are used outside of literary criticism. Of the five transtextual relationships, it is paratext which has enjoyed the most attention from Media Studies and related fields in recent years. However, critical attention has also highlighted how problematic Genette's paratextual conceptualization is perceived to be, or, as I understand through Barker's theoretical revisitation of paratext, how misunderstood the concept is because it has been stretched far beyond Genette's original intended purpose which was situated im structuralist "…literary narratology" (2017, p, 239). And, in fact, that Genette has shown a "…lack of interest

in *actual* readers", and not much more in reception and interpretive communities either (p. 239). I note that in contrast, this research project has a high interest in actual readers, but this ultimately did not impede how I ultimately used and refined Genette's paratextual conceptualization in section 5.9 *Refining the paratextual conceptualization*. The rest of this section will review and discuss the theoretical paratextual framework developments, complications, and reformulations in Media Studies.

Many mediums can demonstrate how Genette's print-focused definition of paratext and structuralist ideas about how paratext should behave don't accurately reflect how it is observed to function in media. Demonstrative examples of this can certainly be seen in media forms such as film and television, video games, and electronic literature, and these are discussed in the following sections.

### 2.1.3. Paratext in new media scholarship

#### 2.1.3.1. Print publishing

Before discussing media forms, it is worth noting that Genette's paratextuality is acknowledged as being brittle even in their area of focus: print books. Where Genette relies on notions of coherent authorship and intent by insisting that a paratext is legitimate only when it is created by key producers, they expose their lack of knowledge about book production. Brookey and Gray suggest that a lack of production studies knowledge must have led to Genette's overestimation of the degree of agency that an author has when publishing a book with a press, which they argue is a much messier process than Genette accounts for (2017, pp. 102-103). I note that Brookey and Gray's observation also suggests that defining paratext as something which can only be knowingly and officially or unofficially created and condoned by key producers is too constrained to be able to apply to screen and new media (and arguably print novels too), where spin-offs, mash-ups, fan fiction, and community exchange also impress upon consumers.

Interestingly, within the NaNoGenMo context[7], generated novel authors actually do have virtually full control over the production and web publishing of their works and accompanying materials, which is ostensibly in line with Genette's assumptions. However, control is also increased for readers and users who have the ability to create paratexts in the form of GitHub comments, the ability to copy authors' repositories and reuse code or input data, and create

[7] This is of course a different case for generated novels which are published in print with a publisher.

their own critical paratexts in the form of press articles and social media sharing and commentary.

#### 2.1.3.2. Franchising and transmedia storytelling

In an interview between Brookey and Gray, media and cultural studies researcher Jonathan Gray argues that because Genette sees the paratext as always outside of the text, the language often used to talk about paratext takes the form of "…paratext versus text, but in fact, it cannot be versus the text because it is part of the text" (Brookey and Gray, 2017, p. 102). Gray's widely cited Show Sold Separately: Promos, Spoilers, and Other Media Paratexts (2010) is celebrated for its critical attention to the works which are created and released with film and television. Gray argues that Genette's clear distinction between a central text and its supporting elements is difficult to make in film and television media. For example, in the case of typical Disney films with huge campaigns, an array of merchandise, marketing deals with fast food companies, and advertising on popular children's television programs would have loudly and actively preceded the actual film's box office release. When such a successful campaign creates a curated and compelling image of what the film represents, to the point where the promotional elements which feed into the film and prime its consumption become part of the experience, the elements are part of the text (2010, p. 38). Although I note that Gray has in fact made clear distinctions between the film and advertising and merchandising elements in their example. In this case for Gray, then, Genette's paratext is not able to satisfyingly account for the form and the textual potency which may be present in the promotional elements; it cannot explain how the elements used to create hype contribute to meaning making in and out of the text.

Gray (2010)'s discussion of film paratexts in practice seeks to demonstrate an impracticality of applying Genette's clear-cut distinction between a text and its paratext. Indeed, when designing this research project I assumed that distinguishing between a central text and its supporting elements was perhaps even more challenging in the case of NaNoGenMo novels. Here, pieces such as the code used to generate the text, and the pieces which document the authorial intentions directly feed into the reading of the generated text. Because these pieces appear to be so crucial to reading and understanding the generated text, it seemed difficult to argue that there is one central text in a generated novel project without first collecting evidence from a reader study. As discussed with the studies' results, though, the study participants did not seem to encounter this perceived complication.

#### 2.1.3.3. Digital games

In their doctoral dissertation, game culture researcher Jan Švelch analyzes video game trailers and their reception by viewers (2017). Trailers would be considered as merely supporting elements under Genette, but Švelch's study challenges this by showing how game trailers can hold a textual status where they are regarded as a text in their own right. Evidence is shown in the way that the trailers are elevated through viewers' comments: "…direct acknowledgement of trailer's textuality can be located in contributions that critically evaluate the viewing experience" (p. 134). Švelch's participant "…observations suggest that a video game trailer can be enjoyed independently from its paratextual connection to a video game. The…comments also highlight the fact that the high production values of cinematic trailers for MMOs [Massively Multiplayer online game] can make watching worthwhile even if a viewer does not intend to play the game at all" (p. 135). Švelch argues the need to reconfigure Genette's paratext into a revamped framework which can support video game media and fully realize a more nuanced, richer textuality in paratexts.

I note that Švelch's reception study of game trailers demonstrates why assumptions about emerging media forms cannot be made. Švelch's careful study of viewers' comments evidences that, rather than trailers only being considered in their traditional capacity as merely supporting elements, trailers can in fact be engaging and potentially independent texts. Similarly, assumptions about generated novels and their elements' status as independent texts or supporting pieces cannot be made with confidence without first studying reader reception. Reflecting on my thoughts about paratextual theory that I develop over the course of studying my results, Švelch's approach to studying and validating assumptions about traditionally paratextual pieces and their impact on an audience aligns with my views in principle. Although in contrast to Švelch my conclusions advocate for more simplicity rather than complexity in paratextual theory.

### 2.1.4. Paratext extensions and reworkings

In response to the gaps in Genette's basic, unaltered theory of paratext, researchers have extensively adapted it so that it is malleable enough to accommodate critical inquiry for their specific area of media study, plus to account for audiences, fans, the web, and different creative industries. These extensions and reworkings of Genette's paratextual conceptualization are discussed in this section. After analyzing the results from both of my studies and after being informed by my research process, I return to Genette's paratextual conceptualization and respond to it with my theoretical refinement in section *6.2.4. My minimalist paratextual conceptualization,* as based on evidence from the ways in which paratext

influences potential readers' interpretation and reception of NaNoGenMo generated novel projects.

Paratextuality is brought into the domains of film and television where Gray (2010) stretches its definition to include a much wider possible set of elements which can be considered paratexts. These include merchandise, fan fiction, promotional pages, and web-based alternate reality games. Gray uses a multitude of examples to show that paratextual elements should not be considered as separate from the text - they are part of it. And furthermore, that what is valued as a central work is subjective and expected to change depending on who values it. This subjectivity of value position greatly alters the foundation of Genette's paratext. Gray clarifies their argument in Brookey and Gray (2017) by stressing that paratexts are an integral part of the text "…as a social and cultural unit", which means that "…the hierarchies of value or importance are never predetermined: to different people and different communities, at different times, the hierarchy of value will be different. Indeed, sometimes a paratext that to me is central might be one that you are not even aware of" (p. 102). While Gray's subjective value of paratexts and their centrality certainly seems to make sense intuitively, I was not able to evidence this with statistically significant differences in value judgements between survey study participants who reported having coding skills, and those who reported having literary reading experience.

Conversely to Brookey and Gray, Švelch writes from the perspective of game studies and criticizes Gray (2010) and other research for forming "…overly inclusive and vaguely phrased extended versions of paratextuality" (2017, p. 5). Švelch reserves the term paratext as a classification which only refers to practices, textual forms, and categories; thus, only general established practices and not individual works (pp. 67-68). For example, paratext can refer to all prefaces, but not a specific individual preface. Paratextuality is then proposed as "…an aspect of a text that refers to the socio-historical reality and potentially comments on a text's position and role within this reality"; so 'paratextual' denotes a measure of a text's reference to sociohistorical reality and can be used when comparing texts. However, Švelch is careful to stress that paratextuality is still intrinsically tied to other types of transtextuality and criticizes other theory updates which do not take this into account. This is because textual transcendence should be seen as a quality of all texts, and therefore overlaps and interactions between transtextualities should be recognized (p. 69). This level of complexity is in sharp contrast to my proposed paratextual conceptualization where I argue that the theory is most useful when used as a simple, generalizable model. This argument is fully fleshed-out in the penultimate chapter of this dissertation.

Literature researcher Yra van Dijk focuses on electronic literature to argue that, like the literary work that it helps to shape, the paratext itself must be analyzed and interpreted. They expand Genette's concept of paratext "…to take in the 'texts' that cluster around a digital text and

become part of it, even if there is no authorial consent" (2014, p. 24). They stress that this is all the more relevant in the web context because paratext can at the same time merge with the text, while simultaneously working to expand the online context beyond author and publisher consent. Similar to the researchers discussed above, Van Dijk rejects Genette's hierarchical separation of text and its supporting elements. Instead, they argue for "…us to abandon standard binary oppositions such as between text and paratext, and to assume that the paratext functions to give rise to the new work" (van Dijk, 2014, p. 41).

Van Dijk's paper critically narrates the process of searching for a work of electronic literature online, and identifies the web-based paratexts which are encountered along the way in order to absorb them into Van Dijk's updated theory about what can be considered paratextual elements, and how they may be valued. These include web user tools and computer file systems, author's homepages, and the Electronic Literature Organization's website. For example, Van Dijk illustrates that readers of electronic literature might begin by searching for works through a search engine, but its results are already unintentionally serving to metatextuality frame the work through the queried keywords. I note that the same could be said for generated novels.

Van Dijk shows that the concept of paratext has already been adapted in some electronic literature projects since at least 2010. In their analysis of a multimodal interactive fiction work, Stewart (2010) discusses on-site and off-site web-based paratexts (p. 64), as well as in-file and out-file paratexts (p. 68). Van Dijk clarifies that these types of paratext describe the location of supporting electronic files and their proximity to the main text (2014, p. 27) in software. I consider NaNoGenMo projects to be works of electronic literature, and I note that Stewart (2010)'s terms for indicating the proximity of web and software paratexts are suited to describing the elements which comprise a generated novel. For example, press articles about NaNoGenMo projects can be understood as off-(the GitHub) site paratexts. Similarly, the generated text would be an out-file paratext of the code if the code is considered to be the main and central text in a project. This site and directory-based concept of paratextual proximity was realized to some extent in my conceptual model seen in in *Figure 18: Conceptual Model Showing Different Potential Ways to Interpret A Printed Generated Novel* in section *5.8 Conceptual model*. Here, I use *Project Paratext* and *Wider Paratext* to represent proximity between files on the web.

### 2.1.5. Definition of paratext used in this dissertation

Several times in Van Dijk (2014), paratext is used as a hypernym for different transtextual elements, and I note that this is not unusual to do in Media Studies research. For example, although the paper is titled with the promise of "…Paratext in Digital Literature", the term

metatext is rolled into paratext, and it is not always clear if intertextuality has been separated from paratextuality. This is not a critique of Van Dijk's use of terminology because it is not clear if it would be productive to painstakingly reassign all of Genette's transtextual relationships outside of the literary print novel. Treating the term paratext as a hypernym for transtextual elements is reasonable, especially considering that paratext is now a much more frequently used term than other transtextual terms. I argue that using paratext as a hypernym in its more flexible, expanded sense beyond Genette is not only reasonable, but productive. This helped me to avoid getting mired in what would have been an extra, unnecessary phase of taxonomic labelling during the data collection phase of my research.

I note that Švelch (2017) criticizes paratextual framework updates which do not account for important transtextual relationships; my own use of paratext as a hypernym for transtextual terms will take care not to overwrite important transtextual qualities or relationships that texts or elements perform, wherever this is relevant to the discussion at hand. I will therefore use paratexts as a hypernym as well as name individual transtextual terms where appropriate. This means that my research questions about how paratext influences people's reading of generated novels is using paratext as a hypernym.

As explained in *Chapter 1 Introducing the research*, in this research project I alternate between naming paratext a theory, concept, conceptualization, and an idea. I do not settle on one of these terms because terminological development is outside the scope of this project. Instead, I use my research time to study paratext in action and to better understand how the concept might apply to real-world data and the research questions.

In terms of research design, after completing the literature review I made the decision not to define what my own definition of paratext was until I had collected and analyzed empirical evidence from my study data. As Barker puts it, "…if we do want to pinch Genette's term, there is work to be done to ensure that we have properly shed the theoretical apparatus that has prompted his approach and built an adequate one of our own!" (2017, p. 240). While I did indeed want to pinch Genette's term and have performed a considerable amount of empirical research work to do so, based on the results and my experience conducting the two studies I did not find it necessary to build my own theoretical apparatus, as Švelch (2017) has done for example. *Chapter 4 Reading experiment and survey study* and *Chapter 5 Workshop Study* is where the development towards *6.2.4. My minimalist paratextual conceptualization* can be traced through the research design and process, and the results.

## 2.2. NaNoGenMo community and platform

NaNoGenMo is an annual challenge where the aim is to computer generate a novel of 50,000 words or more, and to share the code which is used to generate it.The challenge is conducted online and administered by volunteers who also frequently take part. For the purpose of my studies, I define a generated novel being a completed NaNoGenMo entry. NaNoGenMo's single rule stipulates that at least one generated work of 50,000 words or more and its source code must be shared at the end of the challenge in order to be considered complete. This rule is explained on each annual challenge's page[8].

The challenge began as an idea tweeted in 2013 by the challenge's founder, internet artist Darius Kazemi. It has been recurring annually since November 2013 and is the largest organized activity for the creation of generated novels in English. Any participant can join the challenge via the web-based version control and code collaboration platform GitHub. GitHub lets users create free, public code repositories where a computer programming project's files and file version history can be maintained. The nature of the platform allows NaNoGenMo to also function as an archive for a large body of generated novels, since every NaNoGenMo entry from 2013 to the present that was publicly posted on GitHub is in theory accessible to anyone with a web browser. For example, the 'National Novel Generation Month' repository[9] links to the 2019 challenge[10] repository. Not all NaNoGenMo works are (only) text-based or resemble a novel; generated poetry and art books are also created. There is no cost or formal registration process needed to participate, no formal categories (although several styles of generated novels have arguably emerged over the years), and there are no prizes or 'best entry' positions. There are no human or programming language restrictions, although the majority of entries are in English.

Using GitHub is not necessary to participate, but it is the platform where the majority of projects are made publicly available, shared, commented on, and marked as completed. For example, the 2019 challenge uses GitHub's *Issues* functionality to link to the 140 NaNoGenMo submissions which were started for that year[11]. Here, the finished entries are accessed by users/readers. NaNoGenMo projects can vary greatly in input, output, and in the approaches

[8] https://github.com/NaNoGenMo/2019#the-rules

[9] https://github.com/NaNoGenMo

[10] https://github.com/NaNoGenMo/2019

[11] https://github.com/NaNoGenMo/2019/issues

and digital tools used. The majority of generated novels are typically created by taking input data as text or data files, and then computationally extracting the desired data from them. The data is then algorithmically processed using different software tools and methods in order to generate many versions of output text, which I refer to as the generated text. A plain text output on its own can be considered a generated novel, and it can contain text characters, a title, a table of contents, chapter titles, as well as pages and page numbers. Input text types include classic novels, non-fiction books, web-based texts retrieved from an API (such as social media APIs), corpora, datasets, images, and so on.

A NaNoGenMo participant may begin a project by describing an idea for a 50,000 word novel. This is done by creating an issue page in the NaNoGenMo repository of the current year on GitHub. They can then declare their intent to participate in the challenge, describe what they aim to create or which tools or input data they intend to use, and post comments about their or other participants' progress. I will refer to an author's explanation of their project (which is usually posted on their issue page, but also seen in README files), as the author's statement because of its similarity in function to the artist's statement. Typically, the participant will then create their own repository to use while working on their entry, and share a link to their repository in their NaNoGenMo issues page in order to connect their entry to the challenge. For example, Janelle Shane's issue page for their 2019 entry #103, *How to begin a novel: Upgraded version*[12], starts with the author explaining that they used their own crowdsourced dataset. They then describe and show samples of a previous NaNoGenMo entry they wanted to improve, and explain that they used a "neural net called GPT-2" to generate their text. Shane then shares samples from their generated novel, and GitHub users are able to post comments.

NaNoGenMo participants spend their free time in the month of November writing computer code with the aim of generating output text, and finally selecting one or more outputs as their final generated novel. The code, input data, and other files needed for the project are typically found in participants' repositories. From a paratextual conceptualization viewpoint these pieces are all author controlled paratexts, but their location on public GitHub repositories enable the easy sharing, reuse, and reworking, which is supported by the challenges' open-source ethos. In Shane's issue page, they include a link which ties the page to their own NaNoGenMo project repository[13]. In their repository, Shane provides additional information about their generated

[12] https://github.com/NaNoGenMo/2019/issues/103

[13] https://github.com/janelleshane/novel-first-lines-dataset

novel project in the form of a README file[14]. The repository also contains their crowdsourced dataset[15], and plain text files containing their generated text[16]. Shane does not include code files but elsewhere links to the neural network model[17] that they trained and ran in order to generate the texts.

NaNoGenMo is also a community of creators and supporters. As well as commenting on current NaNoGenMo projects, the participants and supporter's comment discussions serve to maintain a memory of previous generated works which they reference and associate with newer ones, and they also reuse and rework previous work's elements into new projects, as Shane did in their 2019 entry. They also share samples of NaNoGenMo works on social media, and write press articles discussing what they judge to be the highlights of the year. Thus, the community as well as individual authors are producers of generated novel paratexts. For example, Zachary Littrell, a contributor to editorial book site *Book Riot*, writes a press article about the annual challenge and their favorite entries between 2013 to 2017 (2017). Littrell provides commentary, generated text samples, and links to each project's repository, which make this article an example of a paratext which takes the form of critical commentary.

For a detailed description and analysis of a NaNoGenMo project and its paratextual elements on GitHub, see *Appendix B*.

### 2.2.1. NaNoGenMo paratext questions in relation to the research project

Because there is a lack reader studies about creative generated text in general, it's therefore unclear how generated novel paratexts impact reading, how important each one may be to interpretation and reception, and if different groups of readers receive a text more positively depending on certain paratexts. For example, the NaNoGenMo rule appears to prioritize the generated text and the code because these are the pieces which must be shared publicly, but that does not mean that both these, or only these, are perceived to be the most important to readers and that they conform to Genette's view of there being one prioritized, central text. For

[14] In programming standard practice, a README file is a piece of documentation which contains introductory information or guidance about the code it accompanies.

[15] https://github.com/janelleshane/novel-first-lines-dataset/blob/master/crowdsourced_all.txt

[16]https://github.com/janelleshane/novel-first-lines-dataset/blob/master/iteration150_temperature0p8_victorian.txt

[17] https://github.com/minimaxir/gpt-2-simple

example, it's unknown whether the author's statement as a paratext might be perceived to be more important than the code files. It's assumed that readers' valuing of a generated text is likely to be impacted by the clarifying technical explanations and creative intentions which are typically described in NaNoGenMo author statements, but this is yet to be confirmed by an empirical study. While some readers might consider the code to be the element which is central to a generated novel project, and therefore held to be more important than the generated text, it is not known whether this could be dependent on a reader's familiarity with programming or not. Further, a well-known text used as input or training data for a generated novel might be recognized by readers, and therefore function as a hypotextual relationship. These questions are investigated and discussed in *Chapter 4 Reading experiment and survey study*.

### 2.2.2. NaNoGenMo project pieces

Van Dijk describes web paratext as "…expanding into an infinite online context" (2014, p. 24), and it would therefore have been difficult to identify pieces as individual units if it were not for the fact that most NaNoGenMo projects' GitHub pages tend to have certain information on specific pages. However, this separation of information is by no means a clear rule; segmenting a NaNoGenMo project into distinct paratextual pieces can nonetheless pose some challenge because each project is unique and doesn't necessarily conform to all established conventions. Classifying and pigeon-holing more traditional forms such as the print novel can also be difficult when analyzing each individual work. The reason why I conceptualize NaNoGenMo projects into distinct pieces is so that the paratext can be operationalized not only for the two studies in this research project, but for ease of discussion as well.

In *Table 1: NaNoGenMo project pieces and their relationships* I operationalize and systematically describe each project piece and its relationship to other pieces. Their schema names (piece A, piece B, etc.) are also listed because these will be used throughout the rest of the research project. For some of the pieces (for example *The code*, *The code repository*) the conventional name is used and it is consistent with the term used in NaNoGenMo or generative literature projects more generally. However, the *creator's progress page* is named in order to avoid priming the survey study participants with literary related terms such as 'author'. The table also includes links to examples of each piece. The table is also reproduced in *Chapter 4 Reading Experiment and Survey Study* for ease of reference. For a detailed description of a NaNoGenMo project and how its paratextual elements map onto the table schema, see *Appendix B*.

Table 1: NaNoGenMo project pieces and their relationships

| **Name and link to example** | **Description** |
|---|---|
| Piece A | **The generated text.** This is output by the code (piece B) and is usually presented in a PDF or plain text file format. It can be shown as plain output, or have elements such as a title, table of contents, and chapter headings. |
| Piece B | **The code.** This generates the generated text (piece A). This is typically done by reading and processing an input (piece G), or by training on an input dataset in the case of a machine learning model. |
| Piece C | **The code repository.** This contains all the files which were used to create and describe the project. These include the generated text (piece A), the code (piece B) and the input (piece G). It also includes a README.md file where the project author describes their project. |
| Piece D | **The creator's progress page.** This is a GitHub issue page which links from the *Issues* in the NaNoGenMo page (piece E). Here, the project author declares their intent to participate in the challenge and can document their creative and development process by posting comments. Other GitHub users can also post comments. Each issue page links to the project's generated text (piece A) and its code repository (piece C). |
| Piece E | **The NaNoGenMo page.** Structurally, this page is a GitHub repository and it links to all the issue pages/creators' progress pages (piece D) which participate in the challenge. There is a new repository for every year of the challenge, and it describes how to participate, the challenge rule, and offers some links to resources. |
| Piece F | **Media page.** This is a press article where a critic introduces and discusses the NaNoGenMo challenge and a selection of projects. This page is situated outside of the challenge and outside of GitHub, although links to press articles can be accessed from the NaNoGenMo page (piece E). |

| | |
|---|---|
| Piece G | **The input.** This is a file which is read and processed, or trained on, by the code (piece B) during the text generation process. The input can take many forms. For example, it could be found data such as the text of a classic novel or a spreadsheet of city library addresses, or it could be a list of words written by the project author. |

### 2.2.3. The two NaNoGenMo projects used in this research

Two NaNoGenMo projects are used as study items in both studies; Ranjit Bhatnagar's 2015 entry *Molly's Feed*[18], and Janelle Shane's 2019 entry *How to Begin a Novel: Upgraded Version*[19]. Both authors were emailed and gave their consent for their projects to be used. The generated novel projects were carefully chosen based on their characteristic similarities and differences. I identify technical and cultural characteristics and throughout both studies collect data and analyze the impact that this has on participants' value judgements and reception.

In preparation for use as study items I segmented each project into seven pieces which represent paratextual elements. Each piece is described in *Table1: NaNoGenMo project pieces and their relationship* from pieces A to G, and I will use this schema to refer to specific components of *Molly's Feed* and *Victorian* for the rest of this research project. In this form neither piece is treated as central or hierarchically above another, although some pieces of course have more links than others. For example, I describe *The creator's progress page* (piece D) and *The code repository* (piece C) as having each having links to 3 other pieces. Notably, I initially treat the generated text (piece A) as a paratext in the sense that it isn't automatically assumed to be more important or central than other pieces in a NaNoGenMo project. I do this in order to test whether new readers actually do judge it to be more important or central than the other pieces. However based on my own experience, if a generated novel project is encountered on social media or is otherwise linked outside of GitHub then it is often shared as a sample of the generated text. The impact of the generated text (piece A) and the other paratextual pieces

[18] Ranjit Bhatnagar, NaNoGenMo 2015, project #170. https://github.com/dariusk/NaNoGenMo-2015/issues/170

[19] Janelle Shane, NaNoGenMo 2019, project #103.

https://github.com/NaNoGenMo/2019/issues/103

(pieces B-G) is measured and discussed *Chapter 4 Reading Experiment and Survey Study*.

#### 2.2.3.1. Molly's Feed

Ranjit Bhatnagar's *Molly's Feed* (2015) is inspired by a monologue spoken by Molly Bloom, a character in James Joyce's esteemed 1922 novel *Ulysses*. On their NaNoGenMo project's GitHub issue page, Bhatnagar presents the generated text by encouraging readers to "Follow along as Molly scrolls through her Twitter feed"[20]. To create the generated novel, the author wrote a list of phrases while they read Joyce's monologue, and then used these as search terms to automatically find matching tweets on Twitter. The results are ordered by length with the shortest last, in order to mirror a sense of breathlessness which Bhatnagar notes in Joyce's monologue.

#### 2.2.3.2. Victorian

Janelle Shane revisits a project from 2017 and 'upgrades' the neural network model to create *How to Begin a Novel: Upgraded Version (*2019). In their GitHub repository, the author describes how they crowdsourced the creation of a training dataset by asking people to pick a novel or other form of written fiction and to contribute its first sentence to the dataset. Shane then uses it to fine-tune a GPT-2 neural network model in order to tune it to generate text whose potential styles may appear to resemble those of the opening lines seen in the dataset. This study uses a sample of *Victorian*, a text which Shane generated by inputting a prompt which the model used to generate sections of text which each start with the phrase "It is a truth universally acknowledged"[21] - this phrase is part of the first sentence from Jane Austen's 1813 classic novel *Pride and Prejudice*.

### 2.2.4. Key differences and similarities between the NaNoGenMo projects

Both NaNoGenMo authors have participated in the annual challenge more than once, and they are known for their creative work outside of it. The two projects are both considered to be good; this appraisal is based not only on my own judgments and on the comments left by readers on the project's GitHub issue pages, but based on the fact that both are described in press articles (media pages, piece F) as NaNoGenMo highlights.

[20] https://github.com/dariusk/NaNoGenMo-2015/issues/170

[21] https://github.com/janelleshane/novel-first-lines-dataset

Both Joyce's and Austen's works are generally considered to be literary classics, plus are used in pop-culture references and film adaptations. Where Bhatnagar (2015) explicitly states that they are using text from *Ulysses*, Shane (2019) does not explain that their prompt is from *Pride and Prejudice*. Shane's crowd-sourced dataset, which takes the form of the input data (piece G) in this study contains lines from popular novels which might be recognized by participants. I identify these relationships as hypotextual relationships, especially between the literary classic and the generated novel.

The methods and generation technology used by the project authors are in sharp contrast to each other. Where Bhatnagar (2015) uses a comparatively simple software package which searches for tweets on the social media platform Twitter, Shane (2019) generates their text using GPT-2; a then state-of-the-art and arguably (over) hyped by the technology news media as being a controversial for its potential to be used to generated fake text-based content.

Another potentially significant and interesting difference between the two NaNoGenMo projects used in this study is their code. Although both projects use the Python programming language and are presented using the Jupyter Notebook[22] format, their purposes and intentions are very different. Shane (2019) links to the GPT-2-simple package that they use to fine-tune the model, which features a simple code demo showing how to use the package. Bhatnagar (2015) on the other hand uses code comments and keeps older lines of code to show their development process and narrativizes their progress. Bhatnagar's code is arguably not what would be typically described as clean, concise, and easy to understand, but it is instead telling the story of progressively improving the code (and therefore the generated novel) by not initially showing the reader the final, clearest and most concise version of the code.

[22] https://jupyter.org/

# Chapter 3 Research Framework

Where the previous chapter outlines the concept of paratext and how it relates to the generated novel, this chapter sets out the research design rationale and plan. This begins with discussing how the paratextual conceptualization theoretically underpins the research questions, and then advances to outline an overview of the research steps which make up this explanatory sequential mixed-methods research project. Finally, reader theories are briefly outlined in order motivated the reader-focused research design choices, and the research's constructivist orientation is thoughtfully motivated.

## 3.1. Research design in relation to the research questions

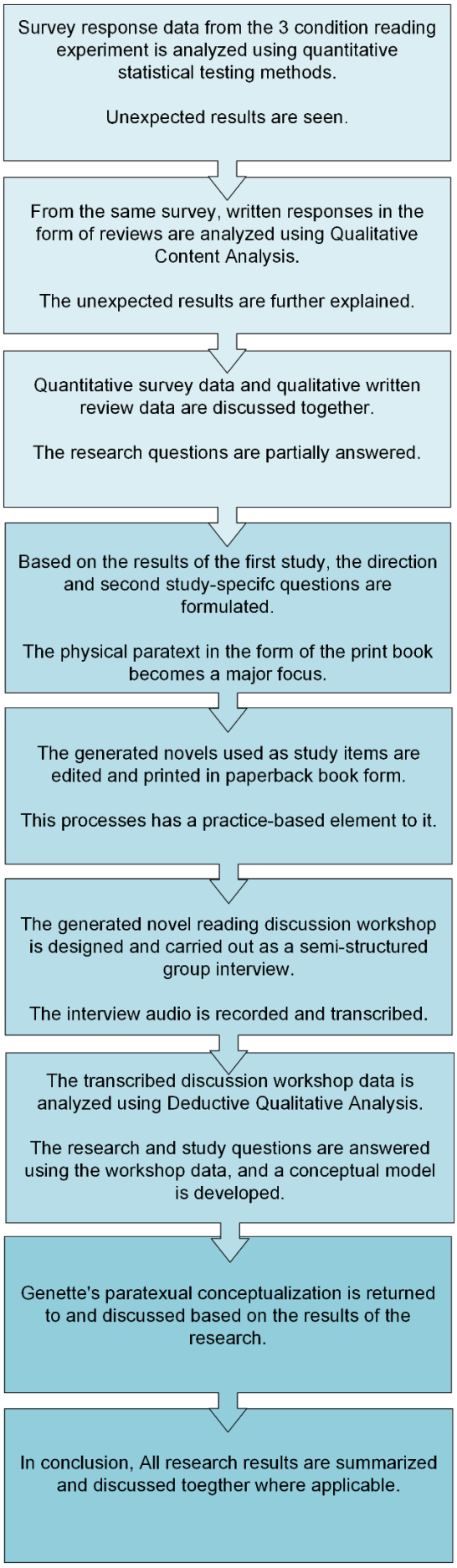
This research project asks in what ways does paratext influence potential readers' interpretation and reception of a generated novel. Broken down into three research questions these are:

1. Which paratextual elements have a central role in influencing the understanding and reception of a project?
2. Which reader skills and experience affect the influence of these paratextual elements?
3. Which technical and cultural characteristics of a generated novel affect the influence of these paratextual elements?

The questions are built around the concept of paratext, which means that there is already one theory underpinning the research project. This means that my research design and methods are not aiming to develop a new theory or identify an existing one that can describe the results. Even with possible challenges or inconsistencies or the need for theoretical refinements, the general theory around what paratext is and how it functions is expected to remain useful and capable of describing the data and answering the research questions. Because there is already a working theory supporting the research questions, then, and because the questions are more specific, the study analysis methods must therefore necessarily have at least some deductive dimensions to them because the research aim is to focus on collecting and analyzing data that can answer the research questions. Thus, although I designed the qualitative aspects of this research to have stages in the methods which allow for substantial flexibility to support the development of inductive analysis, a deductive approach is necessarily in the initial stages of each study because of the specific research questions which are based on paratextual theory.

The research questions can be answered by both quantitative and qualitative methods, with questions 1 and 2 being especially suited to the former because individual paratextual elements and distinct reader skills can be operationalized into categorical variable and counted. I note that Genette's paratextual conceptualization already presupposes that a work and its supplemental or periphery elements are identifiable as distinct pieces. Coincidently, this makes operationalizing paratextual pieces into individual variables and measuring their impact by including them or taking them away from a reading situation easy to do. This ability to easily operationalize paratextual pieces was the starting point for designing this research. The reading experiment and survey study was designed first and uses the two generated novels as study items. The reading experiment was structured into 3 different paratextual reading conditions, where the generated novel samples and their paratexts were operationalized and presented in 3 different combinations. Differences in participant responses based on these conditions were collected using a survey, which also collected written responses from participants in order to have collect qualitative data with which to contextualize the quantitative survey data with. After data collection, the initial unexpected results (as detailed in section *Chapter 5 Reading Experiment and Survey Study*) motivated an explanatory research design because I was interested in further investigating and later collecting additional data in the second study that could explain the unexpected results. A high-level overview of each step in the research design is seen in *Figure 1: Research Steps Overview*.

Figure 1: Research steps overview



## 3.2. Research design: explanatory sequential mixed-method

The quantitative and qualitative nature of the research questions led me to a mixed-method research design which I developed into an explanatory sequential design based on the unexpected emerging results of the survey study questions. For example, it was unexpected that the likert-type question responses about literary value were the only such responses that were, in all cases but one, not impacted by the presence of the generated novel paratext.

Schoonenboom and Burke's paper 'How to Construct a Mixed Methods Research Design' (2017) detail the many aspects of mixed-methods design and use it as a terminological guide to explain my research design. My research design is motivated by a pragmatic approach where I selected data collection and analysis methods based the research questions, contexts, and resources that were available to me rather than defaulting to the textual analysis that I was most familiar with. For example, although I do not have previous experience carrying out survey studies, this was the most practical method of collecting measurable reading data from a large number of participants during the global pandemic.

Schoonenboom and Burke review mixed-method design topologies, and name explanatory sequential design as a well-known design where the "…first phase of quantitative data collection and analysis is followed by the collection of qualitative data, which are used to explain the initial quantitative results" (2017, p. 117). More specifically, the authors also describe a multilevel mixed design where "In these parallel or sequential designs, mixing occurs across multiple levels of analysis, as QUAN [quantitative] and QUAL [qualitative] data are analyzed and integrated to answer related aspects of the same research question or related questions" (p. 118). Multilevel mixed design specifies that quantitative and qualitative results can be integrated at the analysis stage and used to answer the same research questions. Ultimately, this is what I do.

As can be seen in *Figure 1: Research Steps Overview*, the research follows a sequential-dependant design; studies were carried out and analyzed sequentially, and subsequent steps in the research process were based on the findings which preceded it. For example, for the second study I initially intended to conduct online interviews to observe how participants access and use a NaNoGenMo novel's digital. But based on the outcomes of the survey study I altered the second study's research direction to instead foreground the physical paratext. As Schoonenboom and Burke's describe, this is an established path in research design: "Dependent research activities include a redirection of subsequent research inquiry. Using the outcomes of the first research component, the researcher decides what to do in the second

component. Depending on the outcomes of the first research component, the researcher will do something else in the second component. If this is so, the research activities involved are said to be sequential-dependent, and any component preceded by another component should appropriately build on the previous component". (2017, pp. 114 - 115). Despite the dependent aspect of the research design, both studies have equal status meaning that the results of one are not prioritized over the other.

While some results from both studies are compared as throughout this project as needed, all results are integrated and discussed in *Chapter 6 Discussion*. However, comparison between results from the different studies is of course limited because the data from each was collected at different times and using different methods, as expanded upon in *Chapter 6*.

## 3.3. Readers, interpretation, and reception

This subsection motivates the research design choices made in relation to studying reader interpretation and reception. A short summary traces some of the disciplinary lineage around studying readers in the literature and media space, and also flags methods which I use or which are methodologically developmentally related. I critically discuss reader response and methods in relation to my research in *Chapter 6 Discussion*.

### 3.3.1. Reception and value

Understanding how different paratexts and combinations of paratextual elements influence value judgements, and therefore influence reader reception, is a step towards better understanding how generated novel projects function as digital culture artifacts which can be engaged with. In my research, the concept of value began from a cultural studies perspective and where value is subjective and not universal. For example, Brookey and Gray (2017) are of the opinion that hierarchies of value differ between people, between communities, and across time, and that this subjective value also means that the value of paratext can change (p. 102). This means that generated novels cannot be expected to be valuable to everyone, and not in the same ways. For example, a programmer may only be interested in a project's code elements because she is intrigued by the technical challenge but does not consider the project to have any literary merit. To her, the generated text may just represent output which confirms that the code is working as expected. Conversely, her friend may be unable to understand the code but may feel that the generated text reminds them of poetry or a stream-of-consciousness literary style, and they might therefore value the piece as a work of literary experimentation.
Literary, creative, and technical value are all subjective forms of value, and paratext may influence the ratings of each of these values. Together, these values compose a more general

measure of reception in the context of the generated novel space. In the survey study, these three types of value are measured as dimensions of readers' reception because the discourse around generated novels has referred to them as literature, creative, and technology works. I also measure overall value as based on previous research. These values are expanded on in *Chapter 5 Reading Experiment and Survey Study*.

### 3.3.2. Reader response, Interpretive communities, and reception studies

Prominent reader response theorist Wolfgang Iser's work breaks with previous convention by shifting literary theory's sole focus on the text, to include the reader as well. Iser's core stance is that the interaction between the text and the reader is central to the study of every literary work (Iser, 1978, p. 21). Iser's theories are anchored around the implied reader, which is a conceptualization of the reading processes rather than the empirical study of actual living readers. I critically discuss Iser in relation to my research in *Chapter 6 Discussion*.

Another prominent reader response theorist and critic, Stanley Fish, theorizes from the position that the audience of a literary text is relevant to study (Leitch, 2001, p. 2068). Fish is known for the concept of the interpretive community which posits that readers' interpretive strategies derive from educational and professional communities where a reader's training, rather than the text, govern and generate interpretation (p. 2069). Here, interpretive communities do not "…mean to indicate a group of people but rather a collection of norms and strategies held in common" (Livingstone and Das, 2013, p. 4). In their review of key literature on interpretation and reception, media audience researchers Livingstone and Das link Fish's concept of interpretive communities to reception studies (2013). In Livingstone's introduction to the field, they explain that the interpretations of different audience groups and empirical methods such as interviewing, discourse analysis of audience talk, and ethnographic observation are characteristic of the field (2019, p. 1). From the field's outset, reception studies did not share Genette's prioritization of authorial and publisher intent. Reception studies did not

> "…simply assume that audiences would automatically interpret media texts in ways that either their producers or their critics blithely supposed. Hence, the ways in which audiences interpret media texts was recognized as an empirical question, one that demanded the combined analysis of media texts with that of media audiences" (Livingstone and Das, 2013, p. 1)

Livingstone explains that the specific strength of reception studies is the emphasis it places on "…the empirical and diverse reception of specific textual features, conventions, genres, or codes" (2019, p. 1). I have taken some direction from reception studies in my own study designs and analysis. However, with regards to reader response theory as shaped by Iser and Fish an

important difference with my research is that I focus on studying individual, actual readers and not Iser's model implied readers nor Fish's collection of model norms and strategies. And further, I focus on individual readers rather than studying audiences as a whole or communities in terms of social interaction and dynamics. This is expanded on further in *Chapter 6 Discussion*. My studies seek to test how different paratexts may influence readers' reception of a generated novel, and if the readers' literary interpretation skills and/or computer programming skills may enable people to have different reading experiences depending on which paratexts do or do not resonate with them. For example, readers with computer programming skills may judge a generated novel read with its code to be more interesting than one read without. The concept of interpretive communities and reception theory have therefore been useful in motivating and designing the survey and interview studies, despite the differences I outline. I note that Agafonova et al. (2020) have also recognized the relevance of applying reception theory to the study of creative generated texts produced by their Paranoid Transformer project (p. 3).

## 3.4. Ontology and epistemology

This research project is relativist-constructivist in orientation. I draw from Braun and Clarke's (2021) theory chapter to articulate my ontological and epistemological position.

I am of the opinion that a constructivist epistemological orientation is compatible and fits unproblematically with quantitative statistical methods and analysis. This is because statistical results are only valid in a meaningful way or in a useful way if they make sense when interpreted within the discipline, domain, or context that they're measuring or being applied to.
In terms of my epistemological position towards the analysis of the survey's written reviews data, this is perhaps less clear-cut. On the one hand, a constructivist orientation make sense because several reviews' account of how the code or technology works is different from fact. It also appears that some reviews may have assumed that piece G, the Input data, was the generated text itself - this is also different from the actual state of affairs. So in order for these reviews to be considered valid and their most valuable, they could be approached with a constructivist orientation.

On the other hand, unlike other the workshop study data the written review data originates from an reading experiment with a between-group design, meaning that by definition there are aspects of the study, such as the presence and absence of paratextual information and what my specific research questions are, that the participants may not be aware of – but this 'objective truth' of the study is known to me, as the study designer, researcher, and generated novel expert. This might entail that the assumption of there being an objective truth is baked right into

the foundations of the study. And yet, several of the broader content analysis categories primed and discussed in the reviews, such as *Literature and literary* and *Creativity*, are not based on the study itself – they are based on the participants' own interpretation and interaction with their perception of the world as shaped by social and cultural influences. Indeed, this arguably leads to a philosophical orientation question about the paratextual framework itself; on the one hand Genette's privileging of the interpretation put forward by the author and publisher could be seen as a positivist view. But on the other hand, this could also be seen as an acknowledgement of there existing more than one objective truth where the authorial and publisher paratext is trying to steer readers towards one of many interpretations. This last point seems compelling in terms of philosophical rationale; I am satisfied that all aspects of the research can align with a constructivist orientation.

## 3.5. Research ethics and data management

The research has been approved by the University of Southampton Faculty Research Ethics Committee ERGO numbers: 61454.A1 and 50151.A1. All participant informed consent procedures, study execution, and data storage and analysis has been conducted by strictly following the approved research applications. The study data is only stored on university systems and university machines which are password protected and managed in accordance with the university's data management and storage regulations and guidance.

# Chapter 4 Reading experiment and survey study

This chapter begins by reviewing related research studies and then details the design and set up of an online reading experiment where the generated novels are presented in one of three different paratextual conditions. Participant responses are recorded through a survey and written reviews where participants give their personal opinions about the work they read.

The results are analyzed quantitatively for significant differences between paratextual conditions and reader groups across six value dimensions. The results suggest that, while the presence of the paratextual elements in the generated novels have a measurable impact on readers' perceptions across virtually all the value dimensions, the presence of the generated text itself doesn't significantly impact reader valuations. Further, the results show that literary value is barely impacted by the presence of the paratext, which is unexpected. This is further investigated quantitatively through participants' familiarity with the classic novels that have a hypotextual relationship with the generated novels. The perceived ranked importance of each paratextual piece by participants is also analyzed.

The second part of the chapter focuses on the written reviews and uses Qualitative Content Analysis to further investigate the unexpected literary value result. Here, I develop categories to compare differences in the review content between paratextual conditions and reader groups. I conclude analysis by advancing to discussing specific reviews and identify patterns, most notably a developing theme which describes how some reviews reject literary value and relocated it to technical value.

## 4.1. Related work

Based on my literature review in *Chapter 2*, a reader study which aims to understand different groups' interpretations and overall reception of the emerging generated novel form is missing from the literature; current research focuses on academic critics' appraisals. More generally, there is a lack of research focusing on generated novels, and indeed other forms of creative text generation. Of the research which does, the majority relies on the judgments of 1-2 researchers (for example Henrickson 2018, Van Stegeren and Theune 2019, and my own 2019 MSc Web Science dissertation) and do not consider the judgements of groups of readers.

For this research project, Henrickson (2019), McGregor et al. (2016), and Koolen et al. (2020) are the most important related works because each focuses on the opinions of non-expert or ordinary readers and, demonstrates how to design reader reception surveys which attempt to

measure creative, literary, and overall value.

Henrickson's doctoral thesis titled *Towards a New Sociology of the Text: The Hermeneutics of Algorithmic Authorship* (2019), and the subsequent book *Reading Computer-Generated Texts* (2021), center on the complex question of author attribution in the case of generated texts, which Henrickson describes as being created through algorithmic authorship. They research how algorithmic authorship this is perceived by readers. Henrickson considers the generated text within the larger "...modern textual landscape [which is] permeated with various modes of human-computer collaboration". Motivating their research focus, Henrickson argues in 2021 that it is too soon to concentrate research on just the output of text generation systems, because first determining the technology's unique contributions to this "...modern textual landscape" is the fundamental issue currently at hand (2021, p. 4). Henrickson therefore sets their research focus on "...readers' responses to the very concept of NLG [Natural Language Generation technology] itself" (p. 4), and presents their work as the crucial groundwork needed to study generated texts with their readers. Hence, I position Henrickson's research as the most significant and foundational previous work in this project because it has tackled the challenge of introducing the question and the study of reader reception to generated text.

While Henrickson (2019) reviews previous research in the creative space (such as prose and poetry generation), their studies are ultimately designed to collect readers' responses to generated news reports. Henrickson carries out two studies: an online survey with an authorship attribution task, and focus groups. Both studies ask readers to respond to a generated news text which reports political election results. This authorship attribution task is of particular interest to my own research because Henrickson's study design can test possible variability in its readers' responses because it incrementally increases the amount of available paratextual information. My own research design draws on this broader idea of presenting readers with varying paratextual conditions, as well as the reception elements which are part of Henrickson's focus group study.

McGregor et al. (2016) is identified by Henrickson (2019) as a previous empirical study relevant to studying "...ordinary reader's perceptions of algorithmic authorship" (p. 127) and focuses on reader evaluations of generated poetry. McGregor et al. echoes Flores in their belief that the public is becoming more receptive to the idea of machine creativity (2016, p. 51), and that readers in this particular period in the history of technology and art are prepared to engage with computational works without "...losing regard for the inherent degree of creativity" (p. 54). McGregor et al. is run as a survey which tests reader's creativity judgements of generated poems across three different framing conditions.

McGregor et al. (2016) is relevant to my research not only because it demonstrates how a reading and evaluation study with three different paratextual conditions might be conducted as

a survey, but it also serves as an example of why more nuanced questions might be useful to ask in evaluation studies. The authors were unable to show statistical significance in the change of creativity ratings across the three framing conditions, and concluded that "Quality is arguably a somewhat vague category" (p. 58), and that more nuanced questions (among other points) are needed to improve the evaluation process (p. 59). However, it's possible that two of the criteria which McGregor et al. pose as questions - creativity and quality - may be too vague to be used as evaluative questions on their own without being grouped with finer-grained dimensions. Initially, my own survey design addressed this by decomposing creativity and overall quality (as well as literary and technical value) into finer dimensions. This can be seen in *Appendix Table A2: Survey Questions Grouped by Type*. But because of the richness of the qualitative data in this study, the finer dimensions ended up not being useful for the formal analysis. Indeed, I previously did an initial descriptive statistics pass through the more general and finer quality dimension survey questions. I found that the results of the finer dimensions were very similar to the more general dimensions that they were describing. This is a good sign in terms of survey question design, but not necessary to spend time reporting the results of for this particular study, because it would be very repetitive.

Lamb et al. (2018) surveys and recommends evaluation methods for measuring creativity in humans and computers (p. 1). Of particular relevance are Lamb et al.'s sections *8.1 Implementations of Models and ad hoc Tests* and section *8.2 Opinion Surveys, Non-Expert Judges, and Bias* because they have been informative in developing questions for this study's survey. For example, research-based guidance on designing questions for non-expert evaluations of computational creativity works and systems.

Koolen et al. (2020) is situated in computational literary studies and conducts a survey about readers' literary quality judgments of novels, and I have drawn from it to design my survey study. The researchers argue that literary quality can be influenced by "…text-extrinsic social factors" (p. 1). Koolen et al. specifically focuses on the reception of novels by the Dutch general reading public in order to learn whether literary quality judgements (i.e. to what extent do readers consider a book to be literary) can be distinguished from overall quality (to what extent do readers consider a book to be good or bad overall). No definition for literary or overall quality was given to participants because personal opinions about literariness were sought - the same is done in my survey study. The decision to include ratings for overall value is also informed by Koolen et al., although I decompose it into two dimensions and one attention check question in my survey: enjoyable, interesting, and boring.

Koolen et al. (2020) cite three other previous works whose designs have also contributed to the development of my survey questions. Working within media psychology, Busselle and Bilandzic

(2009) create a scale for measuring narrative engagement in film and television programs. They refine a list of aspects of experiencing a narrative which are considered to be the most fundamental and accessible to audiences (pp. 321 - 322). Kuijpers et al. (2014) work within empirical literary studies and also develop and validate a self-report scale using similar methods to measure different aspects of a readers' absorption in the story world of a text-based narrative. In both studies, enjoyment is an outcome or evaluative response to narratives, and I therefore include it as a dimension of overall value in my own survey study. Both studies run multiple rounds of surveys with large sample sizes in order to develop a robust scale using factor analysis, but this approach is outside the scope and available resources of my research.

Miall and Kuiken (1999) are also cited in Koolen et al. (2020) and focus on the empirical study of literary readers, and identify reading time as a literary reader variable in their study on literariness and evaluation. Therefore the amount of time that participants spend on the reading part of my survey has been recorded in the collected data. But it was ultimately not used as part of the analysis presented here because the survey response and qualitative written review data were rich enough without the addition of reading time data.

## 4.2. Method and study set up

This section discusses the methodology and design of my online reading survey study. The survey study was designed by drawing on the related work discussed in the previous section.

The study is a controlled reading experiment carried out as an online survey to test if different paratextual conditions, different generated novel projects, and different skills influence new reader's interpretation and reception of NaNoGenMo projects. It also tests whether the cultural and technical characteristics of the projects influence reception. The paratextual experiment conditions do not reflect real life conditions, but the data from the tested conditions can be used to answer questions that the real world cannot easily control for. This is also why this survey study was followed by the second qualitative study in *Chapter 5 Workshop study* in order to help contextualize and explain the quantitative survey data which was collected under constrained experiment conditions. For example, a constraint on reading time.

The study is designed in three parts which are linked together online: the consent form, the reading part, and the survey part. After choosing to participate in the study through the participant recruitment platform Prolific[23], participants must agree to the consent form or choose not to participate. Next comes the reading part which is designed as a website where

[23] https://www.prolific.co/

participants follow on-screen instructions. They read one of the two NaNoGenMo projects which is presented in one of three paratextual conditions. Each paratextual condition is composed of different combinations of paratextual project pieces – for example, the *Generated Text Condition* only shows readers the project's generated text sample. After reading, the participants then follow a link to the survey part and respond to questions which gauge their value judgements and reception of the project based on the paratextual condition which was shown in the reading part. Participants also respond to questions about their experience with literary forms and technologies or platforms which are related to the projects or to generated text more generally. The collected data enables the impact of the paratextual conditions on reader judgement values and reception to be tested, as well as the perceived importance of each project piece.

For convenience*, Table 1: NaNoGenMo project pieces and their relationships* is reproduced here. Because *Table 1* segments and describes each paratextual element that is considered in the study, it demonstrates how I have chosen to delineate and segment paratextual pieces. It's therefore an applied example of *6.2.4. My minimalist paratextual conceptualization*, which I detail in *Chapter 6 Discussion*.

Table 1: NaNoGenMo project pieces and their relationships

| **Name and link to example** | **Description** |
|---|---|
| Piece A | **The generated text.** This is output by the code (piece B) and is usually presented in a PDF or plain text file format. It can be shown as plain output, or have elements such as a title, table of contents, and chapter headings. |
| Piece B | **The code.** This generates the generated text (piece A). This is typically done by reading and processing an input (piece G), or by training on an input dataset in the case of a machine learning model. |
| Piece C | **The code repository.** This contains all the files which were used to create and describe the project. These include the generated text (piece A), the code (piece B) and the input (piece G). It also includes a README.md file where the project author describes their project. |
| Piece D | **The creator's progress page.** This is a GitHub issue page which links from the *Issues* in the NaNoGenMo page (piece E). Here, the project author declares their intent to participate in the challenge and can document their creative and development process by posting comments. Other GitHub users can also post comments. Each issue page links to the project's generated text (piece A) and its code repository (piece C). |
| Piece E | **The NaNoGenMo page.** Structurally, this page is a GitHub repository and it links to all the issue pages/creators' progress pages (piece D) which participate in the challenge. There is a new repository for every year of the challenge, and it describes how to participate, the challenge rule, and offers some links to resources. |
| Piece F | **Media page.** This is a press article where a critic introduces and discusses the NaNoGenMo challenge and a selection of projects. This page is situated outside of the challenge and outside of GitHub, although links to press articles can be accessed from the NaNoGenMo page (piece E). |

| | |
|---|---|
| Piece G | **The input.** This is a file which is read and processed, or trained on, by the code (piece B) during the text generation process. The input can take many forms. For example, it could be found data such as the text of a classic novel or a spreadsheet of city library addresses, or it could be a list of words written by the project author. |

### 4.2.1. Paratextual conditions and survey variations

The study tests three paratextual conditions: the *Generated Text Condition* (1), the *Paratext Condition* (2), and the *Both Condition* (3). It also tests the two NaNoGenMo projects: *Novel B* representing *Bhatnagar* (2015), and *Novel S* representing *Shane* (2019). This makes a total of six survey variations: B1, B2, B3, S1, S2, S3, as shown in *Table 2: Study Design*. The *Generated Text Condition* only shows participants the generated text (piece A), the *Paratext Condition* only shows participants all project pieces (pieces B, C, D, E, F, G) except for the generated text, and the *Both Condition* shows participants both the generated text and all the project pieces (pieces A, B, C, D, E, F, G). Due to the way in which the Prolific platform is designed, each survey variation had to be run on smaller sub-samples of eight participants each, in order to filter for those with the intended demographics and reported skills. This is represented in the third column in *Table 2* which lists the skills and gender identity of each participant sub-sample, with each survey variation having the same sub-sample combinations. This is an applied example of operationalizing paratextual pieces as described in *6.2.4. My minimalist paratextual conceptualization*, which I detail in *Chapter 6 Discussion*.

Table 2: Study Design

| **Survey Version** (48 participants per survey) | **Maximum estimated time taken to complete** | **Paratextual Content** (96 participants per condition) | **Participant Sub-Samples Grouped by Skills and Gender Identity** (8 participants per sub-sample) |
|---|---|---|---|
| B1 | 25 minutes | Piece A from Bhatnagar (2015). | Literature & code, women+<br>Literature & code, men+<br>Code, women+<br>Code, men+<br>Literature, women+<br>Literature, men+ |
| B2 | 40 minutes | Pieces B, C, D, E, F, G from Bhatnagar (2015). | Literature & code, women+<br>Literature & code, men+<br>Code, women+<br>Code, men+<br>Literature, women+<br>Literature, men+ |
| B3 | 54 minutes | All pieces from Bhatnagar (2015). | Literature & code, women+<br>Literature & code, men+<br>Code, women+<br>Code, men+<br>Literature, women+<br>Literature, men+ |
| S1 | 25 minutes | Piece A from Shane (2019) | Literature & code, women+<br>Literature & code, men+<br>Code, women+<br>Code, men+<br>Literature, women+<br>Literature, men+ |

| | | | |
|---|---|---|---|
| S2 | 40 minutes | Pieces B, C, D, E, F, G from Shane (2019) | Literature & code, women+<br>Literature & code, men+<br>Code, women+<br>Code, men+<br>Literature, women+<br>Literature, men+ |
| S3 | 54 minutes | All pieces from Shane (2019) | Literature & code, women+<br>Literature & code, men+<br>Code, women+<br>Code, men+<br>Literature, women+<br>Literature, men+ |

### 4.2.2. Participants and sample

A benefit of online survey studies is their ability to be easily scaled to a larger sample size. This was fully taken advantage of to create a predominantly quantitative study with a final total of 282 participants. As can be seen in *Table 2: Study Design*, each survey variation has each collected 48 responses (six participant sub-samples groups, with 8 participants in each group). The online participant recruitment platform Prolific is aimed at researchers conducting surveys, and it was used to collect completed responses from 288 participants who were paid a minimum of £5 per hour on average, which is Prolific's recommendation. In total, data from 282 individual participants was usable for the study (n = 288). All Prolific participants complete basic demographic and prescreening surveys, and researchers can filter and invite participants based on these responses and invite them to choose to complete paid surveys. For this study, demographic information about participants' reported fluency in English, age, gender identity, ethnicity, highest education completed, current country of residence, employment status, and student status was collected. Prolific's prescreening survey questions were also used to filter and recruit participants who reported having computer programming skills, and/or reported literature to be one of their top three hobbies or interests. The sample is therefore virtually equally split into three reported skill groups - 92 participants with reported programming skills, 95 participants with a reported literature hobby and no programming skills, and 95 participants with both literature and programming skills. Like the two novels, these three skill groups were

used as variables in the study. It's certain that all participants' reported skill information may not be perfectly accurate or accurately reflect skill or interest groups in lived experience, but this is a known limitation of survey studies and it has been considered be considered during analysis.

In order to improve the gender balance of the participant sample, half of the surveys were offered to participants who identified as *Female*[24] or a gender identity other than *male*, and the other half offered to those who identified as *Male* or a gender identity other than *Female*. Balancing the gender representation in the sample ensures that there is not an under-representation of a gender within each skill group. This is an important design decision because the culture around computer programming can be exclusionary to a range of demographics, including gender, leading to an over-representation of, for example, men in the practice[25]. Therefore the effort was made to balance gender representation during recruitment. Because Prolific can estimate the total pool of active participants who are eligible to respond to a study based on its demographic requirements, it was possible to check whether a participant pool is unbalanced. For example, at the time of writing there was a pool of approximately 22,400 participants who identified as male and reported having computer programming skills, but only about 11,500 participants who identified as female with the same skill. Conversely, about 16,000 female and approximately 4,900 male participants reported literature as one of their top 3 strong interests or hobbies. Therefore my design choice to make the effort to balance gender representation was sound: if I have not, then it is possible that the majority of participants with programming skills could have identified as male, and the majority of literature readers as female. In such a case differences between groups could have been confounded by the unbalanced gender variable. Fortunately, this is not the case in this study because I have accounted for it.

Collecting data about participants' skills and experience may allow them to be grouped based on, for example, familiarity with the input (piece G), as creators and critics or laypeople (for example, people who have created or studied computer generated poetry and those who have not), and as users. Again, I stress that these are not actual groups which necessarily correspond to real-world groups which participants have identified themselves with, but instead should be considered as labels which are given to participants who have responded with agreement to a question. For example, participants who agree with question *A45*, "Before taking this survey, I had used Machine Learning" are considered to have been users of Machine Learning at some

[24] The gender identities were predefined in Prolific pre-screening survey.

[25] Exclusionary cultures in Computer Science and programming have been studied throughout Margolis and Fisher (2002).

point, whether using it as part of their own project, or just completing a tutorial. Questions similar to *A45* are coarse-grained and cannot capture more nuanced distinctions between users, such as senior software engineers and those who have begun to learn to program. However, they are useful in distinguishing between participants who have some experience with a technology or platform, and those who are entirely unfamiliar with it, and this information may correlate with significantly different reception ratings across the three paratextual conditions.

### 4.2.3. Reading part

After consenting to participate in the study, participants are linked to the reading part of the study. This takes the form of a website which instructs participants to read each of the pieces to the best of their ability. I designed six different variations of the website to correspond with the six survey variations. Screenshots of the website showing the S3 and B3 survey variations can be seen in *Table C1: S3 Reading Part Website Screenshots* and *Table C2: B3 Reading Part Website Screenshots* in *Appendix C*.

Several changes were made to both generated novel projects' pieces in order to adapt them for the study, such as removing web advertizings in the press articles (piece F), and ensuring that information location and the amount of information was similar for both projects. All changes made to the project pieces are listed in *Table A3: Changes Made to Pieces* in *Appendix A*. For example, Shane (2019)'s NaNoGenMo entry presents the generated text (piece A) as a plain text file, but I have formatted the text into a PDF file with a book-like layout in order to make it appear more similar to Bhatnagar's generated text PDF.

the words 'novel', 'paratext', and 'author' are not used in the survey study's instructions to participants in an effort to avoid priming and the value-laden associations with those words. However, references to novels which are made by the project authors or press articles (the media page, piece F) are preserved. There is no rule or clear expectation about which part of a NaNoGenMo project or generated novel more generally should be read first, but the order of the pieces shown on the website are not randomized and are always shown in alphabetical order for two reasons. In my own experience, if a generated novel project is encountered on social media or is linked outside of GitHub in another way, then it is often shared as a sample of the generated text. Secondly, the technical task of randomizing the order of the website pages for each participant was too high in terms of research time and resources. Similarly, budget constraints prevented the study from testing the judged importance of each individual piece, but I am nonetheless able to test which combinations of paratexts elicit a more positive reception.

This study website is no longer online, but my GitHub repository containing the website files may be requested for viewing[26]. In accordance with my university's research ethics regulations, there is strictly no research data in the repository.

### 4.2.4. Survey part

After completing the reading part of the study, participants follow a link to the survey part where they answer the survey questions about the project based on their personal opinions and experiences. I built the survey part using the opensource survey web app LimeSurvey[27]. The time that participants spent on each question section is recorded. When the survey is complete, the participants follow a link back to Prolific to record that they have completed the study. If completed correctly (for example, the free response question is written in English), then I approve their submission and the participant is paid. Selected screenshots of the survey part of survey B3 can be seen in *Table D1: Survey Part Selected Screenshots* in *Appendix D*. A print-out of all questions in survey S3 can be seen in *D2: All Questions in Survey S3*, also in *Appendix D*.

### 4.2.5. Survey questions

As shown in the *Measure* column in *Table A2: Survey Questions Grouped by Type* in *Appendix A*, questions measuring literary, creative, technical, and overall value are Likert-type questions. These questions are on a 7-point scale ranging from *Strongly disagree* to *Strongly agree*. Questions about participant's experience and self-perception use the same 7-point scale. Question *Q6* is designed as a qualitative free response where participants are prompted to give their personal opinion and write a short review about the NaNoGenMo project.

In all surveys with more than one piece (B2, B3, S2, S3), two ranking questions ask participants to rank each piece that they read according to which ones were the most important for helping to understand the project, and which ones were the most important to helping to make the project interesting. This is an important question because while the paratextual conditions can describe groups of pieces, the ranking questions can collect finer-grained data about which individual pieces are considered to be most important.

### 4.2.6. Testing and Feedback

Insights about first-time readers' experience and skills were gained from observing and discussing with two creative text generation workshop groups in November 2020, which were

[26] https://github.com/dx9240/survey_website
[27] https://www.limesurvey.org/

designed and ran in collaboration with my then Winchester School of Art PhD student colleague Dr. Noriko Suzuki-Bosco. Next, feedback for improving the first draft of the survey study's website interface and survey questions was obtained from two talk-aloud protocol sessions conducted online. Finally, 16 volunteers participated in an online pilot test-run of all the survey variations to give additional feedback and provide the timing data needed to determine reasonable completion times for each survey variation. The website and surveys were improved and then launched on Prolific.

### 4.2.7. Collecting Data and Sample Demographics

Data was collected from 288 participants in total. Just over 61% of the participant sample are students, and overall, the sample is fairly young with an average age of about 26 years old and a standard deviation of ± 7.935. While the highest education level completed for approximately 35% of the participants is *High school diploma/A-levels*, this is closely followed by about 24% and 22% having completed an undergraduate and a graduate degree, respectively. Because the minimum age for taking the study is 18 years, and because a high percentage of the sample holds student status, this may suggest that there is a high representation of current undergraduate students in the participant sample.

As can be seen in *Figure 1: Participants' Current Country of Residence*, the sample features 27 unique countries of current residence, with the combined residents of Portugal and Poland representing about 39% of the sample. Mexico, South Africa, and Chile are the most frequent residences outside of the European continent, and combined represent approximately 10% of the participant sample. As evidenced in *Figure 2: Participant's Ethnicity*, there is an overrepresentation of White participants in the sample, which means that this dimension of the demographic data is heavily unbalanced. For future iterations of this study, the recruitment stage should undergo additional participant sampling filters to balance representation, as has already been done to balance gender representation. About 48% of participants chose *Female*, and approximately 47% chose *Male* from Prolific gender identity options. Finally, at 2% the third most represented identity in the sample is *GenderQueer/Gender Non-Conforming*.

Figure 2: Participant's current country of residence

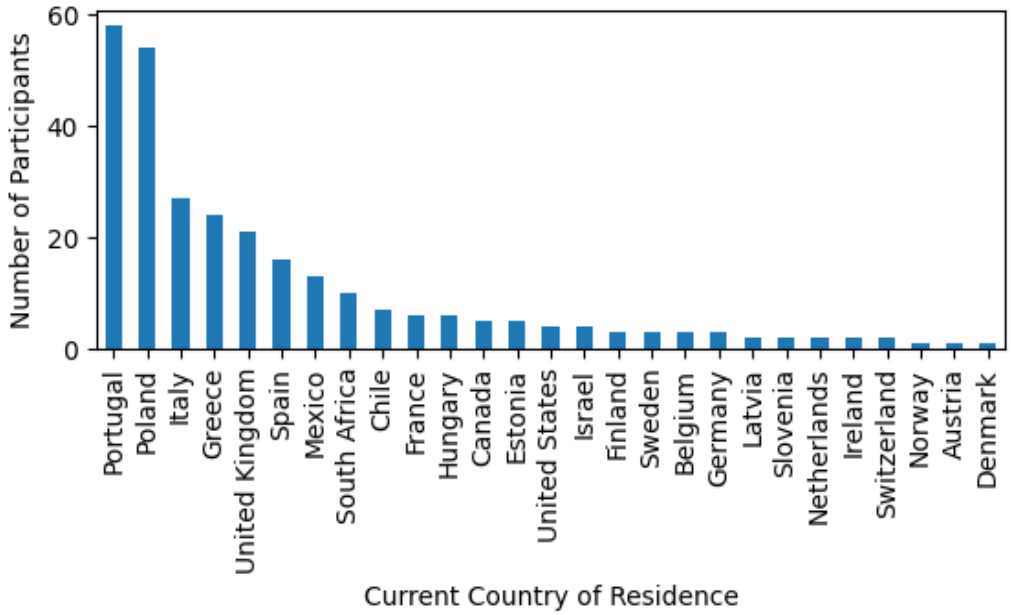
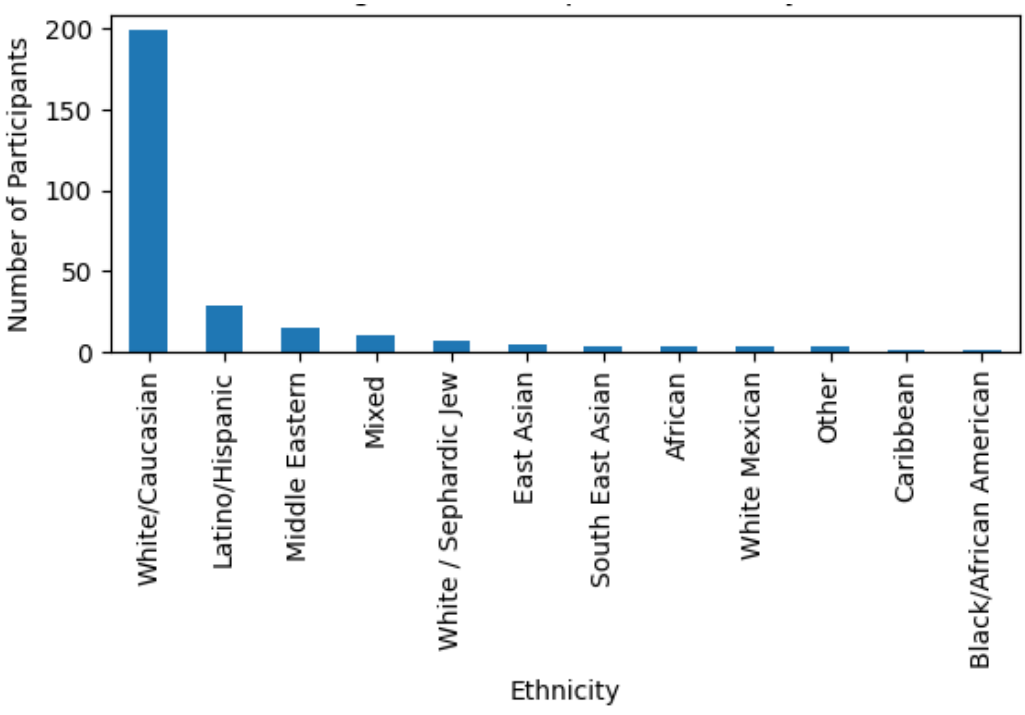



Figure 3: Participant's ethnicity



## 4.3 Statistics and computational tools

This sub-section is written for the benefit of readers who may not be familiar with statistical analyses and would like a brief overview of the terms and methods used in this chapter. Statistical method choices are also motivated here.

The R open-source statistical programming language[28] was used to prepare the data and write code with which to process the analysis[29]. Compared to popular spreadsheet and data analysis applications such as Microsoft Excel and IBM SPSS, R is far more powerful and infinitely flexible because data analysis can be programmed from scratch or reuse existing code or extensions in the form of R packages.

Cleaning and preparing the survey data for analysis was done first and involved tasks such as combining the data from each survey version, checking the survey responses against the attention check questions, and removing all of a participants' response data if one of their responses was not valid. Due to computational processing constraints, some of the variables had to be recoded, or, renamed to numeric values but this did not alter the meaning of data. For example, the 7-point ordinal Likert-type response scale ranging from *Strongly disagree* to *Strongly agree* was recoded to range from 1 to 7 so that statistical analysis could be performed.

### 4.3.1. Sample distribution visualizations and descriptive statistics

The initial stage of data analysis is exploratory in nature. The entirety of the cleaned survey data is selected and grouped, or, 'sliced' according to variables of interest, such as the three conditions, two generated novels, and three skill groups. This means that when each survey question is analyzed it can be sliced according to each of the groups. Each group is visualized as a bar chart so that a quick inspection by eye can understand the spread of data and identify any features which might be relevant to investigate further. To support this investigation and further identify relevant features, two common descriptive statistics which describe and summarize the spread of the data are calculated for each group: the Mean (an average measure) and the Standard Deviation (a measure of deviation from the Mean). When interpreting these values it's crucial to keep in mind that they're describing many, averaged responses on a 7-point ordinal scale which does not have decimal values or, 'in-between' response options. For example, a Mean of 4.5 represents an average response theoretically situated somewhere between *Neither*

[28] https://www.r-project.org/ . R 4.3 is the language version used.
[29] While the code repository is on Github and can be requested for viewing, it strictly contains no research data: https://github.com/dx9240/2024_survey_analysis

*agree nor disagree* (4) and *Somewhat agree* (5), although it is of course understood that no participant was actually able to respond between these two values. Further, if this example also had a Standard Deviation of ± 1.2, then that indicates that the average response fell in a rough theoretical range from *Somewhat disagree* to *agree*. Indeed, while potentially broad individual results may not be very meaningful to answering the research questions on their own, strong patterns which may emerge from these results may be more meaningful in the real-world context.

### 4.3.2. Inferential statistics

Hypotheses which arise from the exploratory analysis stage are tested for statistically significant differences between two or more groups using inferential statistics. The majority of significance testing carried out in this study is the Two-Sample T-Test, which is commonly used to compare the Means of two groups. The Kruskal-Wallis test followed by Dunn's test is also used in order to compare the Means of three groups. Both tests calculate a p-value statistic which is the probability that the null-hypothesis (the hypothesis that there is no difference between two groups) is false and can be rejected. While 0.05 is a commonly used p-value threshold, this study uses a p-value of less than or equal to 0.01[30] (at most a null hypothesis probability of 0.1%) in order to reduce the risk of making type I errors, or, false positives. Similarly, a Bonferroni correction is used after Dunn's test in order to further reduce the risk of type I errors because running several tests can increase the likelihood of producing false positives. Other values are also reported for the T-Test: the t-value (significant differences are more unlikely the closer this value is to zero), Degrees of Freedom (DF. the sample size minus the two parameters in the "Two-Sample" T-Test), and the 95% Confidence Interval (CI. The difference between the two means being compared falls within this range).

The statistical tests are of course only able to evidence whether the null-hypothesis should be accepted or rejected; they cannot be used to directly accept a hypothesis based on one of the research questions. Therefore, several tests are run and their results are discussed in the real-world context and with qualitative data in order to collect evidence to answer the research questions. Thus, with a conservative statistical stance taken against committing type I errors, if a result in this study is significant then it indicates that the evidence for rejecting the null hypothesis is very strong.

An important methodological consideration is this study's choice to use the parametric Two-Sample T-Test on non-parametric Likert-type ordinal data. The sample distribution of

[30] p-value ≤ 0.01

parametric data conforms to fixed parameters. A bell curve shape indicating normal distribution in a data visualization is an example of parametric data because the symmetrical shape represents the areas in which predictable parameters are situated, such as the mean and standard deviation. Non-parametric data on the other hand is not normally distributed, so the parameters are not predictable in the same way. Data from Likert responses is not expected to be normally distributed because it is ordinal, and, intuitively, because Likert questions are designed to collect data about people's opinions or perceptions and so the mean responses aren't necessarily going to be at the central point of the scale. Nevertheless, parametric tests have been proven to be extremely robust, or, reliable even when assumptions are violated, such as using non-parametric ordinal Likert response data instead of the assumed parametric data. "...various distributional assumptions or the use of parametric statistics with ordinal data, may be strictly true, but fail to account for the robustness of parametric tests, and ignore a substantial literature suggesting that parametric statistics are perfectly appropriate" (Norman, 2010, p. 626). Norman emphatically argues that "...parametric methods are incredibly versatile, powerful and comprehensive" and are therefore preferable to non-parametric methods (p. 627). For this reason, the majority of tests for significant differences between conditions, generated novels, and skill groups in this study use the Two-Sample T-Test.

De Winter and Dodou (2010) supports the T-Test choice. The paper details a simulation study on 5-point Likert scale items to compare the test results of the T-Test and its non-parametric equivalent, the Mann-Whitney-Wilcoxon test. The study concludes that t-test can be used with 5-point Likert scale items. Norman's rigorous arguments that "Parametric statistics can be used with Likert data, with small sample sizes, with unequal variances, and with non-normal distributions, with no fear of 'coming to the wrong conclusion'. These findings are consistent with empirical literature dating back nearly 80 years." (2010, p. 631). A final comment on this study's choice of T-Test: The sample sizes compared in this study are of unequal sizes but no corrections, such as the Welch's Unequal Variances T-Test, are conducted because de Winter and Dodou have shown this to be unnecessary through verified testing. The authors conclude that "...for Likert data, the regular *t* test is to be recommended over the unequal variances *t* test." (2010, p. 6). Therefore the Two-Sample T-Test used in this study is suitable for the collected data.

## 4.4. Quantitative results and analysis overview

The quantitative analysis is driven by an investigation to better understand how paratext might function to influence reader reception under operationalized, experiment conditions[31], so that

[31] Or, as close to experiment conditions as is possible online.

the previously outlined arguments put forward by Genette and media researchers can be critically re-examined and discussed. Due to the quantitative nature of the data, it is possible to reason about how a generated text and the paratextual pieces impact readers' reception. This is expected to differ depending on a reader's experiences and skills.

The initial stages of quantitative data analysis focus on the survey question data from the Q1 and groups. These asked participants to respond to statements about literary, creative, and technical value ("The project is literary" - A1, "The project is creative" - A2, and "The project is technical" - A3). The response data from three questions belonging to group Q2 were also analyzed to understand participant responses based on overall quality and their understanding of what they read, as captured by 'enjoyable', 'interesting' ("The project is enjoyable" – A6, "The project is interesting" – A5), and 'understanding' ("The project is understandable" – A12). These questions are split and analyzed across the three paratextual conditions shown to participants (the *Generated Text Condition* – piece A generated text, the *Paratext Condition* - all other project pieces from B to G, and the *Both Condition* – all project pieces from A to G), the two novels (*Novel B* - Bhatnagar's *Molly's Feed* , and *Novel S* - Shane's *Victorian* - novel S), and the three skills (*Skill Code* - coding skills, *Skill Code & Lit* - both coding and literary reading skills, and *Skill Lit* - literary reading skills).

The remaining questions in the *Q2* question group were not analyzed beyond an exploratory first pass. They will not be discussed further in this dissertation because the scope of the analysis has become more focused; the responses to Q1 Literary were unexpected, and so an investigatory path through *Q4 Hypotext Familiarity*, *Q71 Ranking Understanding and Q72 Ranking Interesting* was taken in order to investigate and better understand the interesting results of Q1 Literary. Similarly, all other survey questions will not be discussed further.

In order to begin answering the research questions, potential differences between the paratextual conditions, generated novels, and skills groups were explored. This was done by visualizing the data to see the distribution of responses, and by calculating descriptive statistics to further understand the spread of data. T Two-Sample T-Test or the Kruskal-Wallis test was used to test the groups for significance.

### 4.4.1. Analysis narrative outline

The analysis of the survey questions data begins by focusing on the impact that paratext has on readers' perception of literary, creative, and technical value. As well as the influence of paratext on perceived enjoyability, interest, and understandability. The results show that the presence of the paratext in the project does make a difference across all of these values – except for literary.

It's possible to measure the impact of the paratext by statistically comparing the *Generated Text Condition* and the *Both Condition*, because the only difference between them is the absence and the presence of the other paratextual pieces. Put more formally, where *Both* represents the whole project with all the project pieces, *G* the generated text, and *P* the other paratextual pieces:

$$Both(P + G) – G = P$$

Next, the analysis is performed again to see the impact that the generated text has on readers' perceptions. But unexpectedly, the absence of the generated text from the project doesn't make a difference to reader's value judgements – it doesn't appear to have any impact.

Measuring the impact of the generated text is possible by statistically comparing the *Generated Text Condition* and the *Paratext Condition*, because the difference between then is the absence and the presence of the generated text sample. Put more formally,

$$Both(P+G) – P = G$$

Because both projects have a hypotextual relationship with an esteemed literary work, it's unexpected that perceptions of literariness aren't influenced by the presence of the paratext. Could this be because there is a group of readers who are not familiar with the specific literary works? And further, is it possible that being familiar with the works leads to seeing the generated novel project as being more literary? Surprisingly, no! Perceptions of literary value remain almost immutable.

Flummoxed by the hypotext results the analysis presses on; how important might the other paratextual pieces or their function be to readers? It appears that the pieces which work to explain why the generated novels exists, from the NaNoGenMo challenge page detailing the aim, the press article giving a readers' review, and the creator's progress page are the most important for helping readers to understand the project. Notably, although the generated text isn't considered to be important for understanding, readers prioritize it for helping to make the project interesting. Meanwhile, readers find the more technical of the paratextual pieces, the code, the code repository, and the input data to be the least important overall.

### 4.4.2. The impact of adding paratext: the Generated Text Condition vs. the Both Condition

The data analyzed in this section is the *Generated Text Condition* and the *Both Condition* in survey questions *Q1 Literary*, *Creative*, *Technical*, and *Q2 Enjoyable*, *Interesting*, and *Understandabl*e. These conditions are compared to each other because they represent reader reception when the generated text is presented on its own, versus the impact that it has on

reception when the project is presented as a whole. Thus, the two conditions can be interpreted as measuring the impact of the paratextual pieces. Put more simply: what impact does paratext have when it is added to the generated text?

If there are statistically significant differences between the *Generated Text Condition* and the *Paratext Condition*, then this would be evidence that paratext has an impact on the value dimensions which are measured. In other words, if there is a difference in responses between participants who were shown all the project pieces along with the generated text, and those who were only shown the generated text in isolation, then this would be evidence of paratext's impact on reader reception within the study.

#### 4.4.2.1. Sample distributions

Q1 *Literary*, *Creative*, *Technical*, and Q2 *Enjoyable*, *Interesting*, and *Understandable* were visualized as bar plots in order to investigate the spread of responses. The data was split and visualized across the paratextual conditions, novels, and skills. This is seen *in Figure 4: Literary, Figure 5: Creative, Figure 6: Technical, Figure 7: Enjoyable, Figure 8: Interesting, and Figure 9: Understandable*. In these visualizations, looking vertically the *Generated Text Condition* is represented by the value 1 (the left-most bar charts in each visualization), the *Paratext Condition* is represented by the value 2 (the middle bar charts), and the *Both* condition is represented by the value 3 (the right-most bar charts). Looking horizontally, each visualization represents the conditions in the upper bar charts, the two generated novel groups in the middle bar charts, and three skills groups in the lower bar charts. The legend shows the likert-type responses represented as ordinal values from 1-7. As the most negative value, 1 (seen in red) represents the ordinal response *Strongly disagree*, with 7 (seen in pink) representing *Strongly agree* - the most positive ordinal response.

Figure 4: Literary

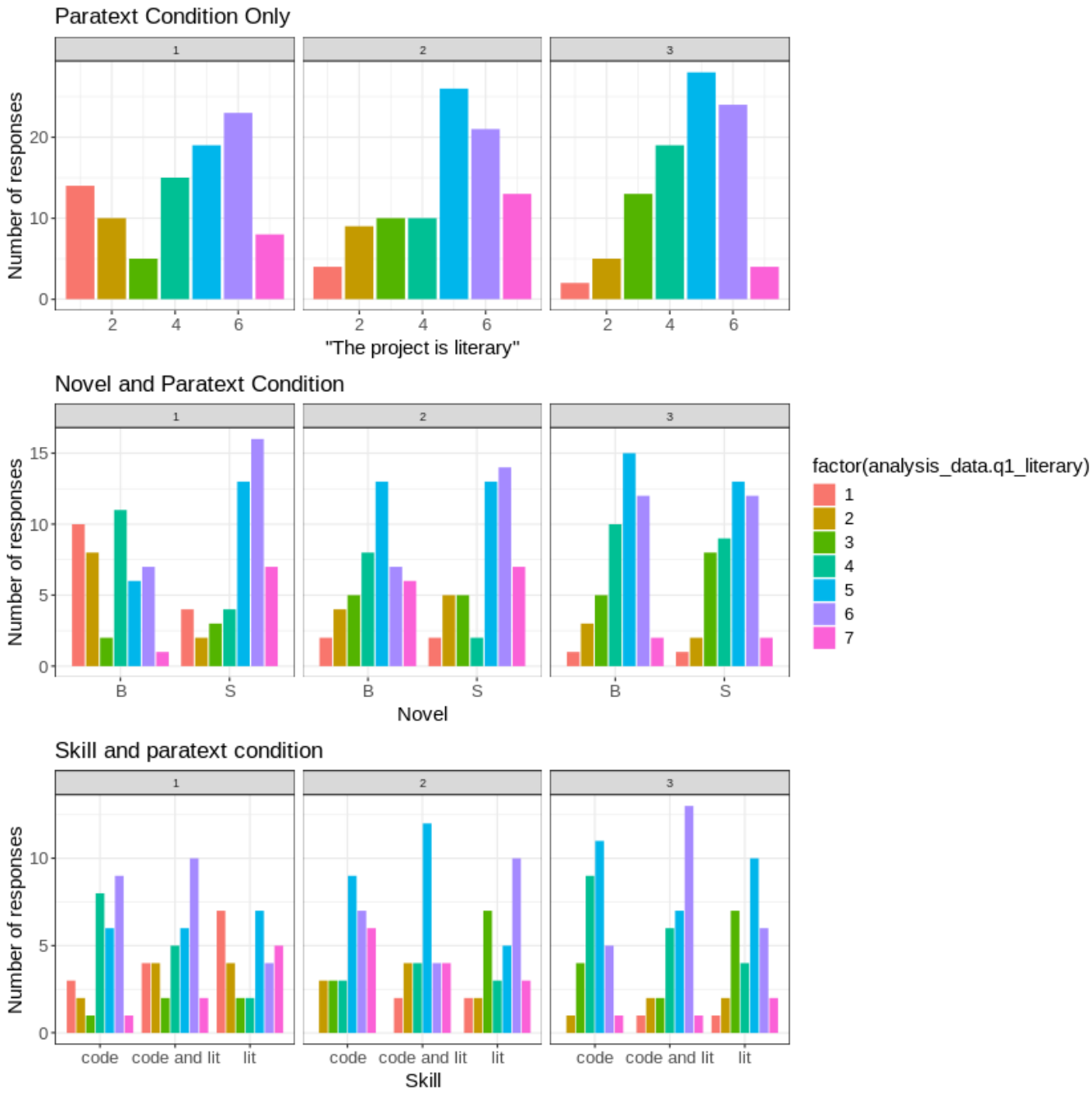


Figure 5: Creative

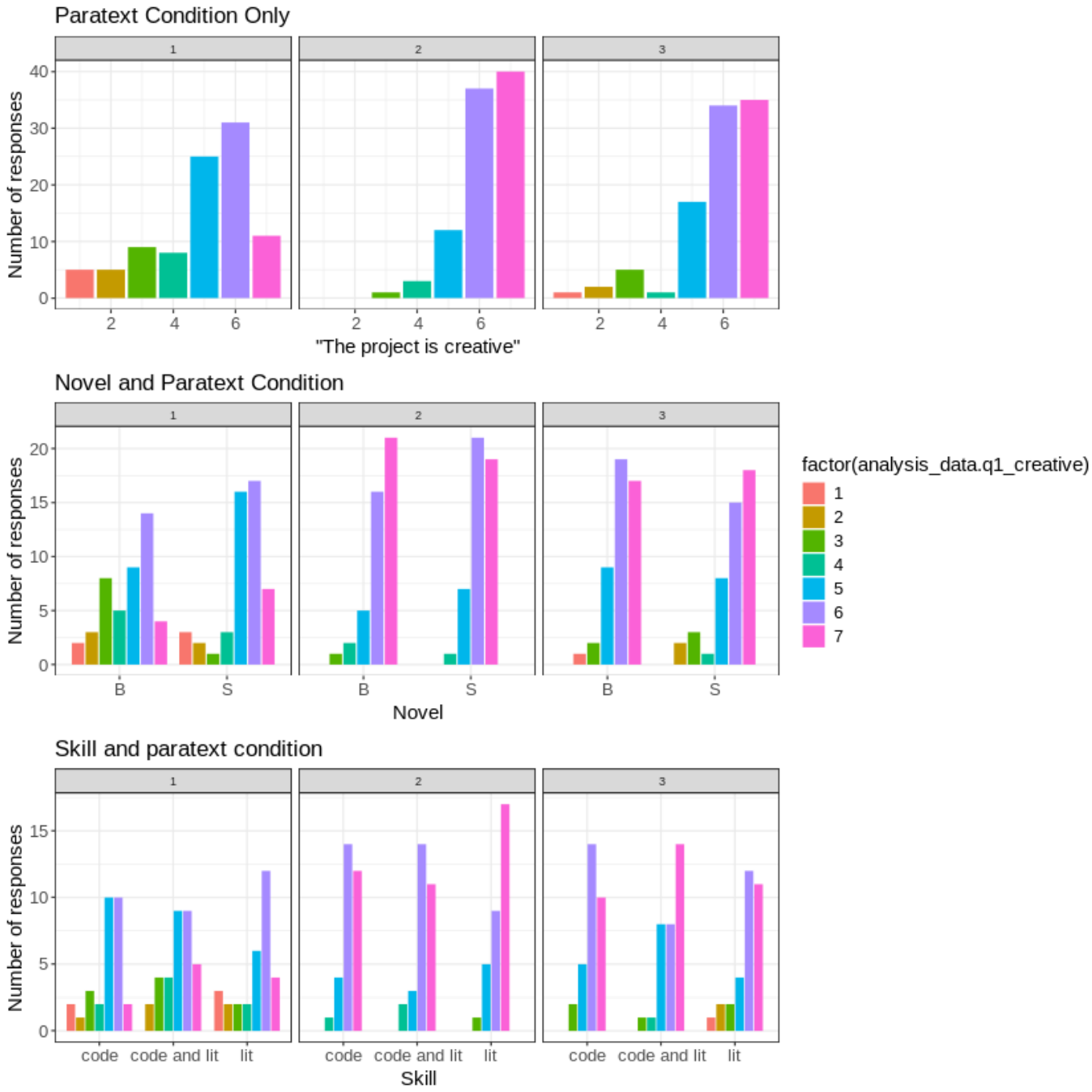


Figure 6: Technical

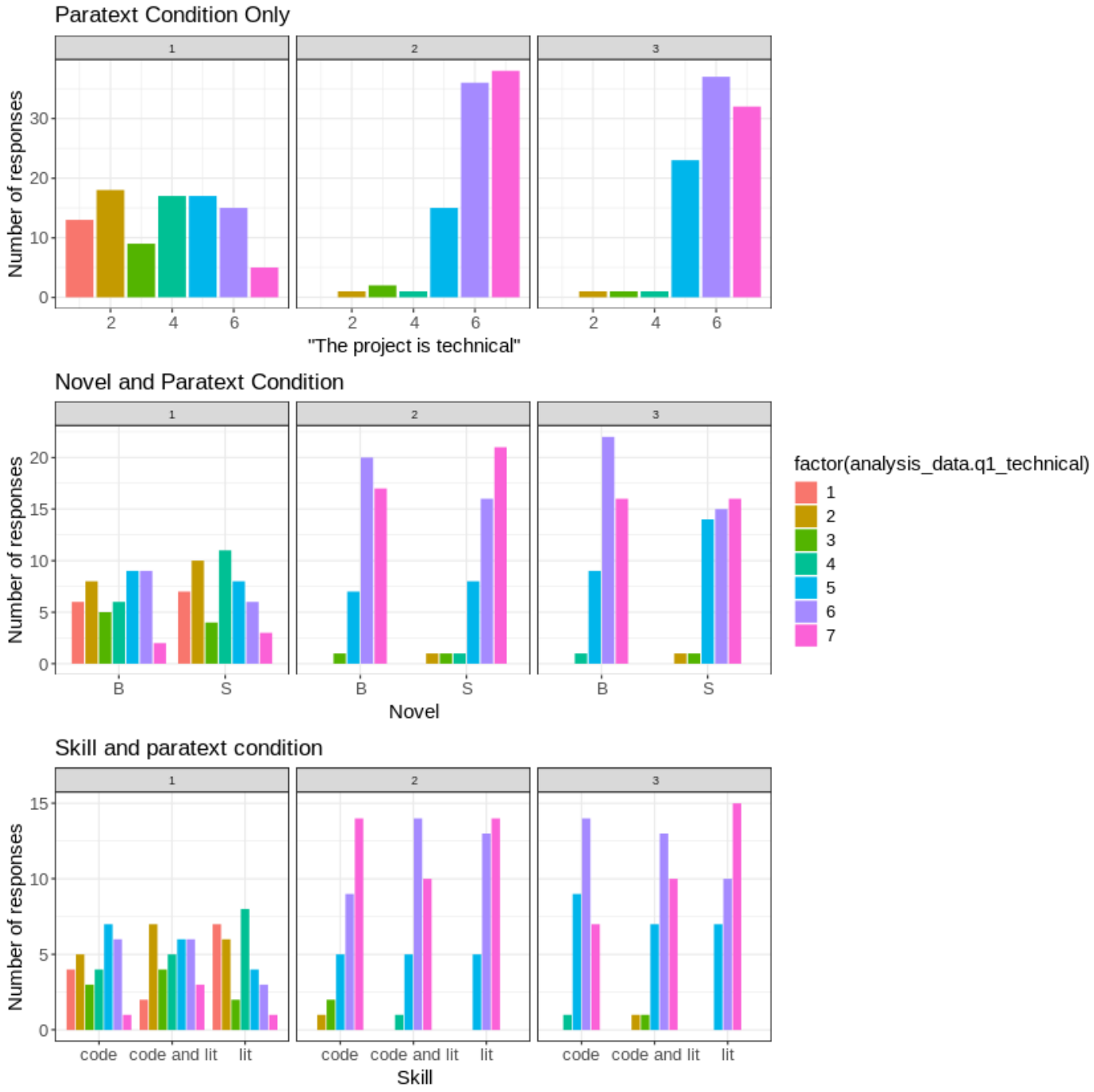


Figure 7: Enjoyable

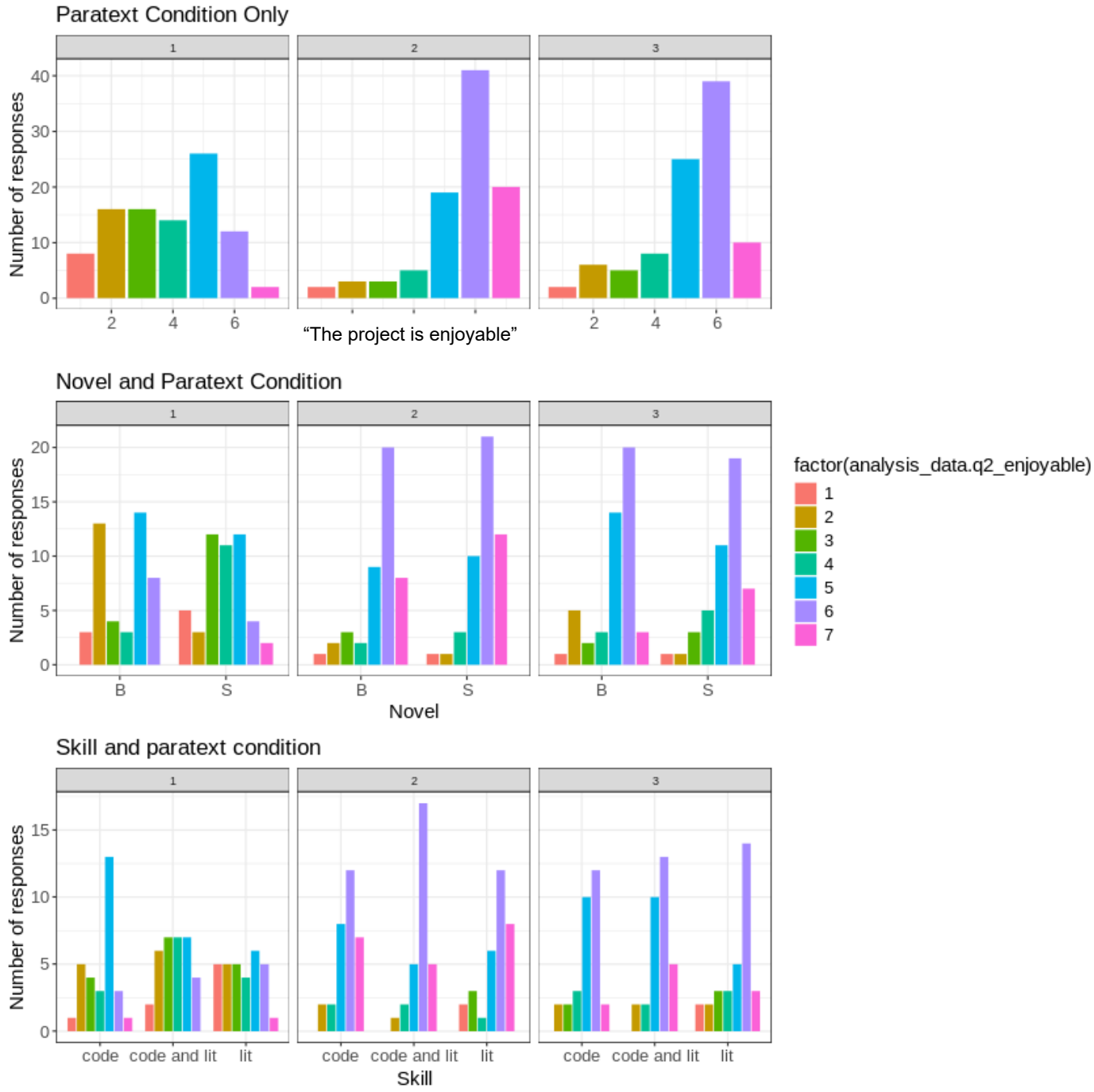


Figure 8: Interesting

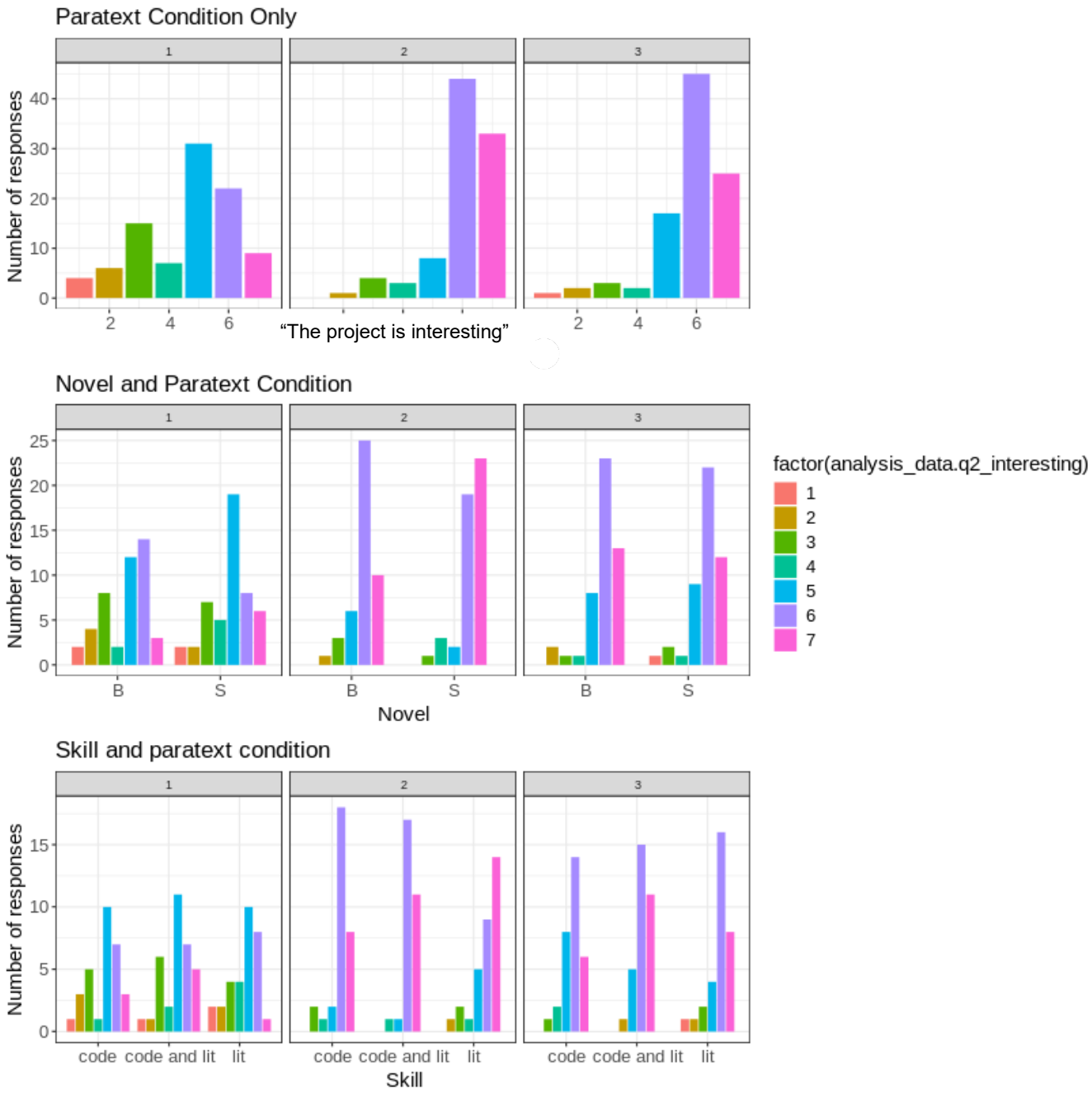
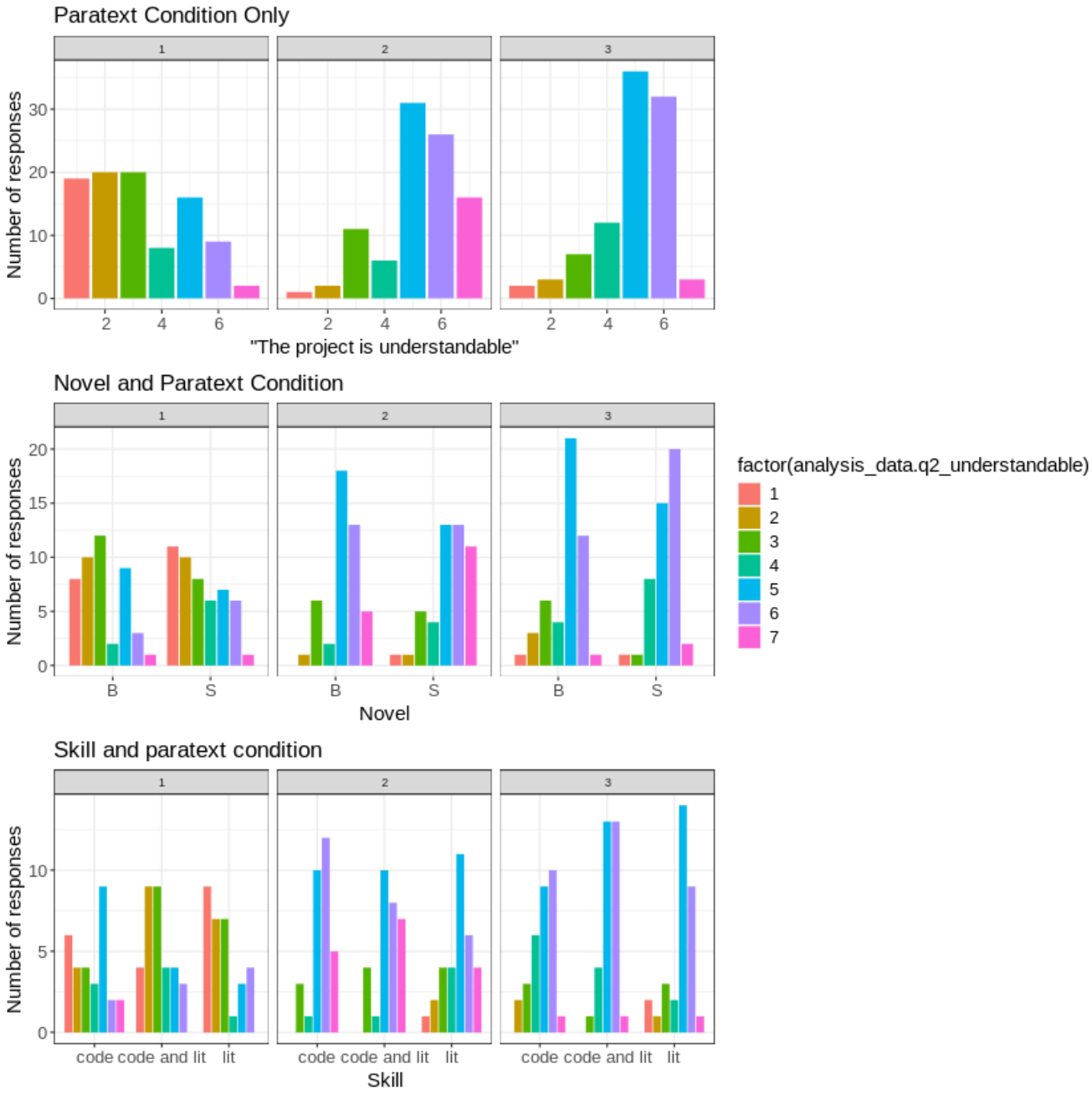


Figure 9: Understandable



As discussed in section *4.3.1. Inferential statistics* normal distribution within the sample was not expected and this is reflected in the vast majority of the visualized survey questions where a range of bimodal, uniform, and off-center distributions are seen. Further, with the exception of *Literary*, distributions appear to culminate towards the positive pole of the plots in the *Both Condition*, but appear to show a greater spread of responses in the *Generated Text Condition*. This may suggest that there is greater agreement in readers' responses when all paratextual elements are present in the reading study. However, this apparent shift towards greater participant agreement is not seen in the distribution of responses for *Literary*.

*Literary* contains groups appearing to show a range of generally bimodal and flatter distributions in the *Generated Text Condition*, the *Generated Text Condition Novel B*, and the *Generated Text Condition Lit Skill* [literary reading skill]). *Literary* also contains groups whose somewhat off-center distribution appears to be characterized by a more pronounced flatness when compared to the off-center distributions in other survey questions. This can be seen, for example, when *Literary* is analyzed per novel across the conditions. In the *Generated Text Condition*, *Novel B*'s bimodal distribution is seen alongside *Novel S*'s which instead lean towards the positive pole. In the *Both Condition*, however, the distributions for both novels appears to 'flatten out', with *Novel B* then appearing off-set towards the positive pole in the *Both Condition*. Thus, the distributions for both generated novels look similar to each other in this condition, which suggests that there may be little difference in responses here. This is in contrast the novels' distributions in the *Generated Text Condition*. In other words, while the distribution of responses in the *Generated Text Condition Novel B* appears to elicit more negative opinion to the statement "The project is literary" than *Novel S* in the same condition does, the difference in value judgements between the literariness of the two generated novels appears to disappear in the *Both Condition*.

In contrast to *Literary*, the other survey questions appear to have clearer differences between their groups. The distribution of responses in *Interesting* ("The project is interesting") is an example of what is generally seen in these other questions; the *Generated Text Condition* shows a somewhat flatter spread which suggests a mixed range of responses. This then gives way to an off-center appearance in the *Both Condition* where the generated novels in *Interesting* clearly lean heavily to the positive pole, suggesting that there may be little difference in judgements of interestingness between the two projects. Indeed, unlike the other survey questions whose distribution plots appear to have clearer differences between the *Generated Text Condition* and the *Both Condition* (as seen in *Interesting*, for example), it is challenging to visually identify clear potential differences between *Literary* conditions and groups. This may suggest that the concept of literariness might be somewhat resistant to the differences between the two conditions, where perhaps neither the generated text nor the paratext strongly impacts reader's

value judgements of literariness. Descriptive statistics were calculated to investigate this further.

#### 4.4.2.2. Descriptive statistics

Generally, the results of the descriptive statistics corroborate the visual reading of the spread of data. *Table 3: Descriptive statistics (Generated Text and Both Conditions)* shows that the mean tends to increase from the *Generated Text Condition* to the *Both Condition*, which appears to indicate a general increase in positive responses when the generated text is presented together with all paratextual pieces. This echoes the lean towards the positive polarity of the plot seen in the *Both Condition* of the visualized distributions. For example, in *Table 3* overall the participants who read the *Novel B* sample in the *Generated Text Condition* responded to "The project is Literary" with a mean of about 3.44, or, between the ordinal scale values *somewhat disagree* and *neither agree nor disagree* (Mean = 3.444 ± SD = 1.865). In the *Both Condition* the mean increased to about 4.65, towards *somewhat agree* (Mean = 4.646 ± SD = 1.36). This means that there's a general trend where the positivity of value judgements increases when the paratext is included with the generated text. But as can be seen in the example, this isn't always a very meaningful increase in a real-world context.

Table 3: Descriptive statistics (Generated Text and Both Conditions)

| Question | Condition | Group | Mean | Standard Deviation |
|---|---|---|---|---|
| Q1 Literary (A1) | Generated Text | | 4.234042553 | 1.93677224 |
| Q1 Literary (A1) | Generated Text | Novel S | 4.959183674 | 1.71948616 |
| Q1 Literary (A1) | Generated Text | Novel B | 3.444444444 | 1.865421663 |
| Q1 Literary (A1) | Generated Text | Skill Code | 4.433333333 | 1.675036455 |
| Q1 Literary (A1) | Generated Text | Skill Lit | 3.967741936 | 2.228360697 |
| Q1 Literary (A1) | Generated Text | Skill Code & Lit | 4.303030303 | 1.895469079 |
| Q1 Literary (A1) | Both | | 4.621052632 | 1.36193657 |
| Q1 Literary (A1) | Both | Novel S | 4.595744681 | 1.377767168 |
| Q1 Literary (A1) | Both | Novel B | 4.645833333 | 1.360375157 |
| Q1 Literary (A1) | Both | Skill Code | 4.580645161 | 1.118755069 |
| Q1 Literary (A1) | Both | Skill Lit | 4.4375 | 1.501343484 |
| Q1 Literary (A1) | Both | Skill Code & Lit | 4.84375 | 1.439295863 |
| Q1 Creative (A2) | Generated Text | | 4.914893617 | 1.617508867 |
| Q1 Creative (A2) | Generated Text | Novel S | 5.163265306 | 1.559205711 |
| Q1 Creative (A2) | Generated Text | Novel B | 4.644444444 | 1.653585025 |
| Q1 Creative (A2) | Generated Text | Skill Code | 4.833333333 | 1.577499841 |
| Q1 Creative (A2) | Generated Text | Skill Lit | 4.870967742 | 1.85727283 |
| Q1 Creative (A2) | Generated Text | Skill Code & Lit | 5.03030303 | 1.446652353 |
| Q1 Creative (A2) | Both | | 5.873684211 | 1.290242475 |
| Q1 Creative (A2) | Both | Novel S | 5.808510638 | 1.377431414 |
| Q1 Creative (A2) | Both | Novel B | 5.9375 | 1.209998242 |
| Q1 Creative (A2) | Both | Skill Code | 5.967741936 | 1.048296109 |
| Q1 Creative (A2) | Both | Skill Lit | 5.625 | 1.660742158 |
| Q1 Creative (A2) | Both | Skill Code & Lit | 6.03125 | 1.062084834 |
| Q1 Technical (A3) | Generated Text | | 3.765957447 | 1.839971551 |
| Q1 Technical (A3) | Generated Text | Novel S | 3.673469388 | 1.830161139 |
| Q1 Technical (A3) | Generated Text | Novel B | 3.866666667 | 1.865963071 |
| Q1 Technical (A3) | Generated Text | Skill Code | 3.9 | 1.8448437 |
| Q1 Technical (A3) | Generated Text | Skill Lit | 3.290322581 | 1.81095828 |
| Q1 Technical (A3) | Generated Text | Skill Code & Lit | 4.090909091 | 1.826260461 |
| Q1 Technical (A3) | Both | | 6 | 0.945313187 |
| Q1 Technical (A3) | Both | Novel S | 5.893617021 | 1.088158447 |
| Q1 Technical (A3) | Both | Novel B | 6.104166667 | 0.778421297 |
| Q1 Technical (A3) | Both | Skill Code | 5.870967742 | 0.805892279 |
| Q1 Technical (A3) | Both | Skill Lit | 6.25 | 0.803219329 |
| Q1 Technical (A3) | Both | Skill Code & Lit | 5.875 | 1.157026222 |
| Q2 Enjoyable (A6) | Generated Text | | 3.829787234 | 1.610705435 |
| Q2 Enjoyable (A6) | Generated Text | Novel S | 3.857142857 | 1.541103501 |
| Q2 Enjoyable (A6) | Generated Text | Novel B | 3.8 | 1.700267359 |
| Q2 Enjoyable (A6) | Generated Text | Skill Code | 4.166666667 | 1.510499651 |
| Q2 Enjoyable (A6) | Generated Text | Skill Lit | 3.64516129 | 1.835726688 |
| Q2 Enjoyable (A6) | Generated Text | Skill Code & Lit | 3.696969697 | 1.468095035 |
| Q2 Enjoyable (A6) | Both | | 5.157894737 | 1.424077077 |
| Q2 Enjoyable (A6) | Both | Novel S | 5.319148936 | 1.353038175 |
| Q2 Enjoyable (A6) | Both | Novel B | 5 | 1.487536876 |
| Q2 Enjoyable (A6) | Both | Skill Code | 5.096774194 | 1.274227575 |
| Q2 Enjoyable (A6) | Both | Skill Lit | 4.90625 | 1.710675682 |
| Q2 Enjoyable (A6) | Both | Skill Code & Lit | 5.46875 | 1.217728584 |
| Q2 Interesting (A5) | Generated Text | | 4.670212766 | 1.582079059 |
| Q2 Interesting (A5) | Generated Text | Novel S | 4.734693878 | 1.524460427 |
| Q2 Interesting (A5) | Generated Text | Novel B | 4.6 | 1.656941322 |
| Q2 Interesting (A5) | Generated Text | Skill Code | 4.633333333 | 1.650148025 |
| Q2 Interesting (A5) | Generated Text | Skill Lit | 4.483870968 | 1.567821583 |
| Q2 Interesting (A5) | Generated Text | Skill Code & Lit | 4.878787879 | 1.556389567 |
| Q2 Interesting (A5) | Both | | 5.810526316 | 1.17866438 |
| Q2 Interesting (A5) | Both | Novel S | 5.787234043 | 1.196665116 |
| Q2 Interesting (A5) | Both | Novel B | 5.833333333 | 1.172981895 |
| Q2 Interesting (A5) | Both | Skill Code | 5.709677419 | 0.972746914 |
| Q2 Interesting (A5) | Both | Skill Lit | 5.65625 | 1.472530738 |
| Q2 Interesting (A5) | Both | Skill Code & Lit | 6.0625 | 1.014014697 |
| Q2 Understandable (A12) | Generated Text | | 3.180851064 | 1.728712405 |
| Q2 Understandable (A12) | Generated Text | Novel S | 3.204081633 | 1.80253978 |
| Q2 Understandable (A12) | Generated Text | Novel B | 3.155555556 | 1.664544103 |
| Q2 Understandable (A12) | Generated Text | Skill Code | 3.633333333 | 1.90250893 |
| Q2 Understandable (A12) | Generated Text | Skill Lit | 2.806451613 | 1.74010258 |
| Q2 Understandable (A12) | Generated Text | Skill Code & Lit | 3.121212121 | 1.494940964 |
| Q2 Understandable (A12) | Both | | 4.947368421 | 1.232265006 |
| Q2 Understandable (A12) | Both | Novel S | 5.212765957 | 1.082191219 |
| Q2 Understandable (A12) | Both | Novel B | 4.6875 | 1.323378172 |
| Q2 Understandable (A12) | Both | Skill Code | 4.806451613 | 1.275914178 |
| Q2 Understandable (A12) | Both | Skill Lit | 4.75 | 1.459120371 |
| Q2 Understandable (A12) | Both | Skill Code & Lit | 5.28125 | 0.851350919 |

#### 4.4.2.3. Inferential statistics

To statistically test differences between groups, A T-Test was performed on each of the groups within the survey questions to compare the means in the *Generated Text Condition* and the *Both Condition*. Because this analysis is running a large number of tests, $p \leq 0.01$ was the threshold used to determine whether to accept or reject the null hypothesis. This small p-value reduces the chance of type I errors (false positives).

A Two-Sample T-Test was run to determine if there were differences in the mean responses between question groups in the *Generated Text Condition* and the *Both Condition*. With some exceptions which are detailed below, the data in *Table 5: Inferential statistics (Generated Text and Both Conditions)* shows a pattern where the differences between the two conditions are significant. However, an intriguing exception to this are the results for *Literary*; it is the only one of the six survey questions which has barely any significant differences between the two conditions. This therefore confirms the exploratory analysis from the previous subsections – that within the quantitative sphere of this study, the responses to the statement "The project is literary" are overall quite resistant to being impacted by the paratextual pieces. This is largely in contrast to the differences between the *Generated Text Condition* and the *Both Condition* for *Creative*, *Technical*, *Interesting*, *Enjoyable*, *Understanding*.

Of course, within the *Literary* exception to the pattern lies another interesting exception – Bhatnagar's *Molly's Feed*. *Novel B* is the only group in *Literary* that has a significant difference between the *Generated Text Condition* (Mean = 3.44 ± SD = 1.865) and the *Both Condition* (Mean= 4.65 ± SD = 1.36). Statistically speaking, this difference in responses is roughly between *Somewhat disagree* and *Neither agree nor disagree* (95% CI -1.87 to -0.53), t(DF) = 91, p = 0). Practically speaking of course, participants were only able to respond to the 7 ordinal responses and not a value between them. Although the difference in responses to *Novel B* is a small change from mildly negative responses in the *Generated Text Condition*, to neutral or unsure responses in the *Both Condition*, based on the low p-value the difference is a very clear one. This difference in responses is based on the inclusion of the paratextual pieces in the *Both Condition*, and it's striking that a difference isn't also seen with *Novel S*. This could be because participants might have differing levels of familiarity with the hypotext (the esteemed literary novels which inspired the generated novels). This will be investigated through the *Q4 Hypotext Familiarity* data in section *4.4.4. Familiarity with the hypotext*.

Two groups within *Creative* had insignificant differences between the conditions, but nevertheless, overall the survey question arguably fits well with the pattern. *Novel S* could not be shown to have a significant difference based on this study's p-value threshold. Intuitively, though, this result does not seem unusual because the mean response of 5.16 in the *Generated*

*Text Condition* is the highest in that question condition, and it's already within the range of positive responses of the 7-point ordinal scale (Mean = 5.16 ± SD = 1.559). The insignificant result for the *Skill Lit* group is more challenging to explain because the quantitative data alone doesn't offer any reasons as to why it is the only skill group in *Creative* without a significant increase in positive responses in the *Both Condition*, especially considering that the participants in the *Skill Code & Lit* group (both literary reading and coding skills) are presumed to have equivalent literary reading skills. Indeed, all three skill groups share very similar means in the *Generated Text Condition*. It is possible that some or all of the additional project pieces in the *Both Condition* have a greater impact on perceptions of creative value for those participants who reported having coding skills, which might account for *Skill Code* and *Skill Code &* Lit having significant differences, and not *Skill Lit*. Analysis of the written reviews may be able to suggest whether there is any basis for this line of reasoning. Finally, it is arguably surprising that the data for *Creativity* and *Literary* are not more similar to each other. Both are complex, and potentially polarizing or even controversial concepts especially a digital context. Analysis of the written reviews may be able to suggest why.

For the remaining survey questions the pattern holds for each group: there is a significant increase in almost all the mean responses between the *Generated Text Condition* and the *Both Condition*.

Question *Enjoyable* was included in the survey as a measure of overall value, and its results clearly follow the pattern where the mean response increases from one condition to the other. In the *Generated Text Condition*, the mean response ranges from about 3.4 to about 4.1 (*Skill Lit,* Mean = 3.65 ± SD = 1.835) (*Skill Code*, Mean = 4.17 ± SD = 1.51). This shows that, from a statistical analysis perspective, most participants *Somewhat disagree* to *Neither agree nor disagree* with the statement "The project is enjoyable". Thus, for most participants engaging with just the generated text sample on its own for the first time, the experience was perhaps not an enjoyable one. However, in the *Both Condition* the mean responses range from 4.9 to about 5.5 (*Skill Lit*, Mean = 4.9 ± SD = 1.71) (*Skill Code & Lit*, Mean = 5.47 ± SD = 1.218). This shows that readers' judgements increase to be within the *Somewhat agree* point on the 7-point ordinal scale when the generated text is presented together with the paratext in this study. This is some indication that in general, a generated novel might be more enjoyable for new readers to engage with when some or of all the project pieces are presented together[32].

[32] While this may be the typical way that generated novels are presented, it is not at all unusual for Literature classrooms to present texts to readers with very little, if any, paratextual information. Of course, the enjoyability of the text may not be a priority in that context.

The difference in the increase of mean responses for *Technical* ("The project is Technical") from one condition to the other is high. The mean responses across the groups in the *Generated Text Condition* range from about 3.3 to about 4, or, between *Somewhat disagree* and *Neither agree nor disagree* (*Skill Lit*, Mean = 3.29 ± SD = 1.81) (*Skill Code & Lit*, Mean = 4.09 ± SD = 1.826). This range across the *Technical* groups jumps to between about 5.9 to about 6.3 in the *Both Condition*, which is within the *agree* response on the ordinal scale (*Skill Code,* Mean = 5.87 ± SD = 0.8) (*Skill Lit*, Mean = 6.25 ± SD = 0.8). Statistically speaking, this is an increase of about 2 points on the 7-point ordinal scale. While the standard deviation in the *Generated Text Condition* is high, it decreased by approximately an entire point on the ordinal scale in the *Both Condition*. This indicates that participants are in higher agreement with each other when presented with the entire project, rather than just the generated text sample. A higher increase in the mean response from one condition to the other across the *Technical* question groups makes sense intuitively, because the paratextual pieces in the *Both Condition* draw attention to the project's digital and algorithmic aspects, as well as the GitHub platform itself. Therefore, they're likely to fit comfortably with readers' concept of technicalness.

Table 4: Inferential statistics (Generated Text and Both Conditions)

| Question | First Condition | Second Condition | Group | T | DF | P | 95% Confidence Interval | Mean of X (First Condition) | Mean of Y (Second Condition) | Null Hypothesis | Significance |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Q1 Literary | Generated Text | Both | | -1.5904 | 187 | 0.1134 | -0.86706379 0.09304363 | 4.234043 | 4.621053 | Accept | |
| Q1 Literary | Generated Text | Both | Novel B | -3.5645 | 91 | 0.00058 | -1.8708770 -0.5319007 | 3.444444 | 4.645833 | Reject | Significant |
| Q1 Literary | Generated Text | Both | Novel S | 1.1399 | 94 | 0.2572 | -0.2696177 0.9964957 | 4.959184 | 4.595745 | Accept | |
| Q1 Literary | Generated Text | Both | Skill Code | -0.40516 | 59 | 0.6868 | -0.8748603 0.5802367 | 4.433333 | 4.580645 | Accept | |
| Q1 Literary | Generated Text | Both | Skill Code & Lit | -1.2923 | 63 | 0.201 | -1.3768596 0.2954202 | 4.30303 | 4.84375 | Accept | |
| Q1 Literary | Generated Text | Both | Skill Lit | -0.98414 | 61 | 0.3289 | -1.4242315 0.4847154 | 3.967742 | 4.4375 | Accept | |
| Q1 Literary | Generated Text | Both | Pos Familiarity | -0.28595 | 88 | 0.7756 | -0.8641111 0.6467198 | 4.391304 | 4.5 | Accept | |
| Q1 Literary | Generated Text | Both | Neg Familiarity | -2.0107 | 84 | 0.04756 | -1.356136191 -0.007500173 | 4 | 4.681818 | Accept | |
| Q1 Creative | Generated Text | Both | | -4.5073 | 187 | 1.2E-05 | -1.378429 -0.539152 | 4.914894 | 5.873684 | Reject | Significant |
| Q1 Creative | Generated Text | Both | Novel B | -4.3227 | 91 | 3.9E-05 | -1.8872499 -0.6988612 | 4.644444 | 5.9375 | Reject | Significant |
| Q1 Creative | Generated Text | Both | Novel S | -2.1454 | 94 | 0.03449 | -1.24239527 -0.04809539 | 5.163265 | 5.808511 | Accept | |
| Q1 Creative | Generated Text | Both | Skill Code | -3.3182 | 59 | 0.00156 | -1.8185029 -0.4503143 | 4.833333 | 5.967742 | Reject | Significant |
| Q1 Creative | Generated Text | Both | Skill Code & Lit | -3.1717 | 63 | 0.00234 | -1.6316036 -0.3702904 | 5.030303 | 6.03125 | Reject | Significant |
| Q1 Creative | Generated Text | Both | Skill Lit | -1.6999 | 61 | 0.09424 | -1.6410047 0.1329402 | 4.870968 | 5.625 | Accept | |
| Q1 Technical | Generated Text | Both | | -10.515 | 187 | < 2.2e-16 | -2.653181 -1.814904 | 3.765957 | 6 | Reject | Significant |
| Q1 Technical | Generated Text | Both | Novel B | -7.6316 | 91 | 2.20E-11 | -2.819882 -1.655118 | 3.866667 | 6.104167 | Reject | Significant |
| Q1 Technical | Generated Text | Both | Novel S | -7.1861 | 94 | 1.56E-10 | -2.833577 -1.606718 | 3.673469 | 5.893617 | Reject | Significant |
| Q1 Technical | Generated Text | Both | Skill Code | -5.4375 | 59 | 1.1E-06 | -2.696276 -1.245659 | 3.9 | 5.870968 | Reject | Significant |
| Q1 Technical | Generated Text | Both | Skill Code & Lit | -4.6881 | 63 | 1.5E-05 | -2.544571 -1.023611 | 4.090909 | 5.875 | Reject | Significant |
| Q1 Technical | Generated Text | Both | Skill Lit | -8.4303 | 61 | 8.02E-12 | -3.661698 -2.257657 | 3.290323 | 6.25 | Reject | Significant |
| Q2 Interesting | Generated Text | Both | | -5.623 | 187 | 6.8E-08 | -1.5403729 -0.7402542 | 4.670213 | 5.810526 | Reject | Significant |
| Q2 Interesting | Generated Text | Both | Novel B | -4.1635 | 91 | 7.1E-05 | -1.8217549 -0.6449118 | 4.6 | 5.833333 | Reject | Significant |
| Q2 Interesting | Generated Text | Both | Novel S | -3.7524 | 94 | 0.0003 | -1.609476 -0.4956044 | 4.734694 | 5.787234 | Reject | Significant |
| Q2 Interesting | Generated Text | Both | Skill Code | -3.1156 | 59 | 0.00283 | -1.767621 -0.385067 | 4.633333 | 5.709677 | Reject | Significant |
| Q2 Interesting | Generated Text | Both | Skill Code & Lit | -3.6208 | 63 | 0.00059 | -1.8370127 -0.5304115 | 4.878788 | 6.0625 | Reject | Significant |
| Q2 Interesting | Generated Text | Both | Skill Lit | -3.0603 | 61 | 0.00328 | -1.9384108 -0.4063473 | 4.483871 | 5.65625 | Reject | Significant |
| Q2 Enjoyable | Generated Text | Both | | -6.0069 | 187 | 9.7E-09 | -1.7642696 -0.8919454 | 3.829787 | 5.157895 | Reject | Significant |
| Q2 Enjoyable | Generated Text | Both | Novel B | -3.6282 | 91 | 0.00047 | -1.856976 -0.543024 | 3.8 | 5 | Reject | Significant |
| Q2 Enjoyable | Generated Text | Both | Novel S | 4.9313 | 94 | 3.5E-06 | -2.0506675 -0.8733447 | 3.857143 | 5.319149 | Reject | Significant |
| Q2 Enjoyable | Generated Text | Both | Skill Code | -2.6027 | 59 | 0.01168 | -1.6451944 -0.2150206 | 4.166667 | 5.096774 | Reject | Significant |
| Q2 Enjoyable | Generated Text | Both | Skill Code & Lit | -5.2872 | 63 | 1.7E-06 | -2.441444 -1.102117 | 3.69697 | 5.46875 | Reject | Significant |
| Q2 Enjoyable | Generated Text | Both | Skill Lit | -2.822 | 61 | 0.00643 | -2.1546815 -0.3674959 | 3.645161 | 4.90625 | Reject | Significant |
| Q2 Understandable | Generated Text | Both | | -8.0959 | 187 | 7.09E-14 | -2.196964 -1.336071 | 3.180851 | 4.947368 | Reject | Significant |
| Q2 Understandable | Generated Text | Both | Novel B | -4.9283 | 91 | 3.70E-06 | -2.1494041 -0.9144848 | 3.155556 | 4.6875 | Reject | Significant |
| Q2 Understandable | Generated Text | Both | Novel S | -6.5849 | 94 | 2.59E-09 | -2.614353 -1.403016 | 3.204082 | 5.212766 | Reject | Significant |
| Q2 Understandable | Generated Text | Both | Skill Code | -2.837 | 59 | 0.00623 | -2.0005445 -0.3456921 | 3.633333 | 4.806452 | Reject | Significant |
| Q2 Understandable | Generated Text | Both | Skill Code & Lit | -7.1282 | 63 | 1.19E-09 | -2.765589 -1.554487 | 3.121212 | 5.28125 | Reject | Significant |
| Q2 Understandable | Generated Text | Both | Skill Lit | -4.8097 | 61 | 1.03E-05 | -2.751573 -1.135524 | 2.806452 | 4.75 | Reject | Significant |

### 4.4.3. The impact of removing the Generated Text: the Paratext Condition vs. The Both Condition

The data analyzed in this section is the *Paratext Condition* and the *Both Condition* in survey questions *Q1 Literary*, *Creative*, *Technical*, and *Q2 Enjoyable*, *Interesting*, and *Understandable*. These conditions are compared to each other because they represent the impact of removing the generated text from the project, leaving only the paratextual pieces. Thus, it can be interpreted as measuring the impact that the generated text may have on reader reception. Put more simply: what impact does the generated text remove when it is taken out of the project?

#### 4.4.3.1. Sample distributions and descriptive statistics

During the study design stage, it was assumed that the frequency of positive responses across all questions would be highest in the *Both Condition*, because all parts of the project would be available to the reader. However, this is not clearly seen based on the visual analysis of the distribution plots where it is challenging to see differences between the two conditions (seen *in Figure 4: Literary, Figure 5: Creative, Figure 6: Technical, Figure 7: Enjoyable, Figure 8: Interesting, and Figure 9: Understandable*). Indeed, *Table 2: Descriptive statistics (Paratext and Both Conditions)* shows that in the overwhelming majority of groups for each survey question, there is a small decrease in the mean responses from the *Paratext Condition* to the *Both Condition*. Bearing in mind that that these results are also accompanied by a large Standard

Deviation when considering the 7-ordinal scale, it may be the case that this difference in mean responses may be too small to detect. For example, in *Table 5*, the mean response for *Literary* in the *Both Condition* is about 4.6, with a Standard Deviation greater than a whole ordinal scale point (Mean = 4.62 ± SD = 1.361). Compared to the question's *Paratext Condition* mean response of about 4.7 and a similarly large Standard Deviation, the differences in participants value judgements between the conditions are very small, and potentially not meaningful (Mean = 4.72 ± SD = 1.683). Inferential statistics were calculated to confirm this.

Table 5: Descriptive statistics (Paratext and Both Conditions)

| Question | Condition | Group | Mean | Standard Deviation |
|---|---|---|---|---|
| Q1 Literary (A1) | Paratext | | 4.720430108 | 1.683459116 |
| Q1 Literary (A1) | Paratext | Novel S | 4.854166667 | 1.725769435 |
| Q1 Literary (A1) | Paratext | Novel B | 4.577777778 | 1.644396669 |
| Q1 Literary (A1) | Paratext | Skill Code | 5.032258065 | 1.559569833 |
| Q1 Literary (A1) | Paratext | Skill Lit | 4.53125 | 1.759478937 |
| Q1 Literary (A1) | Paratext | Skill Code & Lit | 4.6 | 1.734040528 |
| Q1 Literary (A1) | Both | | 4.621052632 | 1.36193657 |
| Q1 Literary (A1) | Both | Novel S | 4.595744681 | 1.377767168 |
| Q1 Literary (A1) | Both | Novel B | 4.645833333 | 1.360375157 |
| Q1 Literary (A1) | Both | Skill Code | 4.580645161 | 1.118755069 |
| Q1 Literary (A1) | Both | Skill Lit | 4.4375 | 1.501343484 |
| Q1 Literary (A1) | Both | Skill Code & Lit | 4.84375 | 1.439295863 |
| Q1 Creative (A2) | Paratext | | 6.204301075 | 0.866767355 |
| Q1 Creative (A2) | Paratext | Novel S | 6.208333333 | 0.77069555 |
| Q1 Creative (A2) | Paratext | Novel B | 6.2 | 0.967658843 |
| Q1 Creative (A2) | Paratext | Skill Code | 6.193548387 | 0.792437372 |
| Q1 Creative (A2) | Paratext | Skill Lit | 6.28125 | 0.958304114 |
| Q1 Creative (A2) | Paratext | Skill Code & Lit | 6.133333333 | 0.860366134 |
| Q1 Creative (A2) | Both | | 5.873684211 | 1.290242475 |
| Q1 Creative (A2) | Both | Novel S | 5.808510638 | 1.377431414 |
| Q1 Creative (A2) | Both | Novel B | 5.9375 | 1.209998242 |
| Q1 Creative (A2) | Both | Skill Code | 5.967741936 | 1.048296109 |
| Q1 Creative (A2) | Both | Skill Lit | 5.625 | 1.660742158 |
| Q1 Creative (A2) | Both | Skill Code & Lit | 6.03125 | 1.062084834 |
| Q1 Technical (A3) | Paratext | | 6.11827957 | 0.987414965 |
| Q1 Technical (A3) | Paratext | Novel S | 6.083333333 | 1.107677911 |
| Q1 Technical (A3) | Paratext | Novel B | 6.155555556 | 0.851617593 |
| Q1 Technical (A3) | Paratext | Skill Code | 5.967741936 | 1.328755582 |
| Q1 Technical (A3) | Paratext | Skill Lit | 6.28125 | 0.728868987 |
| Q1 Technical (A3) | Paratext | Skill Code & Lit | 6.1 | 0.803011573 |
| Q1 Technical (A3) | Both | | 6 | 0.945313187 |
| Q1 Technical (A3) | Both | Novel S | 5.893617021 | 1.088158447 |
| Q1 Technical (A3) | Both | Novel B | 6.104166667 | 0.778421297 |
| Q1 Technical (A3) | Both | Skill Code | 5.870967742 | 0.805892279 |
| Q1 Technical (A3) | Both | Skill Lit | 6.25 | 0.803219329 |
| Q1 Technical (A3) | Both | Skill Code & Lit | 5.875 | 1.157026222 |
| Q2 Enjoyable (A6) | Paratext | | 5.569892473 | 1.346475669 |
| Q2 Enjoyable (A6) | Paratext | Novel S | 5.729166667 | 1.233220716 |
| Q2 Enjoyable (A6) | Paratext | Novel B | 5.4 | 1.452270949 |
| Q2 Enjoyable (A6) | Paratext | Skill Code | 5.580645161 | 1.285150926 |
| Q2 Enjoyable (A6) | Paratext | Skill Lit | 5.40625 | 1.643351709 |
| Q2 Enjoyable (A6) | Paratext | Skill Code & Lit | 5.733333333 | 1.048260738 |
| Q2 Enjoyable (A6) | Both | | 5.157894737 | 1.424077077 |
| Q2 Enjoyable (A6) | Both | Novel S | 5.319148936 | 1.353038175 |
| Q2 Enjoyable (A6) | Both | Novel B | 5 | 1.487536876 |
| Q2 Enjoyable (A6) | Both | Skill Code | 5.096774194 | 1.274227575 |
| Q2 Enjoyable (A6) | Both | Skill Lit | 4.90625 | 1.710675682 |
| Q2 Enjoyable (A6) | Both | Skill Code & Lit | 5.46875 | 1.217728584 |
| Q2 Interesting (A5) | Paratext | | 6.032258065 | 1.067826041 |
| Q2 Interesting (A5) | Paratext | Novel S | 6.25 | 0.956500715 |
| Q2 Interesting (A5) | Paratext | Novel B | 5.8 | 1.140175425 |
| Q2 Interesting (A5) | Paratext | Skill Code | 5.935483871 | 1.030711207 |
| Q2 Interesting (A5) | Paratext | Skill Lit | 5.90625 | 1.352640786 |
| Q2 Interesting (A5) | Paratext | Skill Code & Lit | 6.266666667 | 0.691491807 |
| Q2 Interesting (A5) | Both | | 5.810526316 | 1.17866438 |
| Q2 Interesting (A5) | Both | Novel S | 5.787234043 | 1.196665116 |
| Q2 Interesting (A5) | Both | Novel B | 5.833333333 | 1.172981895 |
| Q2 Interesting (A5) | Both | Skill Code | 5.709677419 | 0.972746914 |
| Q2 Interesting (A5) | Both | Skill Lit | 5.65625 | 1.472530738 |
| Q2 Interesting (A5) | Both | Skill Code & Lit | 6.0625 | 1.014014697 |
| Q2 Understandable (A12) | Paratext | | 5.215053763 | 1.358143413 |
| Q2 Understandable (A12) | Paratext | Novel S | 5.291666667 | 1.472562296 |
| Q2 Understandable (A12) | Paratext | Novel B | 5.133333333 | 1.235828761 |
| Q2 Understandable (A12) | Paratext | Skill Code | 5.483870968 | 1.121634752 |
| Q2 Understandable (A12) | Paratext | Skill Lit | 4.75 | 1.545023228 |
| Q2 Understandable (A12) | Paratext | Skill Code & Lit | 5.433333333 | 1.278019301 |
| Q2 Understandable (A12) | Both | | 4.947368421 | 1.232265006 |
| Q2 Understandable (A12) | Both | Novel S | 5.212765957 | 1.082191219 |
| Q2 Understandable (A12) | Both | Novel B | 4.6875 | 1.323378172 |
| Q2 Understandable (A12) | Both | Skill Code | 4.806451613 | 1.275914178 |
| Q2 Understandable (A12) | Both | Skill Lit | 4.75 | 1.459120371 |
| Q2 Understandable (A12) | Both | Skill Code & Lit | 5.28125 | 0.851350919 |

#### 4.4.3.2. Inferential statistics

A T-Test was performed on each of the groups within each survey question to compare the means in the *Paratext Condition* and the *Both Condition*. Because this analysis is running a large number of tests, $p \leq 0.01$ was the threshold used to determine whether to accept or reject the null hypothesis. This small p-value reduces the chance of type I errors, or, false positives.

Across the large majority of groups in each question, the mean response decreases from the *Paratext Condition* to the *Both Condition*. This decrease is not significant, but it nevertheless suggests that generally, it's not the generated text which is seen with an increase in positive reception – it is the paratext. There are a handful of exceptions to this across the data as seen in table *Inferential Statistics*[33]. A Two-Sample T-Test was run to determine if there were differences in the mean responses between the two conditions. The data in *Table 8: Inferential statistics (Paratext and Both Conditions)* shows that the differences in mean responses between the *Paratext Condition* and the *Both Condition* are not significant. This result is expected based on the exploratory analysis discussed above. Interestingly, the lack of difference in participants' judgements when the generated text is added to the paratext in the *Both Condition* may indicate that the generated text may not summatively increase, and therefore not measurably impact, positive reception across *Literary*, *Creative*, *Technical*, *Enjoyable*, *Interesting*, and *Understandabl*e value dimensions amongst new readers of generated novels. So from a statistical analysis perspective this potentially means that, for example, readers' perception of a project's enjoyableness might not be strongly linked with the generated text.

The data from the survey question *Understanding* raises an important question and line of investigation. Participants' agreement with the statement "The project is understandable" does not significantly increase in mean response from the *Paratext Condition* to the *Both Condition* where the generated text is presented with the paratextual pieces. while this is surprising, it's clear that within this study the addition of the generated text does not impact the understandability of the project. So which of the paratextual pieces, then, do new readers judge to be the most important for helping to understand the project? This cannot be determined from the current stage of analysis because there are several potential candidates which make sense from a theoretical (as opposed to study-based) perspective: Might it be the code (piece B) that is responsible for generating the text and revealing how it is made? The creator's progress page

---

The exception are *Literary, Skill Code & Lit.* And *Interesting, Novel.* And finally *Understanding, Skill Lit.*

(piece D) which explains the intention behind the project? Or perhaps the hypotext/input (piece G) which the author used to inspire and connect the project to existing works or serve as the raw material for the generated text? This is answered and discussed in section *4.4.5.1. Ranking the pieces* for *Understanding*.

While the difference between the two conditions for survey question *Interesting* are not significant, the mean responses to the statement "The project is interesting" in both of the conditions are nevertheless quite positive and comfortably fall within the *Somewhat agree* and *Agree* range on the ordinal scale. This suggests that some or all of the paratextual pieces do work to impact a general sense of interestingness about the project. As is the case with *Enjoyable* however, based on the collected data it is not the generated text that is impacting these judgements. This is arguably a surprising result[34] and it will be explored further in section 4.4.5.2. *Ranking pieces for interest* with *Q72 Rank Interesting*.

Table 6: Inferential statistics (Paratext and Both Conditions)

| Question | First Condition | Second Condition | Group | T | DF | P | 95% Confidence Interval | Mean of X (First Condition) | Mean of Y (Second Condition) | Null Hypothesis | Significance |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Q1 Literary | Paratext | Both | | 0.44543 | 186 | 0.6565 | -0.340763 0.539518 | 4.72043 | 4.621053 | Accept | |
| Q1 Literary | Paratext | Both | Novel B | 4.59575 | 91 | 0.8279 | -0.6881300 0.5520189 | 4.577778 | 4.645833 | Accept | |
| Q1 Literary | Paratext | Both | Novel S | 0.80553 | 93 | 0.4226 | -0.3786445 0.8954885 | 4.854167 | 4.595745 | Accept | |
| Q1 Literary | Paratext | Both | Skill Code | 1.3101 | 60 | 0.1952 | -0.2379375 1.1411633 | 5.032258 | 4.580645 | Accept | |
| Q1 Literary | Paratext | Both | Skill Code & Lit | -0.60377 | 60 | 0.5483 | -1.0512994 0.5637994 | 4.6 | 4.84375 | Accept | |
| Q1 Literary | Paratext | Both | Skill Lit | 0.22929 | 62 | 0.8194 | -0.7235854 0.9110854 | 4.53125 | 4.4375 | Accept | |
| Q1 Literary | Paratext | Both | Pos Familiarity | 1.4877 | 82 | 0.1407 | -0.1770406 1.2270406 | 5.025 | 4.5 | Accept | |
| Q1 Literary | Paratext | Both | Neg Familiarity | -0.46537 | 92 | 0.6428 | -0.7470675 0.4634311 | 4.54 | 4.681818 | Accept | |
| Q1 Creative | Paratext | Both | | 2.0579 | 186 | 0.04099 | 0.01367894 0.64755479 | 6.204301 | 5.873684 | Accept | |
| Q1 Creative | Paratext | Both | Novel B | 1.1506 | 91 | 0.2529 | -0.1906862 0.7156862 | 6.2 | 5.9375 | Accept | |
| Q1 Creative | Paratext | Both | Novel S | 1.7507 | 93 | 0.0833 | -0.05370125 0.85334664 | 6.208333 | 5.808511 | Accept | |
| Q1 Creative | Paratext | Both | Skill Code | 0.95672 | 60 | 0.3425 | -0.2463056 0.6979185 | 6.193548 | 5.967742 | Accept | |
| Q1 Creative | Paratext | Both | Skill Code & Lit | 0.41419 | 60 | 0.6802 | -0.3909260 0.5950927 | 6.133333 | 6.03125 | Accept | |
| Q1 Creative | Paratext | Both | Skill Lit | 1.9361 | 62 | 0.05742 | -0.02130361 1.33380361 | 6.28125 | 5.625 | Accept | |
| Q1 Technical | Paratext | Both | | 0.83906 | 186 | 0.4025 | -0.1598203 0.3963795 | 6.11828 | 6 | Accept | |
| Q1 Technical | Paratext | Both | Novel B | 0.30401 | 91 | 0.7618 | -0.2843790 0.3871568 | 6.155556 | 6.104167 | Accept | |
| Q1 Technical | Paratext | Both | Skill Code | 0.34672 | 60 | 0.73 | -0.4615381 0.6550865 | 5.967742 | 5.870968 | Accept | |
| Q1 Technical | Paratext | Both | Skill Code & Lit | 0.88389 | 60 | 0.3803 | -0.2841872 0.7341872 | 6.1 | 5.875 | Accept | |
| Q1 Technical | Paratext | Both | Skill Lit | 0.16298 | 62 | 0.8711 | -0.3520256 0.4145256 | 6.28125 | 6.25 | Accept | |
| Q2 Interesting | Paratext | Both | | 1.3509 | 186 | 0.1784 | -0.1020788 0.5455423 | 6.032258 | 5.810526 | Accept | |
| Q2 Interesting | Paratext | Both | Novel B | -0.13882 | 91 | 0.8899 | -0.5103110 0.4436444 | 5.8 | 5.833333 | Accept | |
| Q2 Interesting | Paratext | Both | Novel S | 2.0843 | 93 | 0.03988 | 0.02186051 0.90367140 | 6.25 | 5.787234 | Accept | |
| Q2 Interesting | Paratext | Both | Skill Code | 0.8871 | 60 | 0.3786 | -0.2833612 0.7349741 | 5.935484 | 5.709677 | Accept | |
| Q2 Interesting | Paratext | Both | Skill Code & Lit | 0.92012 | 60 | 0.3612 | -0.2396823 0.6480157 | 6.266667 | 6.0625 | Accept | |
| Q2 Interesting | Paratext | Both | Skill Lit | 0.70729 | 62 | 0.482 | -0.456565 0.956565 | 5.90625 | 5.65625 | Accept | |
| Q2 Enjoyable | Paratext | Both | | 2.0374 | 186 | 0.04302 | 0.01306828 0.81092719 | 5.569892 | 5.157895 | Reject | Significant |
| Q2 Enjoyable | Paratext | Both | Novel B | 1.3109 | 91 | 0.1932 | -0.2061333 1.0061333 | 5.4 | 5 | Accept | |
| Q2 Enjoyable | Paratext | Both | Novel S | 1.5443 | 93 | 0.1259 | -0.1172361 0.9372716 | 5.729167 | 5.319149 | Accept | |
| Q2 Enjoyable | Paratext | Both | Skill Code | 1.4886 | 60 | 0.1418 | -0.1663145 1.1340565 | 5.580645 | 5.096774 | Accept | |
| Q2 Enjoyable | Paratext | Both | Skill Code & Lit | 0.91409 | 60 | 0.3643 | -0.3144024 0.8435691 | 5.733333 | 5.46875 | Accept | |
| Q2 Enjoyable | Paratext | Both | Skill Lit | 1.1924 | 62 | 0.2377 | -0.3382445 1.3382445 | 5.40625 | 4.90625 | Accept | |
| Q2 Understandable | Paratext | Both | | 1.4159 | 186 | 0.1585 | -0.1052922 0.6406628 | 5.215054 | 4.947368 | Accept | |
| Q2 Understandable | Paratext | Both | Novel B | 1.6763 | 91 | 0.09712 | -0.0824834 0.9741501 | 5.133333 | 4.6875 | Accept | |
| Q2 Understandable | Paratext | Both | Novel S | 0.29707 | 93 | 0.7671 | -0.4485175 0.6063189 | 5.291667 | 5.212766 | Accept | |
| Q2 Understandable | Paratext | Both | Skill Code | 2.2202 | 60 | 0.0302 | 0.06709093 1.28774778 | 5.483871 | 4.806452 | Accept | |
| Q2 Understandable | Paratext | Both | Skill Code & Lit | 0.5547 | 60 | 0.5812 | -0.3963415 0.7005082 | 5.433333 | 5.28125 | Accept | |
| Q2 Understandable | Paratext | Both | Skill Lit | | 62 | 1 | -0.7509565 0.7509565 | 4.75 | 4.75 | Accept | |

## 4.4.4. Familiarity with the hypotext

The data analyzed in this section is the *Q4 Hypotext Familiarity* (A20) and the *Q1 Literary* (A1) survey question, focusing first on the *Both Condition*, and then on all three conditions.

[34] Rather, it is personally surprising because I find the generated text itself to be terribly interesting.

To follow up on the *Literary* question results shown in table *Inferential Statistics*, where the only significant result was a difference in the *Novel B* group between the *Generated Text Condition* and the *Both Condition* , it is striking that there isn't a similarly significant difference in responses in the *Novel S* group. Why did the inclusion of the paratextual pieces along with the generated text (the *Both Condition*) have a significant impact on the responses to only one generated novel and not the other? This could be because participants might have had differing levels of familiarity with the hypotext - the esteemed literary novels which inspired the generated novels. Further, might participants' knowledge about the hypotext explain the lack of significant differences across the conditions in *Literary*? This is investigated in this section through the *Q4 Hypotext Familiarity* data.

The purpose of *Hypotext familiarity* is to measure new readers' familiarity with either *Ulysses* which is linked to *Novel B*, and *Pride and Prejudice* which is linked to *Novel S*. As discussed in section, *2.2.4. Key differences and similarities between the NaNoGenMo projects,* each classic work has a hypotextual relationship with the project and it was assumed that participants who are familiar with the classic work will therefore be aware of a relationship between them, and that this in turn could impact the project's perceived value. This question is therefore linked to cultural influences. During the study design stage, it was expected that higher familiarity with the hypotext would be seen with significantly higher positive responses to the *Literary* question especially, and that there would be clear differences between the participant skill groups – especially *Skill Lit*. To maximize the chance that the hypotextual relationship was recognize by participants, only the *Both Condition* was analyzed for *Hypotext Familiarity* because the presence of all the paratextual pieces is most likely to prime participants' memory about the classic works. The survey question was phrased slightly differently depending on the generated text sample that participants read:
"I am familiar with [Jane Austen's book Pride and Prejudice] [James Joyce's Ulysses] (for example, I know some details about the plot or the book's cultural status)".

However, in the sample distribution visualizations all *Both Condition* groups are observed to lean very heavily towards the negative pole. Further, the results in table *Descriptive Statistics* show that the Standard Deviation values are high with the majority of values exceeding two points on the 7-point ordinal scale. So it appears that overall, the majority of participants may not be familiar with either of the classic works. Although, there do appear to be more positive responses in the *Skill Lit* plot than in the other two skill groups. In order to test whether participants with literary reading skills actually are significantly more familiar with the classic works than the other skills groups are, another statistical test was run that can compare all three skill group at the same time – a Kruskal-Wallis and Dunn's test.

Figure 10: Hypotext Recognition (Both Condition)

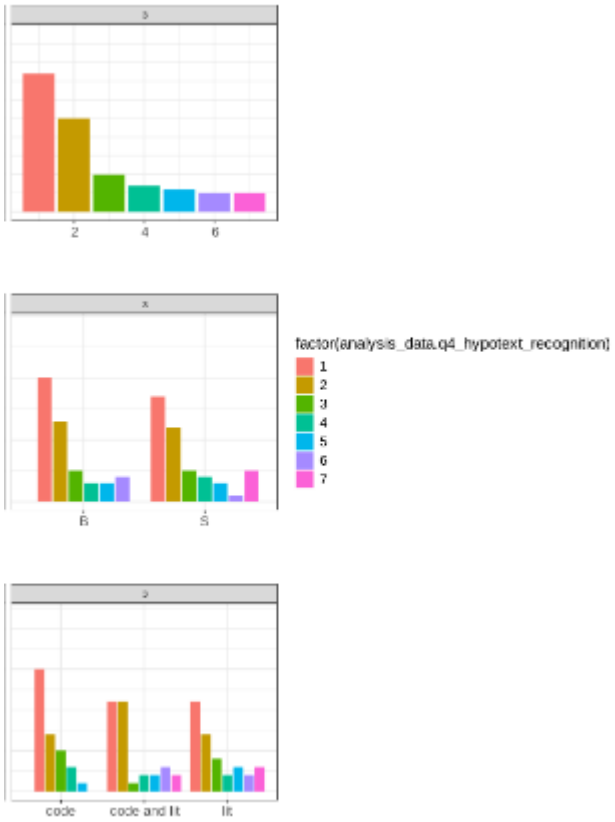



Table 7: Descriptive statistics Both Condition

| Question | Condition | Group | Mean | Standard Deviation |
|---|---|---|---|---|
| Q4 Hypotext Familiarity (A20) | Both | | 3.8 | 2.065934439 |
| Q4 Hypotext Familiarity (A20) | Both | Novel S | 4.319148936 | 2.011989962 |
| Q4 Hypotext Familiarity (A20) | Both | Novel B | 3.291666667 | 2.010169182 |
| Q4 Hypotext Familiarity (A20) | Both | Skill Code | 3.096774194 | 2.005905261 |
| Q4 Hypotext Familiarity (A20) | Both | Skill Lit | 4.8125 | 1.803893638 |
| Q4 Hypotext Familiarity (A20) | Both | Skill Code & Lit | 3.46875 | 2.03175397 |

A Kruskal-Wallis and Dunn's test with Bonferroni adjustment was run to test for significant differences between how each of the three skill groups responds to their familiarity with the hypotext in the *Both* Condition. As shown in table *Kruskal-Wallis Results,* there is a significant between the *Skill Lit* and the *Skill Code* groups (p adjusted = 0.002, Z = -3.211, Chi Squared = 11.219, p = 0). Initially this makes sense because people with a reported interest in literature are perhaps more likely to be familiar with *Pride and Prejudice* or *Ulysses*. However, the aforementioned lean to the negative pole in the visualizations shows that overall the majority of participants are not familiar with the classic novels; this is true even for those participants in the

*Skill Lit* group who selected 'literature' as one of their top three hobbies in the Prolific prescreening survey. In sum, these results suggest that the difference is not as meaningful as predicted during the study design stage. Indeed, practically speaking the difference is arguably not meaningful in the real-world context; *Skill Lit*'s mean response of about 4.8, or roughly between *Neither agree nor disagree* to *Somewhat agree* on the ordinal scale, is quite low if assuming that the participants might be literature afficionados (Mean = 4.81 ± SD = 1.8).

Table 8: Kruskal-Wallis adjusted p-value results (Both Condition, comparing skill groups)

| SKILL GROUPS | **Code** | **Code and Lit** |
|---|---|---|
| **Code and Lit** | 0.6265 | - |
| **Lit** | 0.002* | 0.0233 |

As seen in *Table 10: Kruskal-Wallis adjusted p-value results (Both Condition, comparing skill groups)*, based on the p ≤ 0.01 threshold used in this study, there is no significant difference between the *Lit* and *Code & Lit* skill groups (p adjusted = 0.0233, p = 0.007, Z = -2.42, Chi Squared = 11.219), and the Code and Code & Lit skills groups (p adjusted = 0.6265, Z = -0.81, p = 0.209). It is perhaps worth bearing in mind that this test was run using a Bonferroni adjustment on the p-values, which some data analysts may consider to be too conservative. A less conservative interpretation may choose to consider the difference between *Skill Lit* and *Skill Code & Lit* an important one, which would distinguish *Skill Lit* as the group with the higher familiarity with the classic works. Nevertheless, as already pointed out the level of familiarity is arguably underwhelming. And further, this higher familiarity was clearly not enough to significantly impact the group's perception of the project's literary value when responding to the *Literary* question.

It is possible that the participants who are most familiar with the hypotext (and, presumably, might therefore perceive the project to be more literary) are spread across the three skill groups rather than being concentrated in *Skill Lit*. Would such a group show significantly different responses for *Literary*? To find this out, all the survey data across all three conditions was split into positive (*Strongly agree*, *agree*, and *Somewhat agree*) and negative (*Strongly Disagree, disagree,* and *Somewhat disagree)* hypotext familiarity groups and named *Pos Familiarity* and *Neg Familiarity* respectively. There were very few neutral responses (*Neither agree nor disagree*) and these were not used. Next, in order to determine precisely whether familiarity with the classic work that inspires a generated novel project has any measurable impact on literary reception, or whether this may interplay with the study's different paratextual conditions, the positive and negative groups were split across each of the three conditions and analyzed as the

*Novel* and *Skill* were in the previous section; a Two Sample T-Test was run to test for significant differences in the mean response between the *Generated Text Condition* and the *Both Condition*, and the *Paratext Condition* and the *Both Condition*.

The Results are shown in Table *Grouping By Positive and Negative Familiarity*. *Pos Familiarity* had no significant differences between the *Generated Text Condition* and the *Both Condition* (p = 0.76, DF = 88, t = -2.86). It also has no significant differences between the *Paratext Condition* and the *Both Condition* (p = 0.14, DF = 82, t = 1.49). Similarly, *Neg Familiarity* had no significant differences between the *Generated Text Condition* and the *Both Condition* (p = 0.048, DF = 84, t = -2.01). It also has no significant differences between the *Paratext Condition* and the *Both Condition* (p = 0.64, DF = 92, t = -0.47). Even when looking at the Means for *Pos Familiarity* and *Neg Familiarity* in each of the conditions it's clear that differences are not meaningful because the values are so close to each other.

Table 9: Grouping by positive and negative familiarity

| Question | First Condition | Second Conditio | Group | T | DF | P | 95% Confidence Interval | Mean of X (First Condition) | Mean of Y (Second Condition) | Null Hypothesis |
|---|---|---|---|---|---|---|---|---|---|---|
| Q1 Literary | Generated Text | Both | Pos Familiarity | -0.28595 | 88 | 0.7756 | -0.8641111 0.6467198 | 4.391304 | 4.5 | Accept |
| Q1 Literary | Generated Text | Both | Neg Familiarity | -2.0107 | 84 | 0.04756 | -1.356136191 -0.007500173 | 4 | 4.681818 | Accept |
| Q1 Literary | Paratext | Both | Pos Familiarity | 1.4877 | 82 | 0.1407 | -0.1770406 1.2270406 | 5.025 | 4.5 | Accept |
| Q1 Literary | Paratext | Both | Neg Familiarity | -0.46537 | 92 | 0.6428 | -0.7470675 0.4634311 | 4.54 | 4.681818 | Accept |

These *Pos Familiarity* and *Neg Familiarity* results show that overall within this study, regardless of familiarity or a lack of knowledge about the classic novels *Pride and Prejudice* or *Ulysses*, or regardless of the presence of the generated text, or the presence of the paratextual pieces, new readers' perceptions of a generated novel's literariness are not significantly impacted overall. Thus, the data shows that the perceived literariness of a generated novel is largely immutable within this study.

However, it is possible that the right questions to capture relevant skill differences amongst participants were not asked during the Prolific screening survey, or that the participants who were recruited do not have the skills that the analysis assumes they do. An opportunity for deeper qualitative analysis to better understand the differences between participant skill backgrounds and reception is afforded in the second study in *Chapter 5 Workshop study*.

### 4.4.5. Ranking the pieces

This section describes the analysis of the data and results which are used to directly addresses the first research question - which paratextual elements have a central role in influencing the understanding and reception of a project? The ranking data is particularly interesting because it shows participants' explicit judgements about the paratextual pieces.

The data analyzed in this section is the *Q71 Ranking Understanding* and *Q72 Ranking Interesting* questions from the survey, where participants in the *Both Condition* and the *Paratext Condition* were asked to rank each of the project pieces in order from the most important to the least important. The aim was to collect data which could generally indicate which pieces were prioritized by the new readers overall. The weighted Brute Force algorithm from the RankAggreg R package (Pihur et al., 2009) is used to compute an aggregated optimal ranked list of these paratextual pieces. This is a list of the pieces according to their mean rank in the survey data. Therefore, it can be interpreted as a general indication of which pieces participants may consider to be the most important overall.

The R package used is RankAggreg, which performs weighted aggregation of ordered lists based on their ranks[35]. The task is approached as an optimization problem where "...we first need to define the objective function. In this context, we would like to find a 'super'-list which would be as 'close' as possible to all individual ordered lists simultaneously" (Pihur et al., 2009, p. 2). The goal is thus to compute one optimal ranked list while minimizing the distance, or, the differences between all of the ranked lists that each participant responded with. The optimal list can also be thought of as a potential Mean list. The algorithm outputs the Minimum Objective Function score along with the optimal list, and the score can be used to compare the algorithm parameter options that the programmer uses[36], and to compare the optimal list computations. Therefore a comparatively lower objective function score is, technically speaking, better than a comparatively higher one. Kendall's tau distance[37] was used to compute the distance between the survey questions' ranked lists. This distance is normalized and used to compute the weighting for the Brute Force algorithm[38].

---

[35] The documentation for the RankAggreg package is found here: https://rdocumentation.org/packages/RankAggreg/versions/0.6.6/topics/RankAggreg

[36] I am the programmer and Kendall's tau distance is the weighting parameter that I used.

[37] Kendall's tau distance "utilizes pairs of elements from the union of two lists"; more details can be found in Pihur et al. (2009, p. 3).

[38] The RankAggreg package has one other distance measure that can be selected for the weighting parameter: the Spearman footrule distance. The package authors explain that the Kendall's tau and Spearman's distances are "the two most popular ones...[they] usually produce slightly different aggregated lists which is mainly due to the differences in the two [statistical] philosophical paradigms" (Pihur et al., 2009, p. 2). Indeed, when exploring the

The package offers three rank aggregation algorithms, of which the chosen Brute Force algorithm is the least sophisticated one. It "…simply tries all possible solutions and selects the one which is optimal" (Pihur et al., 2009, p. 2). While the algorithm is very simple, the authors suggest that it is feasible to find the optimal solution with it when the number of possible solutions is not large (p. 5); so Brute Force is a good choice for this study because the ranking question data's small size requires minimal computational resources to run.

The design of the survey's ranking questions forces each participant to make a conscious decision to prioritize individual pieces. These are preceded by the free response written review question, which gives participants the opportunity to reflect on the project before being asked to rank the pieces. While the algorithm computes an optimal ranked list per survey question condition, the list is analyzed from a descriptive perspective where each optimal list is interpreted as a highly likely potential average list (much like a statistical mean value), rather than as a definitive, prescriptive master list where strict rank order is important. Therefore, meaningful groups of paratextual pieces and their general position on the optimal lists are allowed to emerge naturally from the analysis. Intuitively, this descriptive approach to the optimal list results makes sense because the mean ranks, as seen in *Table 10: Project Pieces Ranked for Understanding* and in *Table 11: Project Pieces Ranked for Interesting*, do not always show a large difference between each other.

#### 4.4.5.1. Ranking the pieces for understanding

The *Ranking Understanding* questions asked participants:
"The names of the pieces you read earlier are shown below. The survey would like to know your personal opinion about which of these were the most important for helping to **understand** the project. Please rank the pieces in order of importance."

Following on from the analysis of the *Understanding* survey question and the line of inquiry it raised in section *4.4.2.3. Inferential Statistics* it's clear that within this study the addition of the generated text does not significantly impact the understandability of the project. So which of the paratextual pieces, then, do new readers judge to be the most important for helping to understand the project? The major difference with answering this question using the Ranking

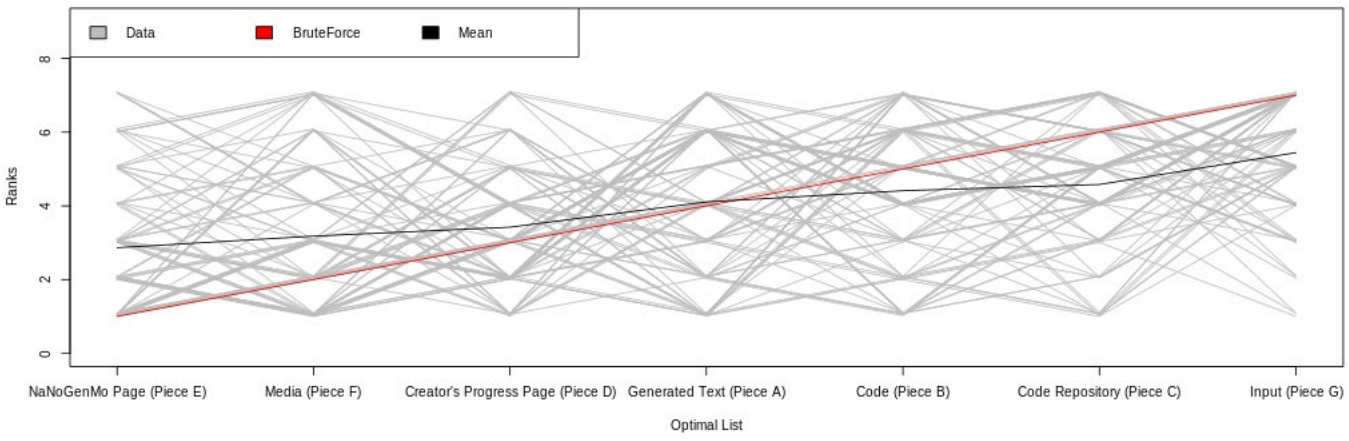
---

package's suitability for analyzing the survey's ranking data, both distance measures were ran and their optimal lists compared. They both produced very similar optimal lists, and indeed both reflected the analysis results presented here. Therefore, because the choice of distance measures did not meaningfully impact the results, Kendall's tau is used in this analysis because it has lower a Minimum Objective Function Score than Spearman does.

Understanding data is that participants are judging each individual project piece, rather than all of them at the same time. Analysis in this section will begin by focusing on the optimal ranked list that is computed based on the Both Condition data, in order to understand which pieces new readers tend to prioritize when they are presented with the whole project. Next, the Paratext Condition data will be analyzed in the same way in order to see whether the absence of the generated text affects the judged importance of the remaining paratextual pieces.

The Brute Force algorithm with Kendall's tau weighting was run to compute the aggregated optimal list from the responses to Ranking Understanding in the Both Condition. The optimal list, mean rank, and the minimum objective function score are shown in table Project Pieces Ranked for Understanding, where the first rank (rank 1) represents the most important piece, and the last rank (rank 7) the least important. *Figure 11: Q71 Ranking Understanding, Both Condition* conveys the results visually: the distribution of participants' ranked list responses is shown with the grey Data lines, the mean for each of these lists' ranks is indicated with the black Mean line, and the red BruteForce line represents the optimal list that is computed based on the other data. The BruteForce line is the same in every rank aggregation visualization in this section because it needs to consecutively intersect through each rank point; it can be seen that BruteForce closely follows the Mean, although it is not possible to follow it exactly as each paratextual piece can of course only appear once in the optimal list. The optimal list for the Both Condition is E, F, D, A, B, C, G, or, The NaNoGenMo page (E), Media page (F), The creator's progress page (D), The generated text (A), The code (B), The code repository (C), and the input (G). The minimum objective function score is about 2.6 (Score = 2.584406); while it's convention to report this score, it's not important for analysis at hand.

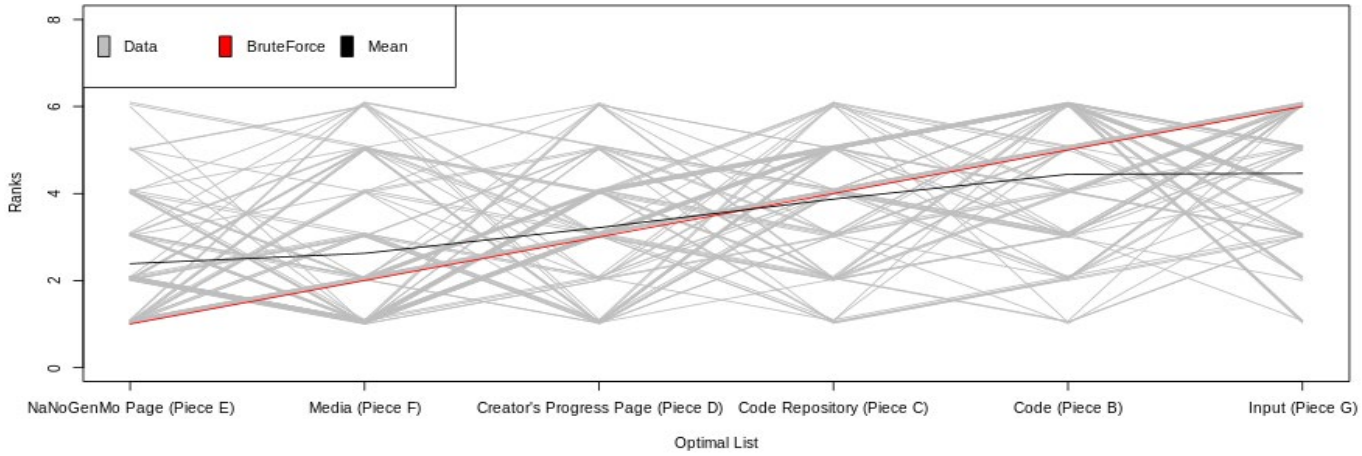
Figure 11: Q71 Ranking Understanding, Both Condition




The same Brute Force algorithm with Kendall's tau weighting was run again to compute the aggregated optimal list for the *Paratext Condition*. Table *Project Pieces Ranked for*

*Understanding* shows the optimal list, mean rank, and the minimum objective function score, and Figure *12: Q71 Ranking Understanding, Paratext Condition* conveys the former two results visually. The optimal list is E, F, D, C, B, G. The minimum objective function score is about 1.8 (score = 1.729219). This score is lower than in the *Both* Condition. Therefore this indicates that the distance (or, difference) between the optimal list and the participants' response lists in the *Paratext Condition* is smaller than the distance between those in the *Both Condition*. Although this lower score is technically better, they're is not considered to be relevant for the current analysis because this study does not aim to evaluate the performance of different models, and so the scores will not be discussed further. Reporting the scores is a convention that is being done for the sake of completeness.

Figure 12: Q71 Ranking Understanding, Paratext Condition




Based on the results from both of the conditions, the NaNoGenMo page (piece E) ranks first, which suggests that it may be the most important for helping new readers to understand the project. This makes sense because the page contextualizes and motivates the reason behind the project by introducing the challenge and explains why it exists and what the rules are. E is followed by pieces F and D: the Media page and the Creator's progress page. These also function to explain the project from a reviewer or critic's perspective, to the author's own aims and motivations, respectively. The E, F, D trio maintain this order in the top half of the optimal list across both of the conditions.

The generated text (piece A) is situated in the middle of the optimal list in the *Both Condition*, and this placement perhaps suggests a neutral or undecided (or daresay, an apathetic) overall feeling towards the importance that the generated text lends to the understanding of the project. While they don't maintain a strict order like E, F, D, the Code, Code Repository, and the Input (pieces, B, C, G), are seen as a set of three at the bottom of the optimal list. Although their ordering changes across the two conditions they nevertheless clearly emerge as a group. This

set perhaps makes sense intuitively because each of these pieces are the 'technical infrastructure' or the technical pieces which computationally realize the generated text. Perhaps for the same reason it makes sense that they may be the least important for understanding; while some readers with coding skills may be able to understand what these pieces do when they are run (although this assumption is not supported by the *Understanding* question results shown back in table *Inferential Statistics*), B, C, G don't explain why the project exists to the extent that E, F, D do[39]. However, this is not so clear-cut because C, the Code Repository, contains a README file which gives some information about the project. Piece G (the Input) remains in the lowest rank position in both of the conditions – it is possible that its function in the project is not entirely clear to readers.

Table 10: Project pieces ranked for Understanding

| **Both Condition** | | | **Paratext Condition** | | |
|---|---|---|---|---|---|
| (Min. Objective Function = 2.584406) | | | (Min. Objective Function = 1.729219) | | |
| **Piece** | **Mean Rank** | **Brute Force Rank** | **Piece** | **Mean Rank** | **Brute Force Rank** |
| E | 2.863158 | 1 | E | 2.387097 | 1 |
| F | 3.178947 | 2 | F | 2.623656 | 2 |
| D | 3.421053 | 3 | D | 3.215054 | 3 |
| A | 4.105263 | 4 | - | - | - |
| B | 4.410526 | 5 | C | 3.870968 | 4 |
| C | 4.578947 | 6 | B | 4.44086 | 5 |
| G | 5.442105 | 7 | G | 4.462366 | 6 |

#### 4.4.5.2. Ranking pieces for interesting

The *Ranking Interesting* question asked participants:
"The names of the pieces you read earlier are shown below. The survey would like to know your personal opinion about which of these were the most important for helping to make the project **interesting**. Please rank the pieces in order of importance."

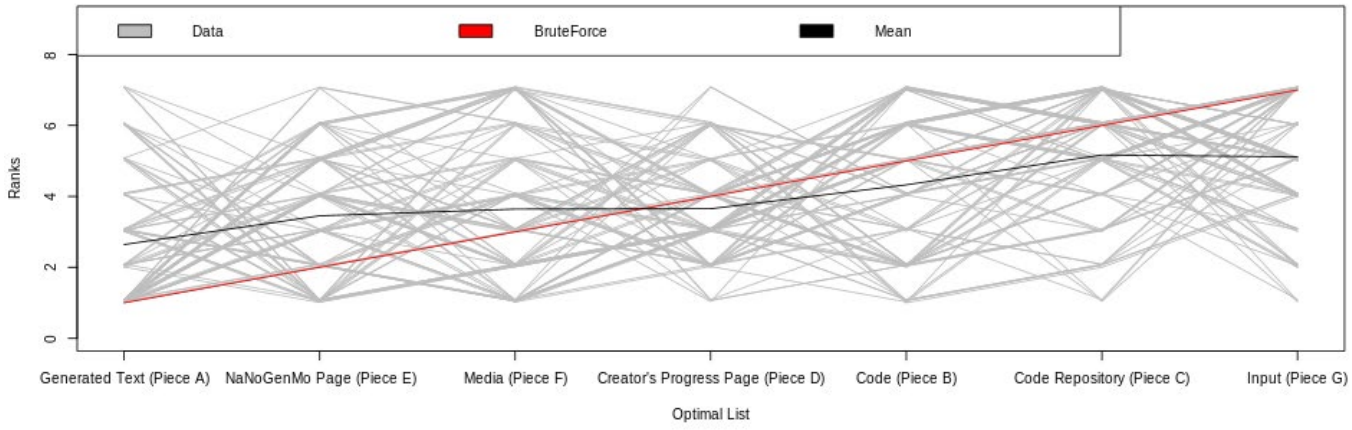
Following on from the questions raised in *4.4.3.2. Inferential Statistics*, the mean responses to the *Interesting* survey question in the *Paratext Condition* and the *Both Condition* was positive overall. The ranked list aggregation analysis in this section will be able to indicate which of the paratextual pieces new readers judge to have the strongest impact on a general sense of interestingness about the project. Mirroring the previous section, analysis will begin by focusing on the optimal ranked list that is computed based on the *Both Condition* data, and then the

[39] I have a background in Fine Art and Literature, so admittedly I find it difficult to understand why a creative piece would need a reason to exist. Nevertheless, this is what the data appears to indicate.

*Paratext Condition* following the same method. As with *Ranking Understanding,* the results will be able to show whether the absence of the generated text in the *Paratext Condition* affects the judged importance of the remaining paratextual.

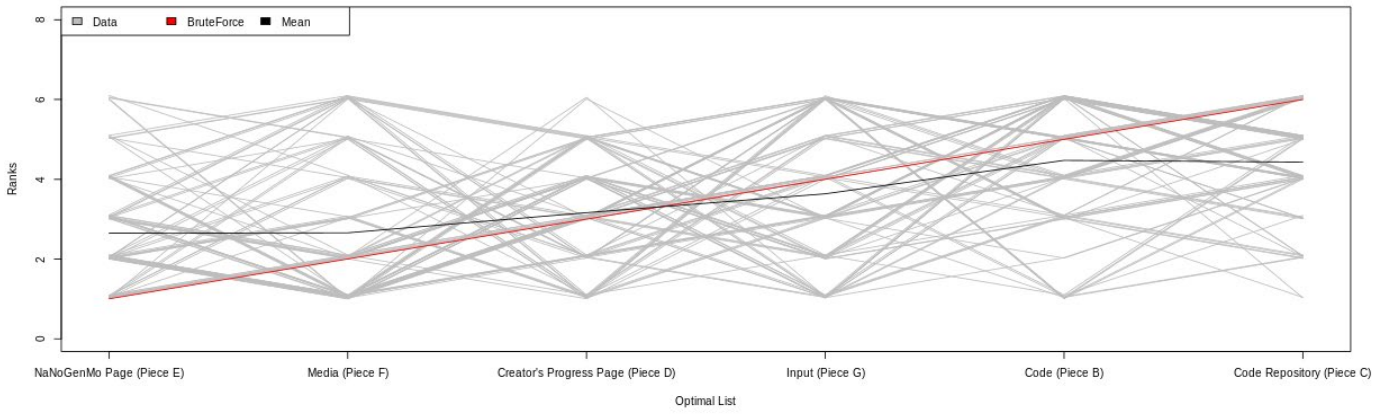
The Brute Force algorithm with Kendall's tau weighting was run to compute the aggregated optimal list from the responses to *Ranking Interesting* in the *Both Condition*. The optimal list, mean rank, and the minimum objective function score are shown in table *Project Pieces Ranked for Interesting*. Figure 13: *Q72 Ranking Interesting, Both Condition* conveys the results visually. The optimal list is A, E, F, D, B, C, G. The Minimum Objective Function Score is about 2.4 (Score = 2.440875).

Figure 13: Q72 Ranking Interesting, Both Condition




The same Brute Force algorithm with Kendall's tau weighting was run again to compute the aggregated optimal list from the responses to *Ranking Interesting* in the *Paratext Condition*. The optimal list, mean rank, and the minimum objective function score are also shown in table *Project Pieces Ranked for Interesting*, and *Figure 14: Q72 Ranking Interesting, Paratext Condition* conveys the former two results visually. The optimal list is E, F, D, G, C, B. The Minimum Objective Function Score is about 1.9 (Score = 1.92678). This is lower than the Score for the *Both Condition*.

Figure 14: Q72 Ranking Interesting, Paratext Condition




The results for Ranking Interesting show that the E, F, D order (The NaNoGenMo page, The media page, and The creator's progress page) is maintained at the top of the optimal list in both conditions, although the trio shifts down one rank when the generated text (piece A) is present in the Both Condition. So while A, the generated text, could be the most important for helping to make the project interesting (as suggested by a low mean rank), the E, F, D order is nevertheless maintained and ranks comparatively higher for interest than the B, C, G set (The code, The code repository, and The Input). Further, when the generated text (piece A) is not present in the Paratext Condition then the project's input data and hypotext (piece G) position moves up in the list. This may be because it replaces A as something interesting to read. Nevertheless, the B, C, G set remains in the bottom half of the optimal list in both of the conditions, meaning that these pieces rank comparatively lower for interest.

Table 11: Project pieces ranked for Interesting

| **Both Condition** | | | **Paratext Condition** | | |
|---|---|---|---|---|---|
| (Min. Objective Function = 2.440875) | | | (Min. Objective Function = 1.92678) | | |
| **Piece** | **Mean Rank** | **Brute Force Rank** | **Piece** | **Mean Rank** | **Brute Force Rank** |
| A | 2.642105 | 1 | - | - | - |
| E | 3.452632 | 2 | E | 2.645161 | 1 |
| F | 3.642105 | 3 | F | 2.655914 | 2 |
| D | 3.652632 | 4 | D | 3.16129 | 3 |
| B | 4.326316 | 5 | G | 3.634409 | 4 |
| C | 5.168421 | 6 | B | 4.473118 | 5 |
| G | 5.115789 | 7 | C | 4.430108 | 6 |

#### 4.4.5.3. All Ranking results discussion

Both the E, F, D ordered group at the top of the optimal ranked list, and the B, C, G unordered set at the bottom are maintained not only across both conditions, but across both questions as well. Therefore the analysis in this section gives some evidence that pieces E, F, D – the

NaNogenMo page, the Media page, and the Creator's progress page – are likely to be the most important pieces for both understanding and interest in this study. This suggests that it is their function within a project, or the information that they convey, that is prioritized by new readers. Like the NaNoGenMo page, the Media page also introduces the challenge but further gives an overview of the generated novel project from a reader's or critic's perspective. The Creator's progress page documents the motivation or inspiration for the project from the author's perspective, as well as how it progresses over the course of the challenge. Mirroring this, the analysis also gives some evidence that new readers likely judge the set of pieces B, C, G – the Code, the Code repository, and the Input - to be both the least important for helping to understand the project, and the least helpful in making it interesting. It is possible that the shifting ordering of this trio could be due to the nature of the survey question; the order of the top-ranking items might be perceived to be more important by the participants than the order of the bottom ranking items.

There is a clear indication that when the generated text (piece A) is present, it is likely to be the most interesting piece. This makes sense intuitively, both because it is unusual and because it appears to be considered a priority just as 'the text' might be from Genette's paratextual conceptualization. However, it seems that the generated text is not important for understanding the project based on both these ranking results and on the *Understanding* question results previously discussed in section *4.4.3.2. Inferential Statistics*. This may in part be due to the fact that it's the first piece presented in the survey where it might trigger the most curiosity as well as the most confusion and frustration, although it's probable that the difficult-to-parse generated text style is responsible for lower understanding. The relationship between difficulty parsing text and understanding it is expanded on in the workshop study, and is explained in *Chapter 5 Workshop Study*.

## 4.5. Qualitative analysis of survey reviews

The Q6 free-response reviews written by the survey participants were analyzed using Qualitative Content Analysis. The survey question asked "What did you think of the project? Please write a short review explaining your personal opinion". Reflecting the major groups used in the quantitative study, the reviews are grouped by a combination of generated novel, skill, and paratextual condition groups. The questions that this portion of the study aims to contribute towards answering are the research questions plus a better understanding of the unexpected results of the Q1 *Literary* question from the quantitative study: why does the presence of the paratext appear to have no impact on participants' literary value judgements from the *Generated Text Condition* to the *Both Condition*? What might be distinctive about responses to Bhatnagar's *Molly's Feed* (*Novel B*) in the *Generated Text Condition*?

### 4.5.1. Method motivation and overview

Out of the many qualitative analysis methods which are used in research, Qualitative Content Analysis was chosen because of the review data characteristics and because this study's questions are more directed in nature and seeking to contextualize the quantitative data, rather than being exploratory or seeking to build theory.

With text-based data comprised of approximately 288 short reviews each ranging from 1 to 6 sentences in length, it is feasible for a researcher to read them all. However, because of the short and often repetitiveness the reviews, as well as the paratextual conditions and generated novels which contextualize their content, the reading and analysis of reviews is made easier with text analysis tools and approaches to help systematically manage the reviews and determine whether they answer the study's question. In their method textbook, Riffe et al. (2019) explains that Content Analysis can be applied to content that has been produced for particular research problems, and experimental conditions or exposure conditions where participants may "…be asked to write or report their post-exposure sentiments" (p. 26). This fits the 3-condition reading experiment and survey study very well, and I further note that the relatively short text format of the reviews is not dissimilar to social media comment posts which are a common unit of Content Analysis.

In Drisko and Maschi's methods chapter which specifically focuses on Qualitative Content Analysis, the researchers note that the method is
"…is most often used descriptively rather than to develop concepts and theory. The yield of qualitative content analysis is most often descriptive categories and themes; conceptualization and theory are not often part of the method. In turn, the aim of coding in qualitative content analysis is not to generate concepts and theory, but instead to describe the meanings and actions of research participants and texts" (2015, p. 105).
The qualitative method's sharp focus on describing the data rather than conceptualizing it fits well with the aims of this review study, because the reading experiment and survey where the reviews originate from has been designed around the paratextual framework; thus the conceptualization and theory behind the study is already known, rendering Qualitative Content Analysis an ideal method to employ.

Drawing from Drisko and Maschi (2015) and Riffe et al. (2019) Qualitative Content Analysis was planned in 3 stages:

#### 4.5.1.1. Stage 1: immersion and initial category development

Closely following Drisko and Maschi, analysis begins with immersion in the review data by carefully reading it and identifying the initial categories to code for:
"The goal is for the researcher to become informed about the content in context, to begin to notice key content and omissions of what might be expected content or perspectives, and to begin to identify connections within the data and preliminary categories." (2015, p. 103).

Because the foundations and research questions of this PhD project are built upon the intertextual framework as first presented by Genette, the analysis of the reviews takes a deductive Qualitative Content Analysis approach[40]. Here, initial categories are developed deductively and determined before the coding beings so that the objectives of answering the research questions and better understanding some of the unexpected results of the quantitative survey analysis might be met.

While Drisko and Maschi discuss the option of developing categories with multiple levels of abstraction and a hierarchy of codes (p. 105), they conclude that "Qualitative content analyses typically use a single-stage method of data analysis" (p. 120). This in alignment with the method described in Riffe et al. (2019). Similarly, this study uses single-stage coding and does not structure codes into hierarchies in order to instead focus on how categories appear together in reviews, since a non-hierarchical structure readily answers the study's questions.

#### 4.5.1.2. Stage 2: Coding and finalization of codes

Due the qualitative focus of the study, each review is coded for explicit or implicit categories, such as specific keywords or latent themes depending on which is most useful in the context of the review data. During coding, a high degree of flexibility is maintained to allow for additional categories to be created inductively as themes developed from the analysis organically. Thus, the finalized categories are a balance of a deductive and indictive process.

#### 4.5.1.3. Stage 3: distribution of category frequencies

Following the finalization of the codes, Schreier (2014a) in Drisko and Maschi "…describes the final step of data analysis in qualitative content analysis as preparing the data in a manner that clearly answers the research question" (2015, p. 109). Because understanding the data heavily

[40] Industry sources have referred to specifically deductive approaches to qualitative Content Analysis as Directed Content Analysis: https://delvetool.com/blog/contentanalysisdirected. Established academic sources such as Drisko and Maschi (2015) and Riffe et al. (2019)do not offer a specific term.

relies on the differences between the groups and conditions, I choose matrix representations to visually convey differences and similarities between groups and categories, and how frequently (or not) a category was written about in each group. The matrix representations are therefore helpful in managing and structuring the reviews so that important relationships between categories can more easily be found, as well as how trending or unique certain categories are throughout the reviews based on their frequency of use.

### 4.5.2. Method and analysis

#### 4.5.2.1. Category development and coding

Based on the 3 stages of Qualitative Content Analysis described above, the survey's reviews were analyzed:

First, as part of the immersion stage each review was read as the survey results were marked completed on the Prolific platform in order to get a general overview of the content and participants' opinions. A polarization of opinions in the reviews was noted. After the survey questions *Q1 Literary*, *Creative*, and *Technical* (discussed in section) had been analyzed, the reviews belonging to each of the 18 review groups were extracted from the survey data as plain text files. Each group contains approximately 16 reviews, for a total of 288 free response reviews (n = less than 288). The grouping is generally maintained from the quantitative survey data and analysis, so reviews are grouped by novel, condition, and skill. Unlike the survey skill groups, these review skill groups are further grouped by novel as well because it was expected that specific details about the generated novels would be written about. For example, group *Skill Code & Lit Reviews B1* represents all reviews written by participants who reported having both literary reading skills and computer programming skills, and who were asked to read *Novel B* in the *Generated Text Condition*. The reviews were read a second time per group and notes were taken.

Category development was performed next. The reviews were read for a third time, this time with the aid of a Keyword-In-Context tool that is part of the software toolkit AntConc[41]. The tool was used to read reviews filtered by keywords to verify my perception that certain themes were present in the reviews as I developed the initial thematic categories. For example, the search term *litera** was used to verify that several participants did indeed feel it was important to refer to literature or the literary in their review. This stage of analysis clearly showed that many

[41]Antconc is an open-source GUI concordance and text-analysis toolkit that is often introduced in Digital Humanities and Corpus Linguistics university courses. https://www.laurenceanthony.net/software/antconc/

reviews referred to more than one theme, and so the decision to use simultaneous coding was made.

Supplementary to this PhD project's research questions, this Qualitative Content Analysis study's research questions are:

In the *Q1 Literary* data, the only participant group which had a significant difference between the *Generated Text Condition* and the *Both Condition* read *Novel B*.

1. In which ways does the review content differ between the two conditions for *Novel B?*
2. What are some of the opinions expressed about literature and literary value between the two Novel B conditions?

The initial categories were developed deductively before formal coding began and include themes which capture what participants were primed and seen to discuss (categories such as *Literature and literary, Creativity, AI, Programming and Algorithms*), reoccurring themes which were noted while reading the reviews (categories such as *Joyce and Ulysses, text quality, authorship and writing, Scary*), and categories which reflect the paratextual framework at work and the research questions (categories *Paratext, Reading Experience*). All categories (including the initial categories) and their descriptions and example reviews can be seen in table *Finalized Categories Codebook below*.

After developing the initial categories, the coding stage began. The text files containing the reviews from the 18 groups were imported into the Nvivo qualitative data analysis software[42] where all the coding and analysis took place. The initial codes were created in the software and coding began by systematically reading each review per group and using a simultaneous coding approach. Each review thus constitutes one unit of analysis, and all approximately 288 reviews were coded using this approach. Any additional categories which developed inductively during coding were also created and coded for in this stage. After each additional category was created, previously read reviews were reread to check whether the new category applied. For example, categories G*ame and puzzle, Art, AI media,* and E*xperiment* developed inductively during coding. Finally, after this 'first-pass' of coding was completed, the reviews in each group were read again to check that coding was consistent across all data. Finally, the reviews were checked once more per thematic category to check that they aligned with the category. The *inspiration* category emerged from this process, and the *Scary* category was renamed from

[42] Nvivo is considered to be a professional analysis tool used in academia and industry. https://lumivero.com/products/nvivo/

*Discomfort and fear* since each coded review contained the term 'scary' - this made for a succinct description.

The finalized categories can be seen in *Table 14: Finalized categories codebook*.

As the sole researcher on this PhD project all reading, coding, and analysis was performed by me, therefore intercoder agreement was not planned for as part of the study design.

Table 12: Finalized categories codebook

| **Category** | **Description** | **Example and Group** |
|---|---|---|
| AI | An explicit reference to the term 'AI' or closely related terms or processes (such as 'neural networks' or 'learning'). The reference is not just about programming or algorithms more generally. | "The project seems interesting, although it does not prove anything in my opinion. The text generated by the neural network is really nonsense and has no style, we, when reading this text, give it meaning and see a specific style in it, because the human mind has a tendency to organize chaos in order"<br>(Skill Code & Lit, S3) |
| AI media | Specific or general references to AI generated works, forms, or genres. | "Generally speaking, I think NaNoGenMo is fascinating. I have read some scripts created by AI, but I did not know it was such a big project. ..."<br>(Skill Lit, S2) |
| Art | The term 'art' or 'artistic' is used, or references to types of art or movements (such as Social Realism), or specific artworks or artists. | "I liked it, it was almost like an avant-garde piece of modern art. I guess it is. I think it was inspiring, even though it wasn't very coherent."<br>(Skill Lit, S1) |
| Authorship and writing | Explicit and implicit references to human authors and human writing in general, computer and AI authorship and writing, and humans/computers sharing | "As a books enthusiastic I'm not a big fan of this project, because it changes the whole purpose of literature. The story does not make sense. Codes can not write like humans. They do not have our minds, |

| | | |
|---|---|---|
| | authorship or writing together. Includes writer as coder, designer, or creator. Includes references to specific authors. References to Joyce only (as opposed to also referencing authorship and writing more generally), are excluded. | our creativity. However, I don't think this is a bad idea. I just think we shouldn't take these types of book seriously or at least we should not give them the same value as real books written by real people."<br>(Skill Lit, B2) |
| Character | Use of the term 'character' or related terms such as 'characterization' or 'protagonist'. Includes specific or implied references to characters (For example, 'Molly' in *Novel B*) or narrators, or other entities identified in the generated text. | "It was very interesting how with generated phrases you can understand the personality and backgrounf [*sic*] of the character. Alas, the generated text was very repititive [*sic*] at times, and it can be boring in some points."<br>(Skill Lit, B1) |
| Creativity | Explicit reference to creativity through related terms such as 'creative', 'creatively', or an entity or work being creative. Excludes less descriptive/more functional uses of the term such as 'the author created' or 'the creator'. | "At first it was not very easy for me to understand what I was reading but then I started over and I thought it was a very intelligently-written piece. Very creative and with meaning behind"<br>(Skill Code, S1) |
| Experiment | Explicit use of the term experiment, or references to laboratory work or processes. | "It was an interesting experiment to see how computer generated stories could be created. However, at this level, from a literature point of view, the result was gibberish."<br>(Skill Lit, B3) |
| Game, puzzle, interaction | Explicit or implied references to the project, or engaging with | "It's an interesting way to play with your mind and your reading comprehension |

| | | |
|---|---|---|
| | suggestions of play, solving, or interactiveness. | ability. Maybe if I had read the book I would have understood better, though" (Skill Lit, B1) |
| Human machine collaboration | References to writers or programmers working with the technology or processes. Potential uses of the project or processes in it for writers, human involvement or intervention. | "Writers, like me feel being greatly helped by writing prompts like these as a guidance tool to further crystallize their vision as to where, and in which format a story should go, yet I don't feel AI can replace my work anytime soon. I mean, one sample page is a huge block text that is textbook eye cancer, something I lambaste my fellow writers for as well." (Skill Lit, S3) |
| Inspiration | Explicit use of the term inspiration, inspire, etc. | "I think it's an amazing and fun idea. It is fantastic for a reader and writer to join this project to have fun and maybe to find inspiration ideas or subjects from the past and from different genres." (Skill Lit, S2) |
| Joyce and Ulysses | Explicit, named references to James Joyce or Ulysses (and various alternative spellings). | "I love it. I like Joyces work so seeing something inspired by it is always welcome. And the novel itself (which i looked into) is pretty fun." (Skill Code, B2) |
| Literature and literary | Specific terms 'literature' or 'literary', or types of literature or poetry such as 'surrealism' or 'beantik'. Specific authors other than Joyce or Austen (such as William Burroughs). | "I think it's very interesting especially in this time we're living that's not full of really high piece of literature, that's halfway between a game and an experiment so i found it really meaningful especially thinking about what a computer can do in our times" (Skill Code and Lit, S3) |
| Monkey | Monkey similes or metaphors. | "Interesting idea, kinda like horde of monkeys with a typewriters, only its |

| | | |
|---|---|---|
| | | monkeys with scissors and glue and specific instructions for what they cut and how they connect words and sentences. Reminds me of DNA CRISPR manipulation, hope that nobody will write a program to create creatures of a genome same way like programmers did with original texts and frankensteined result."<br>(Skill Lit, B2) |
| Paratext | References to specific pieces in the project. Examples of other works being referenced paratextually, or thoughts expressed about paratextual relationships, such as intertextuality. | "I thought the project was intricate and interesting. I have a degree in English Literature and I still read a lot, so the project appeals to my interests. I recognised the Pride and Prejudice reference in the first line, so I was searching for further literary references in the rest of the text."<br>(Skill Lit, S1) |
| Participation | Referring to wanting to take part in a challenge, or in a similar project. References to groups and community. | "I have never thought of using code to write a creative writing piece or any other long passage. For that reason, I found this project very creative and intriguing. I see myself trying to attempt to learn how to write computer generated texts"<br>(Skill Code & Lit, B3) |
| Programming and algorithms | References to the project code, its technical workings, or programming or coding more generally. Although this category with AI if specific technical aspects are | "I think the project was a very clever and creative idea for NaNoGenMo. Using \"Ulysses\" by Joyce as input to create computer generated text is an excellent idea because of how the novel is written: Joyce's stream of consciousness is similar to the style of a text that a computer might |

| | | |
|---|---|---|
| | discussed, such as algorithms, but otherwise distinct. | generate. I enjoyed reading the result and having insight about the process.”<br>(Code & Lit, B3) |
| Reading experience | References to how a participant felt or reacted, physically, emotionally or cogitatively during reading, or speculations about how others may experience reading.<br>Not included are very general references which simply describe something as ‘interesting’ without expanding. | “I have never seen or heard of anything like this before. It is very creative and I find that there was a lot of thinking behind it, but I can see how and why many people would find this boring or uninteresting. The reason being that it is indeed a bit hard to understand and follow and I think a great amount of people (including me up to a point) are not familiar with pc coding and therefore unable to understand a great part of this project. It certainly is a thought provoking project, if you try to pay attention.”<br>(Skill Lit, B2) |
| Scary | Explicit use of the term ‘scary’. | “I think, it was very interesting, but at some places, a little scary to me. I don't know why.. I guess, it's the future, but the texts were at some places a little... strange? I mean, the whole text was strange, but at some places I felt a little uncomfortible [*sic*]” |
| Text quality | Descriptive or interpretive account of the text itself. Not including the act of reading and related and how that felt. | “I found the project interesting and creative. I think it has great potential. Currently the generated text still has too much repetition, rambling without any meaning and sentences that aren't connected.”<br>(Skill Lit, S3) |
| The future | Commenting on or speculating about a general sense of progress over time, a vague | “I think it's really interesting! A computer generated AI which writes 50k words |

| | | |
|---|---|---|
| | point in the future, general future technology and developments, future society. Includes explicit use of the term 'the future' and related, such as 'futuristic'. | novels is something which may revolutionize literature in the future" (Skill Code and Lit, S3) |
| Work and automation | References to a possibility, impossibility, or the implications of (writing or creative) work being performed by AI or computers. | "The project is very interesting, eve though I think creative things should be left for humans, not AI. There will be no place for humans once AI gets the ability to create such things. The text itself was pretty chaotic and difficult to read." (Skill Code and Lit, B1) |

#### 4.5.2.2. Distribution of category frequencies

The distribution of category frequencies stage was performed next to investigate in which ways the review content differs between the two conditions for *Novel B*.

The simultaneous coding approach made it feasible to create co-occurrence matrices to visualize and compare the frequency of the thematic categories in the survey and in the groups. Further statistical comparisons, such as inferential statistics, were not pursued because of the qualitative focus of this analysis.

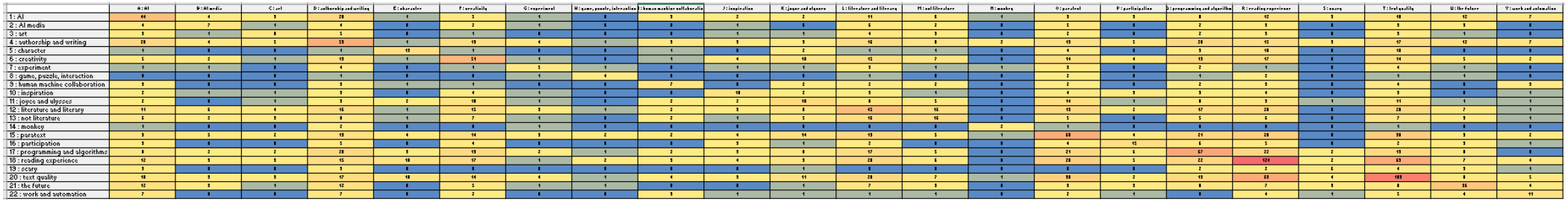
First, a co-occurrence matrix showing the frequency density of categories which occur together in reviews was created – *Figure 15: Category co-occurrence across all* reviews. This was useful to some extent to see which themes participant groups write about together. For example, both reviews in category *Monkey* co-occurs with *Authorship and writing*. While this makes sense intuitively, it was felt that a co-occurrence of categories only was underutilizing the ability to see differences between survey groups or the ability to contextualize the reviews more. For example, both *Monkey* reviews are from the *Skill Lit* and *Generated Text Condition* where participants weren't shown the generated text itself (except for short examples). This contextualizing information is potentially relevant. Therefore co-occurrence matrices visualizing the frequency density of categories for each group were also created.

It's important to note that category co-occurrences do not indicate sentiment or polarity, and all my analysis from this stage onwards bears this in mind. For instance, the *Skill Lit, S2* review that is coded for *Monkey* is simultaneously coded for the *AI* and *Authorship and writing*

categories. My own bias assumes that this co-occurrence of categories in one review might indicate a negative impression of the generated novel project, because the academic debates about AI and authorship that I encounter in my university environment discuss the potential negative impacts and risks of text generation technology and its uses. However, the review contradicts my bias and instead expresses a positive outlook about the prospect of a neural network 'writing' interesting books:

> "*I think that it's a really amazing project, never would have thought of it. It tries to create a neural network that can write a \"somewhat\" novel. It reminds me of the Infinite monkey theorem. In the end this project could be similar, but the network could be adjusted to provide better results. It could even write readble [sic] and interesting books given enough time, just like the monkeys. But not only that, it would be faster than a human writing the same novel. All in all, I think this is a wonderful idea and could provide astonishing results*." (Skill Lit, S2)

Figure 15: Category co-occurrence across all reviews



| | A: AI | B: AI media | C: art | D: authorship and writing | E: character | F: creativity | G: experiment | H: game, puzzle, interaction | I: human machine collaboration | J: inspiration | K: joyce and ulysses | L: literature and literary | M: not literature | N: monkey | O: paratext | P: participation | Q: programming and algorithm | R: reading experience | S: scary | T: text quality | U: the future | V: work and automation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 : AI | 44 | 4 | 3 | 28 | 4 | 5 | 1 | 0 | 3 | 2 | 2 | 11 | 6 | 1 | 3 | 3 | 8 | 12 | 3 | 18 | 12 | 7 |
| 2 : AI media | 4 | 7 | 1 | 4 | 0 | 2 | 1 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 5 | 0 | 2 | 3 | 0 | 3 | 3 | 0 |
| 3 : art | 3 | 1 | 0 | 5 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 4 | 3 | 0 | 2 | 0 | 2 | 3 | 0 | 3 | 1 | 0 |
| 4 : authorship and writing | 28 | 4 | 5 | 53 | 5 | 13 | 4 | 1 | 3 | 3 | 3 | 16 | 8 | 2 | 13 | 5 | 28 | 15 | 3 | 17 | 12 | 7 |
| 5 : character | 4 | 0 | 0 | 5 | 13 | 1 | 0 | 0 | 1 | 0 | 2 | 1 | 1 | 0 | 4 | 0 | 3 | 10 | 0 | 10 | 0 | 0 |
| 6 : creativity | 5 | 2 | 1 | 13 | 4 | 31 | 1 | 0 | 1 | 4 | 10 | 15 | 7 | 0 | 14 | 4 | 13 | 17 | 0 | 14 | 5 | 2 |
| 7 : experiment | 4 | 1 | 0 | 4 | 0 | 1 | 5 | 1 | 0 | 1 | 1 | 3 | 1 | 1 | 3 | 0 | 2 | 4 | 0 | 4 | 1 | 0 |
| 8 : game, puzzle, interaction | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 1 | 2 | 0 | 1 | 1 | 0 |
| 9 : human machine collaboration | 3 | 0 | 0 | 3 | 1 | 1 | 0 | 0 | 7 | 0 | 2 | 2 | 2 | 0 | 2 | 0 | 2 | 3 | 0 | 4 | 0 | 3 |
| 10 : inspiration | 2 | 1 | 1 | 3 | 0 | 4 | 1 | 0 | 0 | 18 | 2 | 3 | 1 | 0 | 4 | 3 | 3 | 4 | 0 | 3 | 0 | 1 |
| 11 : joyce and ulysses | 2 | 0 | 1 | 3 | 2 | 10 | 1 | 0 | 2 | 2 | 10 | 8 | 5 | 0 | 14 | 1 | 0 | 3 | 1 | 11 | 1 | 1 |
| 12 : literature and literary | 11 | 5 | 4 | 16 | 1 | 15 | 3 | 1 | 2 | 3 | 8 | 45 | 16 | 0 | 13 | 2 | 17 | 28 | 0 | 28 | 7 | 1 |
| 13 : not literature | 6 | 2 | 3 | 8 | 1 | 7 | 1 | 0 | 2 | 1 | 5 | 16 | 16 | 0 | 5 | 0 | 5 | 6 | 0 | 7 | 3 | 1 |
| 14 : monkey | 1 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 15 : paratext | 3 | 3 | 2 | 13 | 1 | 14 | 3 | 2 | 2 | 4 | 14 | 13 | 5 | 1 | 53 | 4 | 24 | 28 | 0 | 31 | 3 | 2 |
| 16 : participation | 3 | 0 | 0 | 5 | 0 | 4 | 0 | 0 | 0 | 3 | 1 | 2 | 0 | 0 | 4 | 15 | 6 | 5 | 0 | 2 | 3 | 1 |
| 17 : programming and algorithms | 8 | 2 | 2 | 28 | 3 | 13 | 2 | 1 | 2 | 3 | 8 | 17 | 5 | 0 | 24 | 6 | 57 | 22 | 2 | 13 | 8 | 0 |
| 18 : reading experience | 12 | 3 | 3 | 15 | 10 | 17 | 1 | 2 | 3 | 4 | 3 | 28 | 6 | 0 | 28 | 5 | 22 | 124 | 2 | 23 | 7 | 4 |
| 19 : scary | 3 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 6 | 4 | 3 | 1 |
| 20 : text quality | 18 | 3 | 3 | 17 | 10 | 14 | 4 | 1 | 4 | 3 | 11 | 28 | 7 | 1 | 33 | 2 | 15 | 23 | 4 | 103 | 8 | 5 |
| 21 : the future | 12 | 3 | 1 | 12 | 0 | 5 | 1 | 1 | 0 | 0 | 1 | 7 | 3 | 0 | 3 | 3 | 0 | 7 | 3 | 8 | 36 | 4 |
| 22 : work and automation | 7 | 0 | 0 | 7 | 0 | 2 | 0 | 0 | 3 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 0 | 4 | 1 | 5 | 4 | 11 |

The horizontal (numbered) and vertical (lettered) labels in *Figure 15* read: *(1/A)AI, (2/B)AI media (3/C)art, (4/D) authorship and writing, (5/E)character, (6/F)creativity, (7/G)experiment, (8/H)game puzzle interaction, (9/I)human machine collaboration, (10/J)inspiration, (11/K)joyce and ulysses, (12/L)literature and literary, (13/M)not literature, (14/N)monkey, (15/O)paratext, (16/P)participation, (17/Q)programmimg and algorithms, (18/R)reading experience, (19/S)scary, (20/T)text quality, (21/U)the future, (22/V)work and automation*.

Next, two co-occurrence matrices (one representing the data for each generated novel) were created using Nvivo to visualize the distribution of thematic category frequencies per group. These are seen in *Figure 16: Matrices showing frequency of codes per survey group for all conditions* The order of the categories was reordered manually by eye to present the data in a visually clearer way: categories with denser frequencies (represented by the warmer colors, such as orange) were ordered towards the bottom of the matrices, and the sparser frequencies were ordered towards the top (represented by cooler colors - blue indicates zero occurrences of a category in a group). This ordering by density thus clearly shows which thematic categories are referenced across all groups (such as *Text quality*, *Paratext*, and *Reading experience*), and which are concentrated or only occur in certain groups (such as *Scary*, which only referenced in the *Both Condition*).

Figure 16: Matrices showing frequency of codes per survey group for all conditions

| | A : code_only_reviews_b1 | B : lit_only_reviews_b1 | C : code_and_lit_reviews_b1 | D : code_only_reviews_b2 | E : lit_only_reviews_b2 | F : code_and_lit_reviews_b2 | G : code_only_reviews_b3 | H : lit_only_reviews_b3 | I : code_and_lit_reviews_b3 |
|---|---|---|---|---|---|---|---|---|---|
| 1 : monkey | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 2 : work and automation | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 |
| 3 : experiment | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 4 : scary | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| 5 : art | 0 | 2 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| 6 : game, puzzle, interaction | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 : inspiration | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 |
| 8 : human machine collaboration | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 9 : participation | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 1 | 2 |
| 10 : the future | 1 | 0 | 0 | 2 | 1 | 2 | 2 | 2 | 0 |
| 11 : AI | 0 | 0 | 2 | 0 | 1 | 1 | 1 | 1 | 3 |
| 12 : joyce and ulysses | 1 | 2 | 1 | 1 | 0 | 3 | 1 | 4 | 5 |
| 13 : character | 1 | 3 | 4 | 0 | 1 | 1 | 0 | 1 | 0 |
| 14 : programming and algorithms | 1 | 0 | 2 | 6 | 3 | 9 | 8 | 4 | 11 |
| 15 : authorship and writing | 0 | 3 | 0 | 1 | 6 | 3 | 5 | 3 | 5 |
| 16 : literature and literary | 0 | 3 | 1 | 1 | 4 | 5 | 0 | 5 | 6 |
| 17 : creativity | 0 | 2 | 2 | 4 | 5 | 4 | 4 | 6 | 6 |
| 18 : paratext | 2 | 3 | 5 | 5 | 3 | 3 | 2 | 5 | 10 |
| 19 : text quality | 8 | 11 | 11 | 1 | 5 | 2 | 6 | 7 | 6 |
| 20 : reading experience | 11 | 11 | 12 | 6 | 7 | 7 | 7 | 7 | 3 |

| | A : code_only_reviews_s1 | B : lit_only_reviews_s1 | C : code_and_lit_reviews_s1 | D : code_only_reviews_s2 | E : lit_only_reviews_s2 | F : code_and_lit_reviews_s2 | G : code_only_reviews_s3 | H : lit_only_reviews_s3 | I : code_and_lit_reviews_s3 |
|---|---|---|---|---|---|---|---|---|---|
| 1 : monkey | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 2 : work and automation | 0 | 0 | 0 | 1 | 1 | 2 | 0 | 1 | 2 |
| 3 : experiment | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| 4 : scary | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 |
| 5 : art | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| 6 : game, puzzle, interaction | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 7 : inspiration | 0 | 1 | 0 | 0 | 3 | 1 | 1 | 0 | 0 |
| 8 : human machine collaboration | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 2 | 0 |
| 9 : participation | 0 | 0 | 0 | 2 | 3 | 1 | 1 | 1 | 0 |
| 10 : the future | 1 | 1 | 0 | 2 | 1 | 7 | 3 | 3 | 8 |
| 11 : AI | 0 | 0 | 1 | 5 | 8 | 7 | 5 | 4 | 5 |
| 12 : joyce and ulysses | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 : character | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 : programming and algorithms | 0 | 0 | 0 | 4 | 2 | 5 | 3 | 4 | 6 |
| 15 : authorship and writing | 0 | 0 | 1 | 4 | 5 | 5 | 6 | 6 | 8 |
| 16 : literature and literary | 0 | 2 | 3 | 1 | 5 | 0 | 0 | 3 | 6 |
| 17 : creativity | 3 | 1 | 3 | 2 | 3 | 0 | 2 | 3 | 1 |
| 18 : paratext | 1 | 4 | 4 | 2 | 3 | 2 | 2 | 2 | 2 |
| 19 : text quality | 8 | 11 | 8 | 1 | 2 | 0 | 1 | 8 | 8 |
| 20 : reading experience | 9 | 10 | 12 | 1 | 4 | 6 | 2 | 2 | 7 |

#### 4.5.2.3. Category distribution results and discussion

To restate this study's question; In the *Q1 Literary* data, the only participant group which had a significant difference between the *Generated Text Condition* and the *Both Condition* read *Novel B*. In which ways does the review content differ between the two conditions for *Novel B?*

##### 4.5.2.3.1. Novel B Generated Text Condition

Categories Character, Art, Game-puzzle-interaction, Text Quality, and Reading experience are all more frequent in the Generated Text Condition. Differences between skill groups are also seen here.

Figure 17: Matrix showing frequency of Novel B codes per survey group for the Generated Text Condition and the Both Condition

| | A : code_only_reviews_b1 | B : lit_only_reviews_b1 | C : code_and_lit_reviews_b1 | G : code_only_reviews_b3 | H : lit_only_reviews_b3 | I : code_and_lit_reviews_b3 |
|---|---|---|---|---|---|---|
| 1 : monkey | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 : work and automation | 0 | 0 | 1 | 0 | 0 | 3 |
| 3 : experiment | 0 | 0 | 0 | 0 | 1 | 0 |
| 4 : scary | 0 | 0 | 0 | 0 | 1 | 0 |
| 5 : art | 0 | 2 | 0 | 0 | 0 | 0 |
| 6 : game, puzzle, interaction | 1 | 1 | 0 | 0 | 0 | 0 |
| 7 : inspiration | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 : human machine collaboration | 0 | 0 | 0 | 0 | 1 | 1 |
| 9 : participation | 0 | 0 | 0 | 0 | 1 | 2 |
| 10 : the future | 1 | 0 | 0 | 2 | 2 | 0 |
| 11 : AI | 0 | 0 | 2 | 1 | 1 | 3 |
| 12 : joyce and ulysses | 1 | 2 | 1 | 1 | 4 | 5 |
| 13 : character | 1 | 3 | 4 | 0 | 1 | 0 |
| 14 : programming and algorithms | 1 | 0 | 2 | 8 | 4 | 11 |
| 15 : authorship and writing | 0 | 3 | 0 | 5 | 3 | 5 |
| 16 : literature and literary | 0 | 3 | 1 | 0 | 5 | 6 |
| 17 : creativity | 0 | 2 | 2 | 4 | 6 | 6 |
| 18 : paratext | 2 | 3 | 5 | 2 | 5 | 10 |
| 19 : text quality | 8 | 11 | 11 | 6 | 7 | 6 |
| 20 : reading experience | 11 | 11 | 12 | 7 | 7 | 3 |

As is evident from the bottom of figure *Figure 17: Matrix showing frequency of Novel B codes per survey group for the Generated Text Condition and the Both Condition*, categories Reading experience and Text quality are the most frequent categories across the entire Generated Text Condition. With the absence of other paratextual pieces to write about, referring to the text itself and participants' impressions of reading it makes sense intuitively. Looking at the top of the matrix, it's clear that the left size which represents the Generated Text Condition has an area with higher category frequencies than the right side. There, the Both Condition are seen to be lacking these frequencies and therefore a clear difference is shown between the two conditions.

Category Character is clearly more frequently referenced by the Skill Lit and Skill Lit and Code groups in the Generated Text Condition. This means that more participants who reported having literature as one of their hobbies in the Prolific screening survey wrote more reviews which referred to characters or narrators. Indeed, several of these reviews refer to 'Molly' or she/her pronouns which suggest that the title Molly's Feed may have functioned paratextually to shape the interpretation of the generated text's content. For example:
"I have never red [sic] the Ulysses of Joyce, so i read the text without knowing anything about the \"story\". I think there is a climax of anxiety, the text start with simple thoughts and there is this excalation [sic] where Molly is very upset. I felt more anxious while reading" (Skill Code & Lit, B1)

Categories Art and Game-puzzle-interaction are seen only in the Generated Text Condition. The reviews coded to these categories are simultaneously coded with Reading experience and/or Text quality which means that these participants perhaps felt that they were relevant references to make in the same review. Interestingly, just as with some of the Character coded reviews it seems as if the absence of paratextual pieces may result in more reader attention to the qualities of the generated text. And yet, Joyce's Ulysses is nevertheless referred to:
"I found the project interesting. There seemed to be patterns and repetitions that appeared in the text. Studying the entire piece of text could reveal secrets about computer generated text. I was not familiar with the book, and I think that if I was I would have had more thoughts on this project" (Skill Code, B1)

#### 4.5.2.3.2. Novel B Both Condition

Categories *Joyce and Ulysses, Programming and algorithms, Authorship and writing, Literature and literary, and Creativity* are all more frequent in the *Both Condition*. Differences between skill groups are also seen here. These categories are discussed or expanded upon further in the next section.

While it's of course present in each of the conditions, category *Joyce and Ulysses* is more frequent across the *Both Condition*. This makes sense intuitively because the paratextual pieces in this condition explicitly refer to Joyce's book several times. This category is referenced most frequently by the *Skill Lit* and *Skill Lit and Code* groups, and this is perhaps in alignment with the *Q4 Hypotext Familiarity* survey question (albeit underwhelming) results from section, where the *Skill Lit* group had a significantly higher average mean response to being familiar with one of the classic works.

The frequency for categories *Reading experience* and *Text quality* are reduced in the *Both Condition*. This is interesting because the generated text (piece A) is present here, which suggests that less of the participants chose to refer to these categories as compared to the participants in the *Paratext Condition* who arguably had less potential content to write about due to the lack of paratextual pieces.

Interestingly, the frequency of category *Literature and literary* is highest in the *Both Condition*. This is unexpected because it was assumed that *Literature* would be referenced the most in the groups and condition where *Reading experience* and *Text quality* are most frequent (the *Generated Text Condition*, then) and because the latter two are commonly discussed in secondary and tertiary education literature classes – the former being something that the majority of participants are likely to have encountered. Unsurprisingly, where category *Programming and algorithms* is overwhelming more frequent in the *Both Condition* (where it's also the most referenced category overall), it's referenced the least by the *Skill Lit* group where participants reported not having coding skills during the Prolific screening survey.

#### 4.5.2.3.3. Literary value in the reviews discussion

Progressing onto the next study question; What are some of the opinions expressed about literature and literary value between the two Novel B conditions?

##### 4.5.2.3.3.1. Novel B Generated Text Condition

As already discussed, the *Generated Text Condition* has a smaller number of reviews referring to the *Literature and literary* category. The *Skill Code* group makes no reference to *Literature,* but out of 4 reviews, 3 of these belong to the *Skill Lit* group. 2 of these reviews expand upon their opinion about *Molly's Feed* by referring to Joyce's *Ulysses* to varying degrees, thus highlighting the intertextual or hypotextual relationship between them, and its potential impact on value – namely interest and creativity.

The following review is coded to the *Literature, Paratext, and Text quality* categories. It writes that the generated text is not literature because it lacks the conventional structure of, presumably, a story. The participant assumes that intertextual knowledge about *Ulysses* may have improved their reception along more general dimensions (interestingness), but no mention is made of this potentially changing their view on the 'literature' status of the generated text:

> "*I found it quite interesting, I would assume if I had more knowledge about the book it would be even more interesting. Overall, I found that the project seemed more like a mashup of sentences that aren't necessarily related but can make sense on their own. It seems a bit like a transcript of someones' tweets. I would not, however, consider this literature as it lacks the structure of a text, i.e., introduction, development and conclusion.*" (Skill Code and Lit, B1)

Another review coded to *Literature, Art, Authorship and writing*, *Creativity, Joyce and Ulysses, Reading experience,* and *Text quality* unpacks the question of value further and expands on how it may be impacted by the context or information with which it's presented. The latent theme here is that of authorship where it's implied that *Ulysses* is intentionally written in a particular style by a person, and *Molly's Feed* appears to be contrasted with this; indeed, Bhatnagar nor any reference to a creator, programmer, or designer is made. Instead, it appears that the generated text is ascribed to "…the powers of AI". However, the participant notes a stream-of-consciousness quality in both texts. They appear to point out that studying *Ulysses* in a formal setting impacted their impression of it, and it seems that they are now on the fence about the creative value of the textual quality in *Molly's Feed*, because they seem to consider that the "…knowledge that the text was computer-generated…" Might similarly be impacting their impression of it:

> "*I found it disjointed and hard to read through. I didn't make the connection with Ulysses until it was pointed out. I read and studied Ulysses in college and approached it with a particular mindset, i.e. that it was a creative, artistic piece of work that was deliberately written as a a stream-of-conscious narrative. Approaching Molly's Feed with the knowledge that the text was computer-generated, I judged the stream-of-consciousness, disjointed feel of the text not as creative but as a limitation of the powers of AI. Now I'm kind of questioning my position a little.*" (Skill Lit, B1)

Although not referring to the *Literature and literary* category, additional *Novel B* reviews in the *Both Condition* assume that knowledge about *Ulysses* would have improved or expanded their impression of the project. Thus, these reviews also indicate their acknowledgement of a potentially impactful paratextual relationship between Joyce's book and the project:

"*It's an interesting way to play with your mind and your reading comprehension ability. Maybe if I had read the book I would have understood better, though*." (Skill Lit, B1)

"*I found the project interesting. There seemed to be patterns and repetitions that appeared in the text. Studying the entire piece of text could reveal secrets about computer generated text. I was not familiar with the book, and I think that if I was I would have had more thoughts on this project*." (Skill Code, B1)

"*Kind of interesting but it's probably impossible to understand the informations provided in the text, without grasping the context, which requires reading the book*." (Skill Code, B1)

#### 4.5.2.3.3.2. Novel B Both Condition

As already discussed, the *Both Condition* has a larger number of reviews referring to the *Literature and literary* category – 11 in total. As in the other condition, *Skill Code* group makes no references to that category. *Programming and algorithms* is much more frequent in this condition as well, with 24 references. Category *AI* is also more frequent here, although with a smaller number of total references – 5. The *Both Condition* reviews coded to the *Literature and literary* category were read again, and the prevalence of the latter two technical categories lead to me identifying a pattern and developing the theme of literary value being relocated to technical value.

4 reviews in the *Both Condition* appear to reject the idea that the project could have literary value. These reviews then seem to (re)locate the project's value in technical or computational areas. While there are of course reviews which refer to the project in a way which doesn't reject literariness, this theme of relocating literary value when the paratextual pieces are presented can be developed from several reviews.

Reviews which appear to reject literary value and relocate the project's value elsewhere range from a gentle and nuanced rejection of the project's literary value to a confident, hard rejection. Not all reviews are necessarily negative about the project, though. Many of these reviews are short and tend to refer to topics or areas without necessarily explaining or elaborating on them, but they nonetheless convey what the participants felt are important points to make when asked what their personal opinion about the project is. When these reviews express doubt about the project's literary value, the doubt appears to be anchored in the nonsensical style or quality of the generated text. Furthermore, within the same reviews, doubts or clear rejections of literary value can be seen next to suggestions or speculations about the project's technical value and value to technology related areas or people. I therefore interpret this rejection of literary value and suggestion of technical or computational value as a relocation of value away from the literary.

One reader's gentle rejection of literary value not only comments on how the project serves as an example of creative technology or Artificial Intelligence in the present, but also begins to acknowledge and grapple with the intertwined and complicated question of triangulating value between the arguably well-established literary value of the hypotext, the generated text's similarity to it, and the reader's own preferences. The review concludes by locating the project's value in a general technology area. The use of the phrase "...tech-y side of things" may suggest that the project does have a multifaceted or interdisciplinary quality to it, but ultimately it seems that the reader moves the project's value away from an unsure literary status, and to a more comfortable (albeit vague) status as a piece which could be valued for its technological significance to Human-Computer collaboration and creativity:

> *I have never finished reading Ulysses because I just could not get through it... but I immediately thought, "this is like Ulysses" as I started reading 'Molly's Feed' (although I didn't make the connection with the character). Comparing the first two pages to the final one, I did find that it felt more hastened, so I was pleased to see that it was designed that way. I suppose the project is an interesting experiment about human creativity vs. computer "design", and it's also fun to see human creativity work kind of alongside AI. I don't think I could read all 160+ pages of 'Molly's Feed' so to be honest I'm not sure about the literary merit of the text (but then, I couldn't read Ulysses either...), but it does have its own value in more tech-y side of things perhaps.* (Skill Lit, B3)

In contrast to the previous, the following review appears to be very confident in its reasoning for why the project does not have literary value, and instead locates value in the project's potential to serve as an exemplar for future computational projects. However, it's unclear whether it's only the nonsensical quality of the generated text which devalues any potential literary status, or if the "...code based" method also contributes to this:

> *The project is rather technical and not imaginative because literature has to do with "imagination,especially creating images inside the human mind.It is technical in the sense that it just creates a novel that is gibberish but on the other hand it creates a precedent for code based project in the future."* (Skill Lit, B3)

The next review again considers the interest or value in the computational aspect of the project, while being clear in its rejection of the project's literary value based on the generated text's nonsensical quality:

> *"It was an interesting experiment to see how computer generated stories could be created. However, at this level, from a literature point of view, the result was gibberish."* (Skill Lit, B3)

The final review from the *Both Condition* which relocates value expresses the theme in a more latent way. I interpret the review's use of the terms 'novel' and 'surrealism' as referring to literariness, especially since the arguments expressed in it are extremely similar to others which reviews make the same points while explicitly referring the literature or literariness. This review appears to acknowledge some degree of merit based in a technology area, and perhaps a limited degree of computer agency, but the 'gibberish' quality devalues it. Interestingly, this devaluing extends also extends to Joyce's novel:

> "*It's interesting to see how far AI can reach, but in my opinion a novel should always have some kind of human intervention. It's true that for surrealism it may be that a computer can do what a writer could. Both examples sound like gibberish to me. Pardon me James Joyce."* (Skill Code and Lit, B3)

There are reviews which run counter to the relocating of literary value to technical areas theme. For example, while this review also describes the text as gibberish, the reader nevertheless makes a point of explaining why they enjoyed the project regardless:

> "*I think it's an interesting project that involves two areas that I didn't think it could be joined together - programming and literature! Even if the text itself is gibberish, the whole process behind it's construction is very fascinating. Besides, the text is rather funny and I had a great time looking into it.*" (Skill Code and Lit, B3)

#### 4.5.2.3.3.3. Novel B Paratext Condition

Interestingly, the rejection of literary value and relocating it to a general technical value theme is also present in 4 reviews also coded to the *Literature and literary* category in the *Paratext* Condition. Although participants were not shown the generated text in this condition, they were presented with the same paratextual pieces as the *Both Condition* where the theme was first identified. Therefore, although the rejection of literary value was also seen in the *Generated Text Condition* reviews, it's the relocation of value away from literariness and to a technical area that is seen in reviews written by readers who have been shown the paratextual pieces (pieces B to G). This similarity seems to align with the *Q1 Literary* survey question results where there isn't a significant difference between the *Paratext Condition* and the *Both Condition*.

While the review below does ostensibly relate the project to art and literature, the reader doesn't seem to consider the relation valuable because of perceived authorship. It is implied that people ("how far we can go with creating AI") are responsible for the 'AI' that in turn is perhaps responsible for a part of the project. But the focus then stays on an opposition between "real humans" and code. It's unclear whether the latter is referring back to the AI or to the

aforementioned “we”. In any case, the importance of passing for “real humans” is cemented with the Turing test reference. But in the end the technical themes are not rejected – they are nevertheless considered valuable in a general area of technological progress:

> *“It's a nice project that tries to explore how far we can go with creating AI that then creates art, such as literature. It's relatively easy - with some basic coding skills and time and motivation - to \"write\" a \"novel\" created by AI. But as far as I have seen, none of these novels would pass an adapted Turing test, i.e. no one could be mistaken thinking that these novels were written by real humans and not programming code. That being said, these projects are important and progress is being constantly made.” (Skill Code and Lit, B2)*

The following review pitches its argument from literature grounds and gives the human/machine authorship opposition as the reason why the project cannot have value as literature. But again, the review takes care to make clear that the project has some sort of value – just, presumably, as a not ‘real’ code-written book:

> *“As a books enthusiastic I'm not a big fan of this project, because it changes the whole purpose of literature. The story does not make sense. Codes can not write like humans. They do not have our minds, our creativity. However, I don't think this is a bad idea. I just think we shouldn't take these types of book seriously or at least we should not give them the same value as real books written by real people.”* (Skill Lit Only, B2)

The following two reviews both refer to the topic of literature and literary but in a way that minimizes these values. This minimizing is not done for the technical value that they both reference:

> “*If I do not clearly understand I cannot have a precise opinion, mostly questions. Seems interesting but, for me, meaningless. A divertimento from a literary point of view or maybe a meaningful technical excercise from a technical point of view.*” (Skill Lit, B2)

> *“The project represents a creative approach to Ulysses, and seems to be intriguing technically. I believe the output might be an interesting piece, not a literature master piece, but interesting”* (Skill Code and Lit, B2)

Finally, the following review is an example of reviews in the *Paratext Condition* which do not express the relocation of literary value to technical value theme. It makes a connection between the project and poetry styles, and wonders if a careful reading method that is successful with the poetry might lead to a positive experience of the project. Interestingly, the ‘technicalness’ of the project is not drawn to – rather literary discussion is:

> *"I don't know anything about coding and programming, but I find interesting to know about books generated by a computer program. I have had the opportunity to read alternative, baroque or arbitrary poems that simply seem to make no sense, but by reading them carefully you can find beauty and meaning. I wonder if the same can apply to these novels generated by a computer code, and the discussion that can develop in the literary community." (Skill Lit Only, B2)*

There is of course the irony of referring to the project as "books generated by a computer program" when every piece in the *Paratext Condition* from pieces B to G have been written by people (even *The Code* and *The code repository* contain explanatory documentation written by Bhatnagar). But although this appears to be common across all reviews, this contradiction does not appear to be written about. This is not surprising, because the participants did not have an extended amount of time to engage with and think about the project. This opportunity is afforded to participants in the workshop study.

## 4.6. Limitations of qualitative and quantitative portions of the study

As discussed earlier in this chapter, unlike in a laboratory where conditions can be controlled and fully documented, the online context of the reading study affords a limited degree of researcher control. This was mitigated by asking participants to stay on the study website and to compete the study on a computer and not a smartphone but could not be ensured.

Because the participants were aware of the PhD project nature of the study via the informed consent form, participant response bias is an inescapable effect in this study, although it is not measured. While the consent form made clear that data was anonymous and the wording in the survey stated that the participants' "personal opinion" was sought, it is nevertheless possible that some responses could have been charitable in sentiment or judgement because they assumed that the study was part of a individual person's PhD project and the PhD research would read the reviews. Nevertheless, the study contains reviews of a scathing nature which suggest that a breadth of opinions have been able to be collected.

As previously explained intercoder reliability is not able to be engage with because I am the only researcher on this PhD project. To mitigate this single-coder limitation, detailed examples and quotes reviews have been included as part of the Qualitative Content Analysis study for transparency. Although the categories and theme are developed from the review data, it is not known whether they could also be developed outside of the study from a wider discourse about new readers' perceptions of generated novels.

With regards to linking the quantitative and qualitative results it is important to remember that the free-response question asked participants about their opinion of the project and not, more specifically, to necessarily explain their survey responses. Based on the terms used in many of reviews it's apparent that the survey questions did have a priming effect on the written responses, but this is expected. Indeed, it is perhaps somewhat helpful as alternatively asking the participants to write a review before going through any of the survey questions could have risked situations where participants would have trouble thinking what to write about.

The free-response review data was collected in 2021 from participants who were recruited through the Prolific platform. While generated text and GPT models have been known and discussed in the AI research community and in some media before 2021. Therefore the data collection predates the 2022 launch of OpenAI's ChatGPT[43] which, along with other consumer-facing Large Language Models, have since become a frequent topic in the media and a publicly available tool. This means that some content of the reviews may be different if the data were collected now. Indeed, even if hypothetical new participants would not have knowingly engaged with commercially available LLM generated text, the media and other pop-culture attention to the topic would very likely have (arguably paratextually) shaped their interpretation of the projects in the study. Thus, the data and results could be interpreted as a limited 2021 snapshot in the larger context of layperson's opinions and reception of (creative) text generation and the related technology, and how paratext (media, explanations, marketing copy, training data) might impact opinions, reception, and even user experience.

## 4.7. Integration of results and chapter discussion

This section summarizes the results of the statistical analysis of the survey questions and the qualitative results of the free response reviews, and discusses how these results support or differ from each other. The research questions are revisited based on these results and the rationale for the second study is discussed. The discussion is generated while keeping in mind that all results have been gathered from a controlled online reading experiment using a survey participant recruitment platform, so the generalizability of results are of course limited. The novel contribution made by the survey study nevertheless stands: to empirically study and evidence the impact of paratext.

The results from section 4.4.2. *The Impact of Adding Paratext: The Generated Text Condition vs. The Both Condition* indicate that with the general exception of the *Q1 Literary* survey question,

[43] https://chatgpt.com/

there is a significant increase in the mean responses to the overwhelming majority of value dimensions when the paratextual pieces are present in the project. This raises a further question: why are participants' judgements of the generated novels' literary value barely impacted by the presence of the paratext? Indeed, differences between the *Novel B* conditions review content can be shown, but not differences between the survey's *Literary* value judgements. While it's possible that quantitative methods are not sensitive to aspects which capture the differences between the *Literary* question conditions, this seems unlikely because *Creativity* (another complex and culturally laden concept in the survey) does appear to be impacted by the paratextual pieces. The Qualitative Content Analysis results show that groups which read *Novel B* reference the *Literature and literary* category in their reviews more frequently when the paratextual pieces were present – the same is true for groups who read *Novel S*. therefore it appears to be a concept that participants felt was relevant to introduce and discuss or even refute. Curiously, while Genette's paratextual conceptualization has a disciplinary tradition rooted in literary studies, it seems rather ironic that the current data suggests that literariness is distinct in its reluctance to be influenced by paratext. Or at least, based on the influence of the kinds of paratexts examined in this study. This is unexpected and it may be because the digital and formal study context of the survey is too far detached from the more familiar expectations, interactions and trappings of what some people may associate with reading novels and literary works. Further qualitative research where participants might explain their valuation judgements around the concept of literature and literariness in more depth is required to understand these results better.

While the analysis of the survey questions was not able to show significant differences in skill groups across conditions which were meaningful in context, the Qualitative Content Analysis did show differences between participants skills which seem intuitive based on assumptions about domain knowledge. Groups comprised of participants who are assumed to have literary reading skills (*Skill Lit* and *Skill Code and Lit)* had a higher frequency of references to the *Character* category in the *Novel B Generated Text Condition* than the *Skill Code* group do. In the *Novel B Both Condition*, it was the groups with reported coding skills (*Skill Code* and *Skill Code and Lit*) which had a higher frequency of references to the *Programming and algorithms* category. As statistical significance testing was not used to compare the category frequencies across conditions, the quantitative and qualitative data is not directly comparable. But the differences between skills indicated in the reviews analysis do suggest that qualitative methods are able to capture differences between participant skills and paratextual conditions.

The results from section *4.4.3.The Impact of Removing the Generated Text: The Paratext Condition vs. The Both Condition* indicate that there is no significant difference in the mean responses to the value dimensions. To a small extent this uniformity is perhaps mirrored in the

study of the reviews, where the relocating literary value to technical value theme that I develop in the *Both Condition* can also be seen in the *Paratext Condition*.

Overall, the participants' positive familiarity with *Ulysses* and *Pride and Prejudice* cannot be shown to impact literary value judgements (except in the case of *Novel B*, which is discussed further below). This is interesting because several reviews belonging to *Novel B* in the *Generated Text Condition,* where no additional paratextual pieces were shown, appear to assume that knowledge about *Ulysses* would have improved their reception of it. This may well have been the case for individual readers, but it runs counter to the section 4.4.4. *Familiarity with the Hypotext* survey results. From a broader group perspective (rather than an individual reader one) this is possibly an example of participants' consciously made opinions (as seen in an explicit statement or response to a direct question) being different from their more latent responses. In short: perhaps a difference between explicitly believing A but then doing B without necessarily realizing it. So to reiterate; a difference in believing that knowledge about *Ulysses* might increase value judgements versus familiarity with *Ulysses* turning out to have no impact on judgements of literariness. While the reviews did of course not state that they assumed the literary status of the project might increase if they were more knowledgeable about Joyce's novel, it is an assumption on my part that perceptions of literary value surely could have increased. Although I do not personally believe that the concept of literariness is tied to any hard and fast requirements, I did assume that certainly the modernist and post-modern aspects (or even simply parallels with) the generated texts and their paratexts surely could have moved the needle on literary judgements. Certainly, considering that such connections with literary themes are no less a crude literary mechanism or laundry list of literary properties than some of the other requirements that reviews have identified.

Focusing on the *Generated Text Condition* and the *Both Condition*, the results from section *4.4.5 Ranking the pieces*, suggest that the NaNoGenMo page (piece E), Media page (piece F), and the Creator's progress page (piece D) are ranked by participants as the most important group of paratext pieces for both helping to understand the project and for helping to make it interesting. Conversely the Code (piece B), the Code repository (piece C), and the Input (piece G) were ranked as the least important group of pieces for understanding the project and for interest. In these ranking questions participants make explicit decisions about which pieces they value; In the *Both Condition*, when the Generated text (piece A) is present in the ranking list, it's low mean rank score clearly suggests that it has been explicitly chosen as the most important piece for helping to make the project interesting. However, the results for the *Q4 Interesting* survey question in the *Generated Text Condition,* where only piece A is present, show that the mean response across all groups is less than the equivalent of 'Somewhat agree' on the ordinal scale (the highest average mean response in this question group is approximately 4.9).

This suggests that overall, most participants likely didn't agree that the generated text was interesting. But it is not the case that the other paratextual pieces were similarly scored low for interest; the mean response significantly increases to the positive pole in the *Both Condition* which suggests that the projects' interestingness perhaps lies with the paratextual pieces. This may appear to contradict the ranking results where the generated text is judged to be the most important for interest, although the two survey items were admittedly not asking the exact same question. The most intriguing aspect to this is that, unlike the *Q72 Rank Interesting* question, *Q4 Interesting* captures participants' latent value judgements about the paratextual pieces. Latent, in the sense that although participants were not explicitly judging individual pieces, their latent judgements can be deduced by comparing the conditions. So this could be a second example of an area where participants' explicit and latent judgements about paratext may differ.

Further, it is curious that although pieces A, E, D, and F are ranked as the most important for understanding and interest, the *Both Condition* reviews seem to not only write about piece A the most, but also frequently refer to the project's programmed or algorithmics status. On the one hand this is a topic that perhaps aligns most closely with piece B, C, and G which were ranked as the least important for understanding and interest. On the other hand, the pieces which rank for higher importance do also detail the projects' status as a generated work although the human involvement and direction aspects of E, F, and D do not appear to be highlighted or discussed as much as the computational aspects. For example, participants had the possibility of writing more about Bhatnagar's choices, the twitter users who wrote the tweets, composing with found-text, or adaptations - but the majority of participants chose not to. Indeed, the prevalence of writing about the generated text (piece A) could be because conventional approaches to secondary and introductory tertiary education literary studies emphasize the close-reading of texts and discussion of literary devices (as described in Culler's 'The Closeness of Close Reading', 2010), although this is by no means the exclusive approach for an entire semester. It might also be because it was the first piece to appear in the reading study.

### 4.7.1. Considering the research questions

The PhD project research questions are revisited and considered with the current results. My general research question asks: in what ways does paratext influence potential readers' interpretation and reception of a generated novel? This is broken down into three specific research questions:

*1. Which paratextual elements have a central role in influencing the understanding and reception of a project?*

This is answered by the *Q71 Rank Understanding* and *Q72 Rank Interesting* questions where pieces A, E, F, and D (Generated text, NaNoGenMo page, Media page, and Creator's progress page) were explicitly ranked as the most important for helping to understand the project and helping to make it interesting. And yet, the *Generated Text Condition* responses for the *Q2 Interesting* survey question suggest that on a more latent level, piece A on its own without any other paratextual pieces is arguably received rather poorly in terms of interestingness. To add further complexity, *Ulysses* or 'the book' was referred to in the reviews several times so it was presumably considered to be worth writing about, despite a sample of the classic text (piece G, the Input) being ranked as one of the least important pieces overall. Additional qualitative research may be able to unpack this further.

*2. Which reader skills and experience affect the influence of these paratextual elements?*

This remains in the process of being answered, because the current analysis is unable to show a compelling difference between presumed literary reading skills and reported coding skills, and the paratextual conditions. From a statistical perspective, previous knowledge or familiarity with the classic work which bears an intertextual relationship with the project could not be shown to impact literary value judgements. However, several reviews did assume that knowledge or understanding about programming, or knowledge or familiarity with *Ulysses* would have improved their impression of the project. This suggests that at least some participants expect paratext to impact reception when it is understood. Interestingly, this expectation is not able to be supported by the quantitative results in this study.

*3. Which technical and cultural characteristics of a generated novel affect the influence of these paratextual elements?*

This remains unanswered because the qualitative data items in this first study are too short to be able to expand analysis into the more complex areas of culture and technology. However, literature and literariness, and conceptualizations of the novel and the book do appear to be deeply set by participants' cultural beliefs and expectations. This is an area that needs to be researched further because it is not clear whether these expectations can be satisfied with different paratexts.

# Chapter 5 Workshop study

This chapter presents the second study of this research project: the reading discussion workshop study. It is comprised of two components. The first is a workshop inspired by reading group practices. For this I designed and arranged the printing of physical paperback books which contain the generated text of the two NaNoGenMo projects. For the first component of the study, the books were independently read and annotated by the workshop participants in their own time, who then met together in the workshop to discuss their reading and impressions. I ran this workshop component as a semi-structured group interview.

For the second component, I transcribed the workshop audio recording and performed Deductive Qualitative Analysis on the data to develop themes and a conceptual model. The model describes the different potential ways that a generated novel can be interpreted in relation to the paratextual pieces. The model works towards understanding the role that paratext plays along with reading strategies in the interpretation and reception of printed generated novels. This is intended to be a working conceptual model rather than a definitive one. Finally, this chapter concludes by answering to the research questions based on the discussion workshop data.

In terms of study design, the discussion workshop was the site and means of data collection, and the Deductive Qualitative Analysis is the component where the data is analyzed. Analysis focuses on the readers as individuals, and not as a social group of readers nor on group dynamics.

## 5.1. Workshop motivation

The design of the workshop study was inspired by reading group practices. Specifically, an online academic reading group that I joined for some sessions reading James Joyce's 1939 *Finnegan's Wake*, where a shared digital copy of the text was annotated by group members. I note that although Henrickson (2019) runs focus groups where participants discuss a computer generated news article, I do not consider this to be related work. Both our study items and research questions differ; Henrickson focuses on how readers attribute authorship to functional news report texts, and not on generative literature which prioritizes creativity and experimentation above factuality and functionality.

Strategically, the workshop study is a means of porting the generated novel as a print book into an arguably 'real-world' environment (albeit under study conditions). Here, the affordances of

the physical paratext and the generated novel's potential fit with existing ways of engaging with written works can be tested, and its (or the readers') limits be articulated. In a very literal sense, then, the generated text is moved further away from its digital GitHub publishing platform, and further away from the online formal survey context; one might say that the generated text is severed from its NaNoGenMo and GitHub paratexts. It is instead installed inside the physical paratext of the codex form and presented in a reading group context. Thus, I took on the role of editor and (self)publisher to paratextually relocate the generated text away from an arguably 'technical' context and to, as I presumed, a much more literary context. However, as will be explained in section *5.6 Answering the research questions,* literary value was also challenged here.

Printing generated novels as a physical books has already been done many times both commercially and by NaNoGenMo authors for personal use. For example, the generated novel detailed as a NaNoGenMo project example in *Appendix B* is printed as a hardcover book. Therefore, the most unique aspect of this study is not in the printing of the generated novels, but in primary data collection and analysis of readers responding to generated novels as received through the trappings of physical print paratext and a reading workshop environment. This shift to a physical mode and a discussion activity is in response to the findings from *Chapter 4 Reading experiment and study set up* where the results were not able to show that paratext impacts perceptions of literary value. Also, in contrast to the previous study where responses were gathered from a large number of laypersons, the workshop participants were a group of 7 people. These included 2 professional writers who had previously created works of electronic literature, and subject specialists in areas relating to interactive and entertainment media research.

## 5.2. Workshop set-up and semi-structured group interview

### 5.2.1. Creating the print versions

#### 5.2.1.1. Practice-based dimension

The preparation of the printed generated novels, and perhaps the discussion workshop itself have a link to my practice-based work. I have participated in 3 NaNoGenMo projects: issue #57 *Color Visions*[44] (2019) independently, and issue #57 *The Apollo and the Dragon-King: wild and*

[44] https://github.com/NaNoGenMo/2019/issues/57

*semi-wild rabbits*[45] (2020) with collaborators, and issue #58 *The Year 2020: Now oil the joints of my hand at that moment that there is no love*[46] (2020) with collaborators. The latter two projects were part of generated novel making workshops that I designed and ran with my friend and collaborator Dr. Noriko Suzuki-Bosco. Each of the generated novels created in the workshops can be found through NaNoGenMo's GitHub pages along with the collaborating workshop participants' names. *The Apollo and the Dragon-King* was created with a group of fellow Winchester School of Art Postgraduate Research students, and it is the most developed example of my practice-based work with the generated novel form. As that workshop took place during the global pandemic, we decided to amuse ourselves and celebrate our project with a generated novel launch party online where we playfully negotiated and experimented with how we could engage with the generated novel.

Because of the many animals referenced in the input data, this resulted in a prominent recurring animal theme in the generated text. We therefore agreed to 'fancy-dress' as animals for the book launch. Over video conferencing, we each explained the in-text inspiration for our animal outfits and I then kicked off the launch with a dramatic reading of the *The Apollo and the Dragon-King*. We then discussed the aesthetics, entertainment, and potential functional applications of an 'audio book' recording of the generated novel. Finally, we played a textual scavenger hunt game to find as many unique animal references as we could within a given timeframe. Certainly, *The Apollo and the Dragon King* is by no means the crème de la crème of the generated novel world, and it was perhaps its novelty to most of the participants and the fun, social aspects of creating it and engaging with it that rendered the experience enjoyable. I further pushed and explored the boundaries of how generated novels could be engaged with by printing it as a hardback book, and then recording almost an hour of myself reading the first few chapters of *The Apollo and the Dragon King* aloud in an exaggerated comical voice which was based on my impression of the verbose and rambling '19th century-ish' textual quality (I do indeed find myself quoting phrases from the generated text to this day). The printed hardback and audio was presented as *The Apollo and the Dragon-King: wild and semi-wild rabbits*. *Sound Installation Version* (Lesia Tkacz, 2023) at the Winchester Gallery show *More Than a Thesis: Different Approaches to PGR Study* from 29 June - 29 July 2023.

Thus, my initial ideas for this research project's discussion workshop can be linked to the practice-based work that I have carried out and developed (and further intend to) in parallel to this research project. Indeed, as with the generated novel making workshops in 2020, Dr.

---

[45] https://github.com/NaNoGenMo/2020/issues/57

[46] https://github.com/NaNoGenMo/2020/issues/58

Suzuki-Bosco also greatly contributed to the second half of the discussion workshop where she gave an introductory presentation about her Artist's Books specialism in order to help contextualize the second task. This task was a creative reworking and crafting session where participants used spare copies of the printed generated novel books as physical crafting materials to create their own piece in response to their impressions of the generated novels. This creative reworking phase will not be discussed further as it is no longer within the scope of this PhD project, but it can be part of a future publication.

#### 5.2.1.2. Bookishness

I use the term 'bookish' to describe some of the editing and presentation decisions I made while creating and arranging for the printing of the physical paperback generated novels. Van Dijk also uses the term in *The margins of bookishness: Paratexts in digital literature* (2014) to talk about electronic literature on the web, but they don't define the term. Contemporary literature and culture scholar Jessica Pressman greatly further develops the term in their book (2020), where it has been used to describe book-like and book referential objects such as bookshelf wallpapers, smartphone covers, decors, and so forth. It further encompasses a cultural obsession with the print book object. Although this is not exactly how I use the term to describe how generated novels in print form tend to borrow heavily or even emulate the form of the traditional print novel, Pressman's link to book-like artefacts is nonetheless useful. Pressman's

"...bookishness is about class and consumerism. It is about constructing and projecting identity through the possession and presentation of books. the difference here is that unlike the shelf of leather-bound but never-opened canonical texts, books no longer need to be owned or physically displayed in order to do the work of self-construction" (p. 9)[47].

Thus, attempting to construct and project a generated novels' literary identity through a physical paratext presentation is what I aimed to do when I was editing and constructing the print versions; I was constructing a bookishness of the generated novel as others have previously done. Pressman also connects literariness and bookishness. In their example of a *Pride and Prejudice* duvet cover and the actual novel, Pressman identifies both as examples of bookishness to highlight a sense of literariness:

"Neither the duvet cover nor the words "Pride and Prejudice" are the book *Pride and Prejudice*, instead, the bedspread and the beloved novel, I would say, combine to serve as an example of

[47] While this could certainly be an intriguing critical lens with which to view NaNoGenMo, generated novels, and the reasons for creating them, this is flagged for future research as it is currently outside the scope of this research project.

bookishness. Put differently, we see bookishness in the books we read; it is a literary mode even as it is also a way to commodify (as in the duvet cover) literariness" (p. 10).

Reasoning along similar lines, I assumed that having *Victorian* in physical bookish form, and the classic novel's first line starting each chapter, plus giving the participants an extended reading time along with the annotation task and discussion activity would increase the generated novel's perceived literary value. To some extent it did, but this seemed limited. This is discussed in section *5.6.5. Answering Study Question 2*.

#### 5.2.1.3. Editing and printing

To the best of my knowledge, no known print book versions of *Molly's Feed* and *Victorian* had been created prior to the workshop study, so I made further document layout, typographical, and print publishing choices in order to prepare the generated texts for print as physical books. This process gives a practice-based dimension to my research.

Because of the research aims of this study, I chose the paratextual trappings of the print books to appear plain or neutral in appearance like inexpensive paperback novels. I did this to try to direct participants' impressions to the printed generated text, rather than on aesthetically compelling book covers and other paratextual elements which are expected to impact reception and interpretation of traditional print novels. These 'plain novel' design choices were made with the assumption that it was not possible to design a print book whose physical and typographical elements have no impact on readers' impressions. Indeed, print works (including generated novels published in print) either exist within a consumer market with a range of marketing and production budgets, or exist outside of the consumer market as self-publish books, zines, chapbooks, unauthorized 'pirated' books, and so forth. Potential readers understand these differences because they are also consumers who are primed to make value judgements, and so designing a physical book whose paratextual elements have no impact on readers is unrealistic. For this reason I aimed to design the print versions of *Molly's Feed* and *Victorian* in the aesthetic style of a commercially published plain and inexpensive paperback novel; not a high marketing budget best-seller with an aesthetically compelling cover designed to attract readers, nor a tell-tale self-published work made using the default LaTeX or word processing font and layout with a limited use of graphic design principles. This aim was successful in some aspects but struggled in others (such as my mistake of adding extra blank pages to *Victorian*). As detailed in *Appendix A, Table A3: Changes Made to Pieces*, the generated text of both *Molly's Feed* and *Victorian* were previously edited to alter the document layout so that they would resemble each other more closely, and indeed to more closely resemble a typical NaNoGenMo PDF. I made further document layout, typographic, and other print publishing choices in order to prepare the generated texts for print as a physical book.

The typesetting system LaTeX[48] was used to create the digital layout for the print books. Although LaTex is free software that is predominately known for its use in document preparation in the sciences and engineering, it is also frequently used in NaNoGenMo projects as a method of preparing PDFs which present the project's generated text. I used an existing template from the *memoir* LaTeX class[49] (similar to the concept of a package in R, as explained section *4.3. Statistics and computational tools*) which supports the preparation of a document in a 'fiction book' form. The elements which I used from the template and filled with the appropriate information for each book were a half-title page, a title page where I added the name of the project and its creator's name, a copyright page with the project's original publication year plus "Print Edition 2022", and a Creative Commons Attribution ShareAlike 4.0 license (to preserve the sharing spirit of NaNoGenMo), and a table of contents for *Victorian* only. The generated text was copy-pasted into the LaTeX template and page numbers were added from the first to the last page of generated text. The publishing platform that I used added an additional final page after the last page of generated text which contains a barcode and a series of alpha-numeric codes, such as 'PB' for paperback. Serif font was used. Because of my inexperience, I added an additional blank page before each chapter in *Victorian* because at the time I mistakenly thought it was typical to do for print. Indeed, the workshop participants did comment on the unusualness of the extra blank pages, but I did not comment on my mistake.

Although additional pieces from each project could have been incorporated into the physical book, such as the author's statement presented as a preface, I chose not to do this. My aim was to minimize the range of paratexts in the print book (note that the physical elements such as the book covers and the table of contents are paratexts) so that I could study whether participants felt that the books were lacking important information or lacking additional bookish elements (such as a preface).

The next step in creating physical copies of the generated novel was to select book covers and to print proofs. Lulu.com[50], a print-on-demand and self-publishing press and sales platform owned by Lulu Press Inc. was chosen because of its ease of use and popularity. This afforded me fast results and the opportunity to use a typical online publishing platform that is widely available to NaNoGenMo participants. The main takeaway from this publishing experience was the low quality of the binding. Proofs of 3 books with different paper and cover texture options were ordered: matte, glossy, white, and cream pages, plus paperback, hardcover, and a linen-wrap hardcover with dust jacket. Both the hardcover and linen-wrap covers were considerably

[48] https://www.latex-project.org/
[49] https://ctan.org/pkg/memoir?lang=en
[50] https://www.lulu.com/

more expensive than the paperback, but were nevertheless produced to a low standard with poorly finished spines and inside covers. I judged this to detract from the bookish aura that these more expensive covers were expected to have, so both generated novels for the workshop study were ultimately printed with matte paperback covers and 60# white paper. A 140 mm x 216 mm "digest" format was selected from the publishing platform's list of preset options.

As shown in *Appendix A , Table A3: Changes Made to Pieces*, *Victorian*'s GitHub repository does not contain a PDF file which can be used to interpret how the author would like the generated text to be presented; it only contains a plain text file where the GPT2-simple model's 99 blocks of output (each being about approximately 730-860 words in length) is printed, with each block separated by the characters "==========". I therefore interpreted these separations to delineate individual sections of text, which I chose to present as chapters in order to preserve the separate sections and to construct a stronger sense of bookishness. My assumption was that this would bring greater cohesion to the elements which make up the book object, increase the likelihood of participants' acceptance of the work as a physical codex form, and that it would afford easier navigation and reference to specific parts of the text during discussion inside and outside of the workshop. This latter point was confirmed during the workshop discussion.

The added chapters and table of contents are paratextual to the generated text, and they did eventually lead to the participants discussing about the potential puzzle-like aspects of the books. Along with the page numbers, these elements also aided participants in quickly communicating which passages they were referring to. Conversely, despite the ease of navigation that chapters might have afforded *Molly's Feed*, I chose to preserve the unbroken block of generated text not only because the PDF and HTML document presenting *Molly's Feed*'s generated text do not separate it, but because Bhatnagar's NaNoGenMo project is intended to be in dialogue with James Joyce's final part of *Ulysses* where the text is broken only by physical pages.

The cover design for both books is one of the Lulu platform's default templates. It features a white band over a muted green background (*Molly's Feed*), or a white band over a dark blue background (*Victorian*). Each white band only bears the book's title and creator, and the design is repeated on the back of the books without any text. Finally, each book's spine is white and bears each book's title. Just as with the interior layout, both books are designed to look identical in as many respects as possible, except for the cover color so that participants could easily distinguish and refer to them. I tightly controlled the similarity of their appearance because the books are first and foremost study items, and are therefore designed to have as many controlled

and comparable elements, or, variables, as possible in order to focus readers on the generated text itself as accessed through the physical codex form as paratext.

I chose not to include a publisher name on the books (this would have been a name referring to NaNoGenMo, or a fictional publisher) because this would again have introduced additional paratexts to the physical books and could have made understanding the impact of the physical codex form itself more difficult. The workshop participants did not comment or enquire about publisher information, presumably because it was understood that I had printed them myself. Interestingly, at least one of the participants assumed that I was the creator of the generated texts and that the authors' names printed on the books were fictional. When seen side-by-side, the similarity between the covers suggests that the books may belong to the same series within a publishing imprint. This was not explicitly said during the workshop, although participant P3 did comment on their appealing design:

> *"…it makes an attempt to appear in the traditional sense of what a literary, a piece of literature or fiction book would look like. But after having this conversation I can't help but think about P4's Ikea book example { P4: mmm}, and I feel like it just does a really good job of masking as a book {laughs} in my opinion. Although I would like to make the comment, I actually really liked the style of the books, like the way that these appear. Like these definitely would fit in my opinion {L: visually?} yea visually I actually really like them. When I opened the box and looked at them I thought 'ooo, quite like that, might just put that on my book shelf' {laughs}.*
>
> *P7: It's a definite penguin classics kind of style {group laughter}*
>
> *P3: Yea, that's the vibe I get! It's definitely wouldn't be lost you know in the classic fiction section in Waterstones {laughs}."* (p. 18).

After the workshop participant P6 clarified that they did not consider that the generated novels might belong to a series, although they pointed out that the context of the workshop study may have been why: "…I don't think I ever thought about whether they looked similar or not. But if I saw them in a bookstore, maybe - I guess seeing them in a research context impacts that quite a lot" (personal communication, 7 January 2023).

20 copies of each book were printed, with a per unit cost (excluding shipping) of £4.69 for *Molly's Feed* (181 pages), and £10.92 for *Victorian* (528 pages).

When performing the editor and publisher related tasks I prioritized building a sense of bookishness into and around the objects (hence why I chose to add a table of contents and

chapter titles) and tried to stay aligned with authorial intention in terms of presentation. The addition of a table of contents, chapter headings, and blank pages in *Victorian* are perhaps the most significant alterations that I have made to the project while performing the editor and publisher role to adapt the generated text for print.

## 5.3. Setup and running the workshop

In addition to the research project's research questions, I developed two study questions in order to design the workshop around and to help direct data analysis:

1. As a physical paratext, in what ways does the generated novel in the form of a printed book impact reading and interpretation?

2. Do formal reading group activities such as increased time spent reading, analyzing passages, and group discussion impact readers' perceptions of literary value?

The workshop was composed of 3 main tasks: a pre-workshop reading and annotation task that participants completed in their own time, a discussion task where participants responded to my and each other's questions and comments, and finally a creative task where participants were asked to rework one or more fresh (unannotated) copies of the books based on their impressions or the discussion session. This second half of the workshop was a creative reworking and crafting session where participants used the printed books as crafting materials to create their own piece in response to their impressions of it. This creating reworking phase will not be discussed further as it is no longer within the scope of this PhD project, but it may be used in a future publication.

### 5.3.1. Recruitment and reading instructions

Local participants were recruited for the study by word-of-mouth and email, and were contacted based on their similar or adjacent research, or based on their creative interests in relation to electronic literature or digital media more broadly. University persons, such as PhD students and lecturers, were contacted through their university email, and local creatives through the contact details listed on their professional websites or social media profiles. Once a participant had expressed interest in the workshop, informed consent and the delivery method or postal address for delivering the printed generated novels was collected from participants through an online survey using the same LimeSurvey platform as in the survey study. The workshop was run on a university campus that was easy to reach using public transport, and the workshop room was selected specifically with accessibility in mind. 7 of the 8 participants who agreed to take part in the study were able to participate on the day. Each received a £5

refreshment voucher to use during the workshop break, and approximately £70 as a token of thanks for their participation and to assist in covering public transport costs.

Once consent was given, I packaged both generated novels and an instruction letter in a small, plain cardboard box (seen in *Appendix E.1 Workshop instructions*) and arranged for the box to be delivered based on each participants' preferred delivery method. The instructions asked that:

> *"Before the workshop, you should read each book for at least an hour each. You are encouraged to make notes, comments, highlights, bookmarks, etc. directly onto the pages of the books to document your reading and impressions. The books are unusual, so it's perfectly fine if you find that reading in the conventional sense is challenging. Just do your best and see if you can a way of engaging that works for you"* (*Appendix E.1 Workshop instructions)*.

#### 5.3.1.1. The workshop participants

The workshop participants were highly educated with backgrounds in the creative and digital media industries, Fine Art, and technology. While I made every effort to recruit participants who were outside of my immediate social and professional circle, the 7 workshop participants were all persons I had met before, and each participant knew at least one other person at the workshop. While this initially worried me in terms of the ability to elicit a diverse range of opinions, it turned out to be an advantage because the discussion was lively, flowed quickly, was confidently opinionated with some sharply contrasting views, and overall appeared to be a very enjoyable discussion experience for all with frequent laughter. It certainly did not take on the formality of, for example, an academic conference workshop.

The workshop discussion has been transcribed and is available upon request. The page numbers citing each workshop excerpt refer to the transcript document. Participant names have been anonymized in this dissertation document. Identifying information has been redacted, but information pertaining to professional or academic expertise has been retained because of its relevance to interpreting and contextualizing participant responses. Each of the participants briefly introduced themselves at the start of workshop, and these introductions are reproduced in *Table: Participant pseudonyms and background*. My identity as the interviewer is not anonymized and can be identified by the initial L.

Two audio recorders were used to record the workshop and were placed on the workshop table where all participants could see the devices. The workshop began early because all participants were ready. After study consent information, safety and amenities information, and explaining the planned schedule and breaks, I introduced the task as a reading group discussion to discuss

what the participants thought of the two books they read. I invited any preliminary questions, and then began. The discussion ran for just over 1 hour and 20 minutes. The interview topic guide can be seen in *Appendix E.2 Workshop interview topic guide*. The workshop ran and concluded as planned. The annotated copies of the printed generated novels that each of the participants read were collected as research data, although these have not been analyzed because the workshop discussion itself was rich in detail and was able to yield enough data to answer the research questions.

Table 13: Participant pseudonyms and background

| **Pseudonym** | **Participant's Own Introduction** |
|---|---|
| P1 | P1: Hey my name is P1 I completed my PhD in *year redacted* and my research was around artist books, and trying to see the book as a social medium. Like a way of bringing people together through the making and sharing of the books. (p. 1) |
| P2 | P2: Lovely. I'm P2, I'm a lecturer in computer science. So actually, most of my research has nothing to do with art and I look at, I look at what people think about privacy and stuff like that and data. (p. 2) |
| P3 | P3: My name's P3, I am a PhD student, I'm based mostly in music, my research looks at how music streaming platforms are impacting our understanding and valuations of music. (p. 2) |
| P4 | P4: I'm P4, and I am doing a PhD in computer science as well, and looking at interactive narrative stuff. And I have also made a generated novel before, and so I have a personal interest in this, as well as previously I worked at a book shop, well ran a bookshop so, yea this is very much my jam. (p. 2) |
| P5 | P5: I'm P5, I'm in the final year of my PhD, hopefully {group chuckle}, and my background is computer science and AI and my PhD is looking at how AI can help interactive story write-, storytelling authors write stories. I'm pretty interested in all this as well {chuckle}. (p. 2) |
| P6 | P6: I'm P6, I just submitted my PhD, woo hoo { P2: chuckle} which is about the relationship between player and video game in narrative video games. (p. 2) |
| P7 | P7: I'm P7, I write interactive fiction, a bit like P4, I've also dabbled in sort of generated text. So one of the things I've done was ten million procedurally generated invocations to a fictional god. (p. 2) |

## 5.4. Deductive Qualitative Analysis method motivation and overview

This PhD project's research questions are focused on testing and refining an existing theory – Genette's paratextual conceptualization. It necessarily follows, then, that the workshop data is shaped by a deductive lens and that my analysis' entry point into the data is best initiated deductively. As a method which also uses coding and the development of themes, Deductive

Qualitative Analysis was therefore selected as the best method through which to analyze the workshop data.

Qualitative method developer Jane Gilgun[51] introduces Deductive Qualitative Analysis (DQA) as an update of and departure from analytic induction which emphasizes researcher reflexivity, and stems from the same school as Grounded Theory. It enables concept-guided descriptive research which is done with by developing a set of sensitizing concepts from the onset of the study which are balanced with a case analysis component where the researcher searches "...for data whose meanings might lead to modifications, refutations, and reformulations of concepts and hypotheses, both in the initial material and the material that researchers develop over the course of research..." (2019, pp. 2-3). Where the initial sensitizing concepts based on existing theory might help researchers to see what they otherwise may not have (p. 7), Gilgun stresses that DQA's requirement to carry out the refuting case analysis component ensures that researchers challenge their presuppositions and emerging findings (p. 6). This results in a tighter fit between the conceptual material which is developed by the research and the data that formed the foundation of the theorizing (p. 7). Indeed, the initial sensitizing constructs may be altered or removed as analysis develops if they ae found to no longer reflect the data. Relevant ones develop into themes.

More recently, Fife et al. (2023), and Fife and Gossner (2024), clarify Gilgun's case analysis by naming it 'negative case analysis'. The latter define it as negative cases in the data where participant experiences or instances sharply defy, counter, or mis-align with the guiding theory, its constructs, or predictions (p. 5). This is the term and definition that will be used in this dissertation.

Fife and Gossner (2024) is a method primer on DQA and I use it to structure the analysis and interpretation of the workshop data and findings. The researchers explain that DQA can be carried through a constructivist lens (which aligns with the epistemological orientation of this research project), and summarize their conceptualization of DQA as a direct way to "...operationalize theory, intentionally incorporating both deductive and inductive analysis, and emphasizing negative case analysis to prevent premature conclusions or confirmation bias" (p. 2). This is the conceptualization that I rely on because it not only describes my rather literal operationalizing of the generated novel projects' linked pages into individual paratextual pieces, but also my need to code and develop themes inductively in order to be able to test and challenge theoretical assumptions about how paratext functions. Similarly, as contrasting opinions were vocalized in discussions between participants during the workshop, the negative

[51] https://ssw.umn.edu/emeriti/jane-gilgun

case analysis component of DQA is the ideal methodological mechanism for homing in on important points in the data.

For my study I draw from Fife and Gossner (2024) and develop code into themes. However, they do not elaborate on the practical aspects of coding, and Fife et al. (2023) do not appear to define their units of analysis for coding. It appears that in the latter case transcripts have been coded for units of meaning rather than more prescriptive units such as sentences or a window of words. I do as Fife et al. and code the data according to units of meaning.

### 5.4.1. Deductive Qualitative Analysis steps

Fife and Gossner (2024) outline the DQA analysis steps (*Table 1: Processes and Outcomes of Deductive Qualitative Analysis*, p. 3). I use these steps to structure my DQA analysis within the larger workshop study and PhD project. The final 2 steps are detailed for the remainder of this chapter:

**Developing research question and selecting guiding theory**

This stage is reflected in *Chapter 2 Background and literature review and Chapter 3 Research Framework* , where the PhD project research questions are stated and discussed, as well as the unpacking of Genette's paratextual conceptualization and how subsequent scholars have critiqued and revamped this conceptualization. As explained previously, my research aims to test the impact of paratext on the reception of generated novels, and I expect to either challenge or refine the conceptualization in the context of this emerging form.

**Operationalizing theory**

Building from this project's research questions, the results and questions raised by the survey study, and the reading group workshop's semi-structured interview guide, initial and provisional sensitizing constructs were developed in order direct early analysis.

**Gathering purposive sample**

While DQA studies can assemble a sample of multiple cases, I only focus on the workshop data. I designed, organized, interviewed, recorded, transcribed and anonymized the workshop audio. I also created memos. The transcribed text is the data that I analyze here.

**Coding and analyzing data**

I incorporate immersion into this step to closely familiarize myself with the data by both listening to the audio and reading the transcript. Then, following Fife and Gossner, I alternate between an initial deductive approach to coding and inductive coding, including

negative case analysis, and inductive themes are developed (2024, p. 3). Coding iterations are conceptualized as 2 phases, although there isn't is a strict, prescriptive split between the two (p. 6):

**Early Analysis** – A first iteration of coding. This beings with deductive coding based on the initial sensitizing constructs and the paratextual conceptualization. Additional codes are allowed to develop inductively as evidence of new concepts beyond the sensitizing constructs are identified. Negative case analysis is also begun. I perform reflexive memoing at the end of this early analysis stage to better understand and link together the data, and I create a codebook of the sensitive constructs and themes developed at this stage.

**Middle Analysis** – A second iteration of coding. Theme development begins gradually as well as a continuation of negative case analysis. I create a codebook which shows the codes developed from this second iteration. During memoing I map the relationships between some developing themes which begins to construct an initial conceptual model. In this manner primary themes are identified.

**Theorizing**

I perform a third iteration through the data to finalize the codes, further develop primary themes, plus identify supporting and contradicting evidence for the initial conceptual model. The model is finalized to convey a working, theoretical refinement of how paratext may interplay with reading strategies in the interpretation and reception of printed generated novels. My conceptual model therefore is a product of an empirical examination and revision of "… existing theory such that it is more precise or accurate to the present sample, which may include altering or replacing components of the theory if there is evidence to do so (p. 8).

## 5.5. Method

After transcribing and anonymizing the workshop audio, I performed Deductive Qualitative Analysis (DQA) on the transcribed workshop data using the Nvivo[52] qualitative data analysis software tool to code and manage the data through each coding iteration.

### 5.5.1. Operationalizing theory

Supplementary to this PhD project's research questions, this DQA study's questions are:

[52] https://lumivero.com/products/nvivo/

1. As a physical paratext, in what ways does the generated novel in the form of a printed book impact reading and interpretation?

2. Do formal reading group activities such as increased time spent reading, analyzing passages, and group discussion impact readers' perceptions of literary value?

The semi-structured interview topic guide questions which were developed for the workshop formed the basis of the initial sensitizing constructs. These are *Author Related*, *Creativity*, *Generated Text Content or Qualities*, *GitHub Section*, *Literature and Literary*, *Physical Book*, and *Reading Experience*. Their descriptions are shown in *Table 14*: *Early Analysis and Codes*.

### 5.5.2. Early analysis

I begin my immersion in the data during transcribing and memoing. Because the transcript is of a group discission the text data has a relatively free structure. Therefore I made the decision to code data in units of analysis which encompasses enough information to convey a meaningful instance when the unit is read on its own. This means that all or part of a participant's turn in a discussion could be coded as a unit, or a unit could contain an exchange featuring more than one participant; capturing the evidence or idea expressed in one unit is prioritized over a unit of analysis with a prescribed length or structure. As the sole researcher on this PhD project I was the only coder and was not able to discuss my coding with a colleague working with the same data. To mitigate this, I use memoing to reflect on the analysis at each stage of the DQA study.

The sensitizing constructs were created as initial codes in Nvivo, and then coding began by alternating between deductive coding to the sensitizing constructs, and inductive coding. A simultaneous coding approach was used. Throughout this first iteration new codes were developed inductively when instances in the data did not clearly relate to the sensitizing constructs. For example, code *Ethics* was developed inductively because I noticed that participants made references and discussed what they judged to be offensive or dangerous content in the generated text or input data. This suggests that in these instances participants felt that it was a relevant point to share with the group and so I created the code to describe these ethical concerns. *Table 14: Early Analysis Codes* shows the sensitizing constructs and codes which have been created by the end of the early analysis iteration. As can be seen in the table, the sensitizing construct *GitHub Section* did not have anything coded to it because even though some references were made to information that is found on the GitHub pages, I interpreted the participants to be speak about it rather generally and not focusing their points on the specific GitHub pages and examples from them. For instance, participant P2 said that:

> *"I found the, the GitHub, for this and I had a little look at that and sort of the ideas behind it. And I hadn't read the book that it's based on, so I didn't really get that reference. But having read about it, and read the wikipedia page for that book, the concept then made a bit more sense { P1: chuckle}, I must say."* (p. 6)

Rather than explaining where on Bhatnagar's GitHub the "ideas" behind the project were, what they were, or commenting on them, the participant instead chose to also refer to Wikipedia and *Ulysses* when describing their reading path. This led to the inductive development of the code *Paratext* which is more inclusive and flexible in that it encompasses references to the project paratext or describing accessing it plus existing works which are referenced as an example or comparison to the generated novel projects. The sensitizing construct *GitHub Section* was therefore removed because it did not reflect the data well. This early analysis phase concluded with memoing where I considered and reflected on the research questions. The full list of codes which were developed inductively during this initial interation are *Code or Project Workings*, *Ethics, Fictional Stance, Framing,* and *Functional Uses*, *Paratext, Puzzle/game/play*.

### 5.5.3. Middle analysis

For the second coding interation, I draw on an interation approach that I used in the previous section during Qualitative Content Analysis - using a co-occurrence matrix to systematically immerse myself in and traverse the data through a different path based on overlapping codes rather than a linear path through the transcript. To do this, I generated a co-occurrence matrix in Nvivo and used it to systematically read groups of excerpts which had been simultaneously coded to two codes. I reasoned that if an excerpt had been coded more than once then it was possible that some additional complexity, such as an interplay between two or more concepts, may be expressed by it. This was an inductive approach as my focus was on developing codes that would capture new, perhaps more latent ideas. Because many of these codes express more specific points or more complex relationships than some of the initial codes do, gradual theme development took place. *Table 15: Middle Analysis Codes* is a codebook showing which codes were developed by exploring code overlaps. Some codes were also developed simply by me becoming more immersed in the data and were not related to initial code overlaps, and this is noted in the table as "Developed during analysis". I used my process of leveraging simultaneous coding to revisit data as a general guide rather than a rule. For example, although there is no overlapping coded text between *Creativity* and *Framing*, I nevertheless investigated a potential relationship between these codes based on my interpretation of the data where readers have said that they think creatively to frame the generated text for fun and for sense-making.

Reflecting on my approach, I note that it mirrors Bingham (2023) where a second cycle of coding "…can be applied to the data within the generated first cycle codes. In other words, in the second cycle, the researcher can further analyze the coded text of the first cycle, adding a second layer of coding to the initial first cycle codes." (p. 3). I did this systematically by drawing from a Content Analysis iteration step in order to engage with the data through a different path on the second iteration.

At the end of this second iteration phase I engaged in memoing by answering the research questions again, writing memos about specific excerpts from the transcript which I felt exemplified points made by participants, and furthering theme development by noting and sketching relationships between codes on paper. It is at this stage that the importance of some codes and their relationships were advanced as potential primary themes as I worked them into an initial conceptual model. For example, I connected *Framing* and *Reading and Interpretation/Sensemaking Strategy* with a readers' creativity relationship; I judge this to succinctly represent some of the primary experiences and opinions that many of the participants expressed and discussed.

#### 5.5.3.1. Theorizing

My conceptual model developed through 5 drafts and the Deductive Qualitative Analysis advanced into a theorization stage where I was able to answer the research questions and reached theoretical saturation in terms of developing the conceptual model and primary themes.

The analysis developed from the workshop data will first be discussed in relation to the PhD project's research questions and this study's questions by drawing on evidence from the workshop transcript. The findings developed from answering the research questions will then be presented in a further refined manner by reporting the primary themes and how they were developed from the initial sensitizing constructs and inductively developed codes. Next, the working conceptual model which developed from analyzing and developing the relationships between the codes, themes, and the workshop participants' reported reading strategies and discussion about the generated novels is presented in order to further refine the analysis. In each of the analysis discussions negative cases will be highlighted and discussed where applicable. Finally, these findings are compared with Genette's paratexual conceptualization in order to propose a theoretical refinement. This is a major novel contribution to research.

## 5.6. Answering the research questions

My general research question asks: in what ways does paratext influence potential readers' interpretation and reception of a generated novel? This is broken down into three specific PhD research questions:

1. Which paratextual elements have a central role in influencing the understanding and reception of a project?
2. Which reader skills and experience affect the influence of these paratextual elements?
3. Which technical and cultural characteristics of a generated novel affect the influence of these paratextual elements?

Supplementary to this PhD project's research questions, this Deductive Qualitative Analysis study's questions are:

1. As a physical paratext, in what ways does the generated novel in the form of a printed book impact reading and interpretation?
2. Do formal reading group activities such as increased time spent reading, analyzing passages, and group discussion impact readers' perceptions of literary value?

### 5.6.1. Answering research question 1

Which paratextual elements have a central role in influencing the understanding and reception of a project?

It appears that authorial intention, and latently the technical workings of the project, have a considerable role in impacting the understanding and reception of a project. The printed book form of the generated text was discussed, quoted, described in detail and compared frequently during the workshop. An idea for an additional paratextual piece that could be situated in the physical book was also discussed during the workshop – a critic's framing piece.

#### 5.6.1.1. Latent links to technical aspects

On a latent level participants spent a fair portion of the discussion talking about the generated text content or its qualities while making speculations and factually correct or incorrect, or unverified assumptions, about the technical aspects which led to the production of the qualities or textual artifacts. So participants seemed to be talking and reasoning about technical aspects which they arguably could have chosen to find and read about, but appear to have decided not to (although at least one participant did mention that they were not sure whether they were

"meant to Google it [the project] or not for the study" (p. 11). In this latent sense then, the content of the technical aspects of the project paratext seems to be referenced (but not actually read) by participants quite frequently, and arguably in more specificity in terms of actual examples and quotes than specific references to authorial intention. Indeed, during the workshop I was surprised that the participants did not seem more interested or curious about the code and other GitHub pages when I showed them towards the end of the discussion. Therefore, from my observer view point it appears that the participants found that pointing out and talking about the generated textual content or qualities through assumptions about the novels' technical aspects was conducive to discussing the project. Based on the transcript and the pace of conversation and laughter in the workshop audio, it also seems that the participants enjoyed engaging in this discussion, where perhaps the factual correctness of their technical-related explanations and potential for verifying them by reading the GitHub pages, or asking me, was perhaps either not important or not interesting. Rather, it could be that the *speculation* about the technical aspects and workings of the project was the interesting part. Thus, I see this as a latent connection to the technical pieces because the workshop participants are aware that the information exists and can be found, but instead they appear to be interpreting the generated text through their own or a pop-cultural idea of what the technical paratext is or indicates. It is possible that my analysis is overly sensitive to instances where participants make unverified assumptions or factually incorrect statements about the technical aspects of the project, because this seemed prevalent and intriguing to me when I was analyzing the survey review data.

Participant P6 explains that while they're not knowledgeable about the technical aspects of the input data and GPT2 model, they're also not interested in finding out. Rather, it seems that thinking about what these might be during reading was sufficient. Later in the discussion P6 says that they enjoyed reading *Victorian:*

> *P6: Yea I was, while I was reading it I kept thinking about what would the source material have been. Like not, not knowing or understanding anything about the actual, you know whatever, whatever the technical bit of it is that gets done with the source text. But I kept thinking, okay I wonder what, like where it's getting all this like, stuff from { P1: mmm}. Yea so especially with* Victorian. *Like with* Molly's Feed*, I, I was fairly certain I didn't really want to know {group chuckle} anyway because it was just going to be some random twitter things probably.*(p. 6)

Participant P4 speculates about *Molly's Feed* and constructs their sentences in a way that appears to personifying or give agency to an algorithm which makes a selection and display decision from multiple sources. Based on the code that Bhatnagar shows in their project, this isn't strictly how it functions, especially in terms of algorithmic decision making. But the point is

that P4 and the other participants seem satisfied with this reading of the generated text with a latent link to its assumed (not verified) technical workings:

> *P4: I think it's, because it's making a decision somehow to, of what, what to pick next to, to display, right? And, and they seem to go along in similar, so obviously it veers wildly off into all kinds of tangents but you can sense in some like, {inaudible. Passage?} pages that it, it repeats some of the same key- you know words pop up again and again. I might have some of them. Oh yea yea yea so like on, for instance on page 19 you can see that they used the word pretended. Or I pretended, like about 4 times in a row. And I think it was probably drawing, 5 times in fact, I think was probably drawing from different sources to do that { P1: mmm}. But, I mean, and so. Yea. So you can see if you look like the structure has some kind of coherence in it somewhere, yea.* (p. 10)

Similarly, participantP2made an assumption about a technical aspect of *Victorian* based on a textual artefact that interested them, but the interest did not seem to extend to actually verifying it:

> *P2: I wondered about that. They also, they all end mid-sentence { P1: mmm}, and I wondered if that was {P7: mmm}, I guess an artifact of how it was created like maybe they'd cut each block of generated text down to a particular length or something. { P4: yea, I think that's exactly it} So I was interested just whether that was a bug or a feature.* (p. 11)

It appears, then, that the potentially interesting aspect in reading or analyzing the generated text without an authorial intention framing or another creative framing is in making assumptions about how the technical process and technical aspects produced particular textual qualities and features. This speculative reading strategy, I propose, could be conceptualized as 'solving' textual curiosities by speculative explanation based on what the participant already knows or assumes about the project and technologies it uses, rather than looking to the GitHub pages to confirm their assumptions.

Of the participants who did search for a project's GitHub pages (such as participant R) they seemed to choose to not discuss specific parts or aspects of the GitHub pages which convey authorial intention to the same level of detail that generated text was discussed. On the one hand this makes sense because the GitHub pages were not present in front of the participants during the entire workshop, but it is nevertheless jarring that for a concept that was explained by participants to be important for interpreting and valuing the generated novel, that there wasn't more, detailed, and substantial discussion about specific authorial designs. For example, rather than explaining and linking the textual quality to speculation about technical aspects, it could have been linked more to authorial intention, or discussions about adaptations more generally. One might point out that this was actually done during the workshop where participants

discussed how they conceptualized *Molly's Feed* as a stream of social media posts, but I feel that this is a weaker, less clear link between the generated text and specifics about authorial intention when compared to the much explicit links participants made between qualities and artefacts in the generated text and technical aspects.

It seems that participants' flow of conversation (in part directed by my semi-structured interview) was focused more on their own interpretations and ideas rather than on an expansion of authorial, hypotextual, or intertextual topics and questions. I was personally surprised that questions or readings of the code and its workings were not discussed more or analyzed by the participants. Similarly, the projects' input could have been analyzed, but was not done nor asked to be done before or during the workshop even when the code was presented (albeit briefly).

##### 5.6.1.1.1. Expectations about code interpretation

My expectations that the code or input could have been read or analyzed and discussed as part of engaging with a generated novel is based on how I have come to understand the engagement approach from an Electronic Literature and Critical Code Studies approach, both of which are shaped through an academic lens. The latter is an interdisciplinary field closely linked to Electronic Literature which approaches computer code from the idea that it is culturally relevant and can be read and critiqued as a text, such as by applying a critical theory lens to interpreting it. For example, Electronic Literature scholar and Creative Code Studies pioneer Mark Marino demonstrates how a piece of computer malware code can be interpreted through a critical reading of heteronormativity (2009). And like a text, "...[the code's] meaning is determined not only by the programmer's intention or by the operations it triggers but also by how it is received and recirculated [by other audiences]" (Marino, 2020, p.4). They explain that even in the mainstream media, there's a growing sense of the significance of computer code (p. 3). But based on my data, I'm not able to show that the participants ascribe a high significance to nor specific interest in the generated novels' code.

Digital art and culture researcher Richard A Carter writes from an Electronic Literature and Software Studies vantage point and focuses on twitter bots in Carter (2020). The discussion throughout shows that code analysis and critical contextualization is an established part of academic reading. One of the goals in Carter (2020) is to demonstrate how "...generative digital writing [or, generative literature] can be assessed critically" to demonstrate that such an analysis can be a worthwhile pursuit that rewards "...extended academic criticism" (2019, p. 990). In what Carter presents as a typical example, they analyze Liam Cooke's *poem.exe* source code as well as samples of the bot's generated text as part of the approach to interpreting Cooke's creative work (pp. 991-992). Similar is seen in Montfort (2014). Digital media scholar Nick Montfort is well known in the Electronic Literature community as an author of generative

literature works[53], including poetry and generated novels. Their attention to a work's code is documented in Montfort (2014) (a technical report from the MIT Trope Tank, an academic lab[54]) where their 2013 NaNoGenMo novel *World Clock* is discussed. Presumably borrowing the term from Aarseth (1997), Montfort describes their works as "...not well-understood as traditional texts, and are better thought of as cybertexts. Even better, perhaps, they should be thought of as computer programs, programs with extremely specific materialities, programs that are absolutely inseparable...(in the case of World Clock) the "zoneinfo" time zone database." (2014, p. 5). Thus, Montfort's approach to presenting their work for an academic audience is to highlight the computational, code aspects. But again, this level of interest or engagement with the code cannot be shown in my data.

#### 5.6.1.2. Authorial intention and a proposed critic's framing

When the subject of authorial intention was present in the discussion, this was talked about positively and it seems that all participants valued authorial intention in principle. But authorial intention seemed to be talked about rather generally instead of making specific references to, or examples from, the creator's progress page on GitHub. Participant P2 describes how finding out about the author's intention improved their enjoyment, contributed to interest, and seems to imply that human intention is connected to creative value:

> *P2: I suppose I, I liked the Molly's Feed more after I understood in what sense it was created, and the logic behind it. So that whole process I guess felt kind of creative. And I guess the thing about Molly's Feed is that it has absolutely no coherence really, on a, on a global level. Because even once you know that it was based on a certain set of keywords that are supposed to reflect the final chapter of, of Ulysses or whatever, it, it doesn't, you can't really tell that from reading it, like, like we said you couldn't tell what keywords had been used. But, I dunno I guess I did enjoy it, after I knew in what sense it was creative and what like, the, the human who had been involved in it was trying to do { P1: mmm}. That felt nice to sort of see the result of their work from sort of, interest point of view.* (p. 14)

To contrast and nuance this, Participant P7 explained that they wanted to form their own interpretation of the text without being directed in a particular way. But even though they

---

[53] For example, generative poem *Taroko Gorge* (2009) and *Hard West Turn*, entry #119 in the 2017 NaNoGenMo challenge.
https://nickm.com/taroko_gorge/
https://github.com/NaNoGenMo/2017/issues/119

[54] https://tropetank.com/

avoided finding out contextualizing information about the project, this was expressed in a way that didn't seem to devalue the information as being less important than initially forming one's own interpretation. Earlier in the workshop, participant P5 seemed to feel that authorial intention was not an important part of their interpretation:

> **L: Is that that important for, to enjoy the reading if, if you did, that there is some intentionality here with how the themes or the style or whatever was structured? Or was it not?*
> *P5: I think for me it's not important at all. Especially with AI art the fun is more interpreting it than figuring out the intention. {inaudible} yea.* (p. 10)

Indeed, it seems that they did not begin the workshop thinking about authorial intention:

> **L: How did we conceive of that when we were reading? Did you feel like you needed an author for this? Did you have an author in mind? How did you sort of, conceptualize this, as you were reading or engaging with the book?*
>
> *P5: I mean I definitely sort of, like my thoughts sort of wandered to what was the model trained on. And like I was like, imagining it as an AI writing it, and I mean not necessarily thinking of it as an author, and like, I'm just kind of reading it like the text is there for me to interpret. But when, like when I see patterns like what P4 said I sort of think about was this trained on domain specific data and stuff like that.* (p. 5)

But P5 nuances their thoughts on authorial intention towards the end of the workshop discussion, and makes the point that generated novel authors would want to manage readers' interpretation paratextually:

> *P5: I mean it depends. Like with Molly's Feed, all of that, I think it's, I would like, like whenever I read generated text I feel like it would be nice to have the option to choose whether I want to know the intention or not. So it kind of makes sense to put it at the end { P1: mmm}. But I guess it depends on the kind of text like, for Molly's Feed I don't know if I could have read it in an interesting way without knowing the intention. So I feel like as an author, if I was writing it I would have wanted to set that expectation first, for the readers.*
>
> **L: Okay. So it's the difference between I guess what the author would, what the you think the author would want and what you would prefer.*
>
> *P5: Yea yea yea, exactly.* (p. 26)

So although the creator's progress page (Piece D), as an operationalized paratextual unit and distinct and individual GitHub page, was not specifically referred to in the workshop discussion

(instead “the GitHub” was referenced), the concept of authorial intention seemed to be talked about frequently and positively in terms of interpreting the generated text and valuing the project. Later in the workshop discussion, conversation flow and my questions moved in the direction of what sort of additional paratextual pieces in the physical book could benefit the reading experience. Some participants talked about a framing piece to aid in interpreting the generated text, and this began as information about the author’s intention, although it wasn’t necessarily implied that this should be the Creator’s progress page (Piece D):

> **L: Do you think that the, well I guess my next question kinda is, to lead onto that, do you think that something could be added or taken away to the physical book based on what you’ve seen here {indicated screen} that would improve your reception of it perhaps, or your interpretation?*
>
> *P6: I think it would be nice if, if you know, we’re just talking about this how, this how it’s printed as a physical book to have that kind of information at the end, maybe.*
>
> **L: At the end. Which information is that?*
>
> *P6: Just kind of like I guess the author’s intention like the, the bit where they are describing what they did and why. Yea, I mean not necessarily like the whole entirety of the code, but just kind of like the auth-, I guess the author’s statement would be nice to have at the end. But not necessarily at the beginning because then it would possibly sort of distract from, from what was going on {L: okay}.*
>
> *P5: I mean I think for Molly, Molly’s Feed would have benefitted from that intention being written down somewhere. But for Victorian I just kind of like that it’s, I mean, I don’t know if any of the information here would have made me feel different about it, but if they’d come up with some other kind of narrative framing for it, explaining why it’s jumbled text, that might have made it interesting { P1: mmm}.* (pp. 25-26)

While participant P5 agrees with P6 about the benefit of having the author’s intention available, they suggest that an alternative narrative framing (not necessarily the author’s) would improve interest.

Drawing on interpreting Contemporary Art as examples, participants P1 and P2 mention interest and explain the benefits of explanatory information in the context of galleries and museums.P1seems to make a connection between interacting with a conceptual piece and having help interpreting it, but that this is dependent on the individual work. They also point out how *Victorian*’s more book-like structure appears to have aided in interaction or giving

interaction cues to the reader (note that this is the generated novel where I added numbered chapters to the physical book):

> *P1: I feel like if you're interested in the process of how the novel was created then sometimes like the ideas behind it, it can add to it. Like when you go to a museum and you're faced with a real conceptual piece and you have no idea where to even, you know, kind of interact with that. {inaudible} that little blurb like next to it can really help you know { P5: yea} and anchor where, you know, how to sort of understand that piece {P7: mmhm} and create your interpretations around it. But then there are other works where you just want to go in and not be distracted by any other information you know { P5: mmm}, so, yea I think it depends. I think with Molly's Feed that is was really, I Googled too you see, so then, and then I kinda like what it was about, and, and that actually helped me sort of frame { P2: mmm} like what, what I was sort of confronted with. With this one too {indicates Victorian} I had a quick look but yea, but this one I just sort of went into it. Maybe because it was sort of more like a book, you know, it was broken up into chapters, I sort of knew how to behave with it maybe { P5: mmhm} you know? I mean I don't know. So yea, I think it depends you know whether these like these extra information is helpful or not, or like whether it's necessary or not.* (p. 26)

Participant P2 conceptualizes this additional information not as an interpretation, but as a narrative. Here as well they emphasize that this can provide or improve interest although this depends on the individual piece, as well as the method used to create it:

> *P2: It's nice to have that, it's not interpretation is it, but that, narrative about the artifact somewhere. Because if you go to an art gallery sometimes you do just want to look at a picture because it's a nice picture. But other times, actually it's interesting to hear about why it was painted, or how. Particularly if it's like a novel method for making art, like it's made of elephant poo or something like that, it's interesting right? I have a, I have a {laughs} I have a particular, though I didn't pick that method specifically because of these books {group laughter} {L: are you sure?}. I have in mind a particular piece of art I once saw in a Southampton art gallery, which was made of elephant poo. Like, like that's interesting. And actually I think so much of the cultural value {chuckle} in these {indicates books} is, is almost as commentary on these methods that are sort of emerging. And the relationship they DO have to what, you might actually call literature. But that was interesting and like I say when I did Google Molly's Feed and I understood the reasoning behind it, I didn't hate it as much anymore {group laughter}. So {chuckle} yea.* (p. 27)

I asked P2 where this additional information would ideally be located, because I wanted to better understand whether the experience of searching for and discovering the project paratext

online added to R's interest, instead of only reading offline with the physical book. The online vs. physical book wasn't explicitly answered, but the essential response is 'it depends'; P2 answers by making reference to existing canonical works of literature and the different means of publishing them in order to present additional explanatory information:

> *P2: It was nice to be able to go and find it online I guess. But then, I suppose it does depend why you're reading it doesn't it? So I'm, I'm conscious that you can just buy a Shakespeare play as a play, or you can get the copies they have in school which are the Shakespeare play with some notes and explanation about what on earth the man was on about {L: mmhm}. And maybe some history about like why it was written in a certain way for example {L: okay}. But then with, with Charles Dickens actually it's helpful to know that it was written to be serialized and that's why it has the structure it has for example {group mmm}. So maybe it's just a case of working out the right balance of like what's useful to know at the start, versus what isn't. And, and maybe that's a publisher's sort of curation role. Like a, someone who knows more about it, {to L} maybe there's a job here for you {chuckle}, someone who knows more about it who can help guide people into interpreting it.*
>
> *P5: Yea, I think that's true. Like the more abstract something is the more it would benefit from some ideas on how a reader might go about interpreting it. Like, not necessarily that this is what it's meant to be, but, kind of like this is ways which you could interpret it { P2: mmm}.* (p. 27)

Here, neither P2 nor P5 are talking about the author's intention; they seem to be emphasizing an editor or critic's paratext. The participants don't appear to be talking about the same thing, though. Where P2 speaks about explanatory information such as historical context and a pieces' structure, P5 appears to be talking about an interpretive framing which is consistent with their creative narrative framing reading and interpretation strategy as described earlier throughout the workshop. But I understand both participants to be describing paratextual pieces which require their writer to be very knowledgeable about all aspects of a generated novel project, and in the case of AH's idea, require the ability to creativity frame the generated text.

Other participants agreed but stressed that this additional paratextual piece should be located at the back of the physical book. Participant P3 explains how this placement would be able to support the different reading and interpretation approaches which the participants engaged in:

> *P3: I was just thinking, having a small piece of explanation at the back of a book, would probably, in my opinion, would be a beneficial thing. Because then if I want to check I can,*

*but it doesn't mean that everybody has to {group mmm}. It's very similar to for example you went and Googled it, whereas P7 made a point of not Googling it. So there's option. It just, all I think it does is just gives you a little more freedom and if you want to know more, or if you want to experience it completely authentically without, I guess, muddying what your impressions are. Because for example Molly's Feed you went into like a whole different zone. Like you {indicates AH}, you made up a way to sort of make it make sense for you and make it more fun for you { P5: mmm}, whereas maybe I guess someone else would go 'nope, no idea' and maybe would view that more as a wall, maybe don't have those skills to sort of make it a more fun exercise, and would benefit from say, having an explanation at the back of the text or something like that.* (p. 27)

In sum, the participants expressed the value of knowing authorial intention and explanatory information, as well as a creative framing to the generated text which is different from the former two. They emphasized that the necessity of the paratextual information very much depends on the generated novel, which is the same case as other creative works and forms. Therefore I conclude that while the GitHub pages' Creator's progress page (piece D) and the Media page (piece F) might be the most similar paratextual pieces that this PhD project has worked with thus far, the discussion presented here suggests, rather, that a contextualizing and framing piece written by a knowledgeable critic/editor could be the paratextual element that could have the greatest impact on a new readers' understanding and positive reception of a generated novel as a physical book. In terms of which piece has the most central role, this is quite clearly the generated text itself as it appears to be the central site of initial interpretation and discussion from which other pieces are referenced and linked.

### 5.6.2. Answering research question 2

Which reader skills and experience affect the influence of these paratextual elements?

This research question was developed to find out whether the value or importance of specific paratextual pieces are impacted by specialized skills, for example, if understanding the computer results in the reader valuing the code more. But neither the previous survey study nor this workshop study was able to evidence that. Yet, the workshop discussion and analysis indicate that there is a more abstract ability that seems to enable new readers to enjoy the reading – their own creativity.

Based on the reading strategies described in the workshop discussion, the ability to interpret and frame the generated text creatively or narratively appears to be a participant skill that enables a better reception of the generated text. Participants' accounts of rereading the generated text, marking it, and thinking of a narrative vignette to explain why the generated text

content and qualities have the form they do and why the reader is interacting with it requires them to exercise their own creativity, playfulness, and the ability and willingness to engage ergodically with the generated novel. Participants who did this also described the activity as enjoyable and fun.

By ergodically, I mean that the reader is making a non-trivial effort to traverse and interpret the generated text. This concept was coined by Game Studies and Electronic Literature scholar Espen Aarseth, and is traditionally used to describe the process of reading interactive fiction, or, cybertexts[55] where "...nontrivial effort is required to allow the reader to traverse the text" (1997, p. 1). Valdimir Nabokov's *Pale Fire* (1962) is an common example used in literature classrooms because of its extensive and unreliable-narrator-plagued hypertextual apparatus presented in the form of footnotes. Regardless of electronic or physical media, ergodicity captures a degree of interactivity. Returning to Aarseth, they break down their detailed distinctions and potential relationships between ergodicity, the cybertext, the hypertext, and the ordinary text in a model: *User Functions and Their Relation to Other Concepts* (1997, p. 64). Engaging further with Aarseth through the lens of generated novel reading falls outside the scope of this current research project, but I mark it for future research.

Arguably, interactive and ergodic engagement strategies are described by the workshop participants where a second attempt at reading the generated text is described with the reader working to mould an interpretive narrative vignette or fictional scenario to the generated text as they parse it. I note that participants also describe the challenging aspect of parsing the text. Additionally, several participants reported using a writing implement to actively mark textual content, artifacts, or patterns as they parsed the text. While the workshop's reading instruction did ask participants to annotate the generated text, it did not specify that it needed to be done in such an interactive, responsive, and playful manner:

> *P5: I was gonna say, the lack of punctuation in Molly's Feed made it seem more like stream of thought, and because it's sort of like, abruptly ended and became something else. So I mean, I didn't really mind the lack of punctuation, but it was just hard to read it like it was. So I think after reading a few pages I was trying to find an interesting way to read it like, frame it narratively in my head like, pretend I'm a psychiatrist parsing {chuckle} {group laughter} somebody's rambles. And like I thought of, marked stuff based on different ways to read it. So like I marked stuff that felt like, if I'm a psychiatrist this would*

[55] Apparently frequently erroneously, according to the detailed discussion and disputes in Noah Wardrip-Fruin's personal blog post *Clarifying Ergodic and Cybertext* (2005). Here, several respected Electronic Literature scholars join the discussion, including Aarseth. https://grandtextauto.soe.ucsc.edu/2005/08/12/clarifying-ergodic-and-cybertext/

*be interesting to me, because it seems kinda messed up {group laughter}. There's also a lot of mention of spying and … Russians…or like terrorist activities, so I was kind of, pretending I was sort of, parsing surveillance data, and I marked stuff like that, so that was fun {chuckle}.*

*P1: I probably read it quite similar to you. The first time I thought, I've got to read it and like make proper notes, you know like proper. But then I kind of thought well I can't really there's so much of this sort of you know rambling going on. So I sort of split it into first read which was kind of trying to read properly, and the second read was just highlight all the capital letters you know, in yellow. And then the third read was trying to like… I found like spreads that didn't have any markings on it, you know like the page spreads, then I just treated it almost like, empty space and made lots of drawings, like this {shows drawings}{group laughter}.*

**L: What lead you to highlight the, the capital letters?*

*P1: Yea because it was already like, popping out you know? Like visually the capital letters already popping out, and I actually quite liked it. It was like, it was suddenly like 'ha ha ha ha', and like or… And you could kind of imagine it, like visually imagine it in your head, like WHAT? You know and then suddenly you sort of you know like read it like WHAT? So amongst all this sort of like rambling of this stream of consciousness there were these sort of things that kind of anchored or like kind of caught your attention. So I thought well I'm gonna then like highlight it it's gonna, it's already capturing my attention I'll make it even more eye-catching by, you know, highlight it in yellow, yea. So I was really playing with it quite visually really I think. After the first read when I tried to it properly, and then, it sort of failed miserably {chuckle}.*

*P5: Yea I think I did the same thing, like, in the first…reading I was just trying to figure out how should I read this { P1: mmm} and I sort of like read the same pages again and again with like different way of reading it { P1: yea yea yea yea}, it was kind of fun.*

**L: Did anyone else have any other ways of reading that they kind of…*

*P4: Yea so with, with Molly's Feed I took it as a kind of, almost like head swapping between people's kind of circling, like neurotic thoughts. Like as if I was jumping between different minds. I guess like in a, in a feed, like just like people putting their unfiltered thoughts out. But at first, because it wasn't coherent to read it as the thoughts of one person I don't think, so, but it kind of makes it, if you, it's kind of like jumping into lots of different heads.*

> *And sometimes it feels like it's circling back to some of the same characters in a way {P7: mmhm}. Because of some of the same obsessions come back again and again*. (pp. 4-5)

The workshop participants themselves seemed to understand creative framing as an ability. For example, P3 appears to refer to AH's own creative framing of *Molly's Feed* as a skill:

> *… Like you {indicates AH}, you made up a way to sort of make it make sense for you and make it more fun for you { P5: mmm}, whereas maybe I guess someone else would go 'nope, no idea' and maybe would view that more as a wall, maybe don't have those skills to sort of make it a more fun exercise…* (p. 27)

To connect this with the suggested critic's piece, I note that the proposed placement of a critic's framing piece at the back of the physical book is interesting because another genre which is also published in book form that and requires a reader to work hard to creatively interpret and 'solve' a work around a narrative framing is the puzzle book. That genre has the solutions to puzzles located at the back of the book so that readers do not accidently see the answers before they have had a chance to engage with the puzzle or challenge. Sometimes this solution information is located elsewhere, off-site from the physical book. This has a curious parallel not only with the ergodicity described above, but also with participant D's puzzle solving approach to engaging with the generated novel and the conscious avoidance of the project paratext. D's reading strategy also included marking text patterns on the pages.

A relationship between human agency, sense-making and a significant reading experience was made by participant P4:

> *P4: Well, you know actually it did, it did make me think of some, some kinds of like experimental poetry can be very difficult to parse, and you might approach it in a very similar way of like sense making, trying to draw together things. And I guess the only real difference is that you would think that with the poem, some, there's been some kind of human agent trying to put something together { P1: mmm} such that it will afford some kind of significant experience* (p. 19)

Where the value of authorial intention seemed at least in part to be based in the human involvement, in terms of the ability to engage ergodically and to develop one's own creative framing of the generated text, I propose that it was the reader themselves who was considered to be the creative human element, and perhaps this is where at least part of a significant reading experience may lie.

### 5.6.3. Answering research question 3

Which technical and cultural characteristics of a generated novel affect the influence of these paratextual elements?

Based on the workshop discussion data, the answer to the research question is tied up with broader cultural expectations about authorship and novels and books. It appears that these expectations might affect the ability of printed generated novels to be accepted as having literary value. It seems that the generated nature of the projects, despite different levels of human authorial involvement and intention, and design, violates the ability for the projects to be able to be considered along literature or literary lines. Participant P1 appears to point out and reflects on the expectation of authorship when asked whether the generated novels could be considered literature or literary. They also link creativity into these expectations:

> *P1: Yea well I. it's a really good question and I'm just wondering like, you know, if I didn't know that a computer was involved in creating these novels, whether or like I would have walked away thinking 'that's pretty creative you know, what these people have thought of to do'. I mean it is creative but, whether, I mean it's, it's interesting but, exactly it's interesting and it might be like inventive but like whether it's creative... again, you know, I sort of like, wonder. Whether it's literature or like literary...yea I don't know* (p. 21)

I consider the idea of literary value to ultimately have not been accepted by the participants. But nevertheless they do ascribe the generated novels a general cultural value. This value seems to be anchored in their ability to serve as a type of conceptual commentary piece[56]. Participant P5 described and quoted *Victorian*, and perhaps the physical paratext of the book form itself, as a piece which mocks literature and the language found in literary books. Thus P5 perhaps considered the generated novel's status as a commentary piece to be expressed as work of parody:

> *P5: I mean I felt like Victorian was mocking literature than masking it or itself being literature. Like I said, like parts of it was, it felt funny to make {inaudible} the fact that it was mocking the type of sentences that used to be there in like old English books or something. I don't know where I see it but it's like familiar. Like long sentences with like, a*

[56] Having critically studied both Art and Literature, my personal view is that an ability for a work to serve as though-provoking commentary is certainly an attribute of a work of literature and art (be they good or bad examples). Therefore the generated novel's unacceptability under these banners is curious.

*lot of clarification in it and being like 'or this, or that' kind of thing. So it felt like mocking literature.* (p. 19)

Ultimately, in response to the literary question P5 felt that the concept of literature was not well defined, and they instead stated the that generated novels were distinctly 'AI Literature'. Participant P6 also agreed with this term although specified that for them, AI Literature could be a type of literature although they were not sure.

To refer again to participant R's point, the development of their idea about the generated novels' valuation seems to have a shifting opinion about creativity and human involvement or authorial intention, which they initially highlight:

*P2: I suppose I, I liked the Molly's Feed more after I understood in what sense it was created, and the logic behind it. So that whole process I guess felt kind of creative. ...I guess I did enjoy it, after I knew in what sense it was creative and what like, the, the human who had been involved in it was trying to do { P1: mmm}. That felt nice to sort of see the result of their work from sort of, interest point of view.* (p. 14)

And yet, towards the end of the discussion P2 appears to have disregarded the human involvement that they themselves had pointed out:

*Like, these {indicates the books} feel like they have to be understood in context or in comparison to actual literature { P4: yea}{ P5: yea} so in that sense they're sort of orthogonal to it and are of literature but in the same way they're missing something that seems to be quite fundamental to literature, which to me I guess was the idea that you'd read it and you'd learn something about someone else's idea. Or it would be like human communication, which these aren't really.* (p. 20)

Nevertheless, P2 stresses a link between the generated novels and literature, where "*...I think so much of the cultural value {chuckle} in these {indicates books} is, is almost as commentary on these methods that are sort of emerging. And the relationship they DO have to what, you might actually call literature"* (p. 27).

Although it was not tested directly, it seems as if the amount of information about human intention that is available about a generated novel might have little impact on its acceptability as a piece of literature. Perhaps this is dependent on cultural factors, such as expectations about media and their production, which are resistant to paratextual influence. This begs the question of whether inviting a generated novel author to give a talk or public interview about their creative processes and intentions would be able to impact readers' literary and creative valuations, and through the general trend across the studies' data is that this might be very

challenging to impact in general. Speaking reflexively through an anecdote to explain how I partially arrived at this comment, some time after the workshop I gave a guest speaker talk at a for a writer's society event titled *What Has AI Ever Done For Us? A Dip Into Generative Literature* (2023). After presenting at length and detailing the authorial intention, choices, and distinct textual styles of 3-4 exemplary NaNoGenMo generated novels, a colleague and I noted that the audience questions nevertheless seemed to carry a rejection of generated novels as intentional, authored, creative works where the creative and intentional human aspects of the projects were overlooked or not considered – not 'taken onboard'.

#### 5.6.3.1. Folk theories of AI

Interestingly, as described previously the workshop discussion involved many instances of interpreting and commenting on the generated text qualities with links to unverified or speculated workings about the projects' technical aspects. Individual speculations perhaps fall on a spectrum of factual and popular cultural beliefs about how a piece or technology might work (several of the survey's written reviews especially appear to align more with popular cultural beliefs about AI technology rather than factual ones). Thus I propose that to varying extents, folk theories of AI are used by some participants to explain aspects of the generated novel or to comment on it. This is a useful concept because it offers a critical lens for understanding assumptions and beliefs which can shape attitudes towards generative literature, and technologies and policy more generally. My idea to use the term 'folk theories of AI' is informed by the concepts of folk theories of algorithms and folk psychology. While it's possible that other researchers might also be using the term 'folk theories of AI', it is not currently a widely used term and I therefore define it for my own use here.

Folk theories of algorithms is a broader term and encompasses data and algorithmic tools and processes more generally. Ytre-Arne and Moe (2021) is a qualitative study demonstrating this, and uses folk theories of algorithms to understand and critique how datafication is experienced by people through everyday media use. The researchers draw from Human-Computer Interaction (HCI) which also studies technology use and user responses, but where HCI's goal is to improve software design, Ytre-Arne and Moe specifically take a critical stance for analyzing and critiquing technology and user agency (p. 808). Therefore the concept of "…'folk theories' is an approach to analyze the understandings that people draw on in everyday life…[it] centers on revealing the conceptions people hold of how the media works – that is, their theories." (p. 810). And crucially, as I also argue for folk theories of AI, to analyze these conceptions interrogatively and understand them as sociotechnical phenomena. While several of the processes covered in Ytre-Arne and Moe (2021) could be accurately and factually described as using AI technology (such as reccomendation systems), I use the term folk theories of AI to refer to cases where lay-

persons specifically make references and explaintions based on popular beliefs, imaginations, and narratives about AI. I have noticed that this has become increasingly prevalent in popular discourse since the AI 'boom' trend after OpenAI's release of the popular AI chat product ChatGPT for consumers (introduced in section *1.5 Research project Scope*, and discussed further in section *7.3 Limitations and future work*).

Folk theories of AI is concept that Colombatto and Fleming (2024) also arguably work with, but they instead use the term folk pyschology to measure how folk intuitions about supposed AI consciousness differs from expert knowledge. The researchers designed a study where particpants were asked to read a text about ChatGPT, and then read text about consciousness and experiencing before answering survey questions. They explain that their core motivator for understanding folk psychological attributions of consciousness (regardless of whether AI actually is conscious or not) is because public perceptions of AI can shape moral stances, and can affect the extent to which future research can impact the public (Colombatto and Fleming, 2024, p. 1). I understand their research to be focused on folk theories of AI because they're working to understand non-expert's popular beliefs and narratives about AI, and the impact that this might have. Therefore I argue that it's useful to interpret some qualitative results as folk theories of AI because this gives a recognizable form to insights about popular and inaccurate perceptions of technologies and its implications.

### 5.6.4. Answering study question 1

As a physical paratext, in what ways does the generated novel in the form of a printed book impact reading and interpretation?

This answer is linked to the second research question, where it's the printed book form which seems to better catalyze an ergodic approach to reading, and through that, better enable readers' ability to creatively frame and interpret the generated text. Based my observations and confirmation from the workshop participants, the physical paratext enables a better traversal and navigation of the generated text than a digital PDF file does (digital PDFs are a typical NaNoGenMo entry format). Here, 'better' means an ability or willingness to engage with the generated text for a longer amount of time and being able to traverse and find interesting textual qualities, artifacts, or passages faster and more frequently. For example, when asked about whether they felt that the physical book had an impact on their reading experience, participants talked about longer reading, moving through the book in a non-linear manner, and how the physicality lends a more enduring or final quality to the project. This suggests that the digital version has a more ephemeral aura to it:

*P5: I mean I think I'm definitely more likely to read for longer if it's a physical book and not on screen.*

*P4: Yea I agree.*

*P7: You kinda got the assurance that it's not going to be updated at some point { P5: chuckle} like there's going to be a new version generated and then 'oh you know we'll…' like this is now the current GitHub version or whatever with completely different text.*

*P1: So feels sort of more concrete, or like a finished product in that sense.*

*P7: yea.*

**L: An archival copy.*

*P6: It's also quite nice to just be able to leaf through it, especially, especially when, when there is no coherence really and you're like well I'm trying to figure out a way to read this so I'll just flip ahead to like something that, it's nice so it kind of is, it's a lot easier to just like flip through them and stop when it, when it, you know when something jumps out at you. Whereas if it was just a PDF, yea { P1: mmm} that would be not, yea not really working that well I think.* (p. 16)

Although not everyone agreed – participant P2 said that they missed the affordances of digital media (p. 16).

However, the printed book form was not able to impact perceptions of readability; participants P3, P4, and P1 felt that something was promised through the physical form, but was then violated. For P1, this violation of a conventional reading approach seems to preclude a link to literature or (presumably literary) books:

*I mean obviously you guys a really interesting point about it. Because it's a physical book you sort of have this expectation that it should sort of behave like a book. And like when it kind of doesn't, and like you find it really difficult to read then you feel a little bit betrayed, you know. And you think 'well what am I supposed to do with this', and then I think that's why I kind of came up with my own rules and like you know, how to read because it was, it wasn't readable in the conventional sense. So yea, so and to that question of whether it's literary or not, it's like I don't know. Mmm {*L: okay}. Yea I don't know actually. I think I, I would put it in, because like I've got a collection of artist's books you see, and a lot of the artists books are also like not readable in a conventional way. So I think it would go under that, you know. And it wouldn't necessarily sit next to my other collection of books on the bookshelf.* (p. 21)

N's thoughts link to the broader cultural expectations about novels and books as discussed in response to the third research question.

### 5.6.5. Answering study question 2

Do formal reading group activities such as increased time spent reading, analyzing passages, and group discussion impact readers' perceptions of literary value?

This question sought to capture whether treating the generated novels as more conventional reading group works and book objects would enable them to be perceived as or accepted as literature or literary in some aspects. But no, there isn't strong evidence in the workshop study that this is the case. As discussed in the answer to the third research question, it seems that cultural expectations around authorship and books may affect the ability of the generated novel books to be accepted as having literary value. As already seen in previous workshop discussion excerpts participants were either not sure or unwilling to think of the generated novels as having literary value specifically, although cultural value and value as 'AI Literature' was expressed. An important and distinct exception to this, however, is participant S's impression and defence of *Victorian*. Where N's challenged reading perhaps precluded linking generated novels to literary value, P6 appears to link their pleasant reading experience to literariness:

> *P6: So I kind of feel like, well to me Victorian, at least it feels, it feels very literary. Just because for me, clearly not for you, it was a very pleasant read {group laughter}. You know just a pleasant, like the language was just really nice to read. I could like engage, you know I could { P5: yea}, like there were like a lot of like really nice kind of like poetic descriptions in there. So just the style was like, like I'm not, I'm not thinking that it's literature, but it is literary. So Victorian at least. But then maybe like P5 said, maybe that is just AI literature. So maybe it is literature just like a different kind of literature, I guess.* (p. 20)

S at first describes the generated novel as not literature, but they appear to confidently accept it as literary. But, as mentioned before, they ultimately refer the generated novel as AI Literature and perhaps suggest that it could be a type of literature - but this statement seems less sure. Participant P7, a creative fiction writer, pointed out that they would not be interested in reading a literary novel anyway:

> *P7: I kind of feel th-, it's odd because when you say like 'oh you know it's not what someone would think of as a literary novel', I kind of think, well, I wouldn't read a literary novel for fun so in that respect {group laughter} there's certainly more similar than a lot of, you know…*
> (p. 19)

## 5.7. Sensitizing constructs, codes, and primary themes

The sensitizing constructs which remained relevant to the workshop discussion analysis are *Author Related, Creativity, Generated Text Content or Qualities, Literature and Literary, Physical Book,* and *Reading Experience*. These are documented in *Table 14: Early Analysis and Codes*. In the same table, the codes which were developed inductively during the early analysis phase *are Code or Project Workings, Ethics, Fictional Stance, Framing, and Functional Uses, Paratext,* and *Puzzle/game/play*. The table's first column shows the code name, the second column contains the inclusion criteria for the coded excerpts, and the last column shows an example excerpt from the workshop transcript.

Some sensitizing constructs were not developed into the conceptual model, but they were nonetheless useful in structuring the data and answer the research questions. For example, *Literature and Literary*. Similarly, the inductively developed codes which are not in the model are nevertheless useful in answering the research questions, such as *What is a Generated Novel* which contains coded excerpts of instances where the generated novels don't conform to expectations about books and literature, as discussed in section which answers the third PhD project research question.

Table 14: Early analysis codes

| Initial Code | Definition | Example |
|---|---|---|
| Author related | An author or creator is explicitly referred to or implied – human or AI. Authorship. Whether intentionality is important or not. | P7: I, I also Googled the, because I, I saw Janelle Shane and though, hmm that sounds like someone's come up with a vaguely Jane Austen sounding name, and no I, turned out they're both real people {group laughter}. |
| Code or Project Workings | Explicit or implicit references or assumptions about how the code or other technical aspects of the project function. Explaining generated text qualities as artifacts of the generation process. | P5: I mean I definitely sort of, like my thoughts sort of wandered to what was the model trained on. |
| Creativity | Explicit or implicit references to human or machine creativity, co-creativity, the participants' own creativity | P2: Yea, yea. Likewise with Victorian I think thinking about it as maybe a way to get some inspiration. There were some nice sentences there, and I almost wish that you could've said 'yea I kind of like this sentence, maybe there's some others on a similar theme. But that might have been a nice way to, to like use it as a way to prompt your own creativity. |
| Ethics | Explicit or implicit references or questions about offensive or dangerous content in the generated text or input data. | Chapter 3 starts off a bit misogynistic. |
| Fictional Stance | Participant reports taking a fictional stance to the generated text | …You can treat it as if it were fiction, and then you might get more enjoyment from that. Or you know, you could treat it as if it was real. … |

| | | |
|---|---|---|
| Framing | Participants report, or I interpret their reading approach, as using narrative framing techniques. | P2: I don't know if it would, it wouldn't be sentence structure because clearly it's got in its little AI head a few sentence structures that it really loves {laughs} {A: chuckle}. So yea maybe, maybe concepts or, or keywords. |
| Functional Uses | Suggested or implied functional uses of generated novels. | But it's, yea, it's, it's an odd combination of vocabulary and style there, and it's the kind of thing that, you know, if you could find, like, there's got to be a character who would say that. I don't know that Molly is that character, but, it's, yea. I could be a starting point for something else. |
| Generated Text Content or Qualities | General discussion or specific references to the content of the generated text, or its textual qualities and style. | P5: To me it came across as a style. Like specifically AI generated style {inaudible} |
| GitHub Section | References or discussion about specific paratextual pieces on GitHub | No text was coded for this theme; it did not end up being useful to the analysis in terms of sensitizing constructs. Project Paratext proved to be a more useful code.. |
| Literature and Literary | Litera*, poetry, literary movements, styles, or works | P7: I kind of feel th-, it's odd because when you say like 'oh you know it's not what someone would think of as a literary novel', I kind of think, well, I wouldn't read a literary novel for fun so in that respect {group laughter} there's certainly more similar than a lot of, you know... |
| Physical Book | Explicit or implicit references or activities connected to the physicality of the printed generated novel, physical aspects of reading, or print layout, presentation. | Maybe because it was sort of more like a book, you know, it was broken up into chapters, I sort of knew how to behave with it maybe { P5: mmhm} you know? I mean I don't know. So yea, I think it depends you know whether these like these extra |

| | | |
|---|---|---|
| | | information is helpful or not, or like whether it's necessary or not. |
| Paratext | Explicit or implicit references to the project paratext or accessing it. Existing works, such as novels, that are referenced as an example or comparison to generated novels. | I found the, the GitHub, for this and I had a little look at that and sort of the ideas behind it. And I hadn't read the book that it's based on, so I didn't really get that reference. But having read about it, and read the wikipedia page for that book, the concept then made a bit more sense { P1: chuckle}, I must say. |
| Puzzle, Game, Play | Explicit or implicit references or activities where the generated text content and/or the physical book form are engaged with playfully or ludically, as a puzzle or game. Examples or ideas using the generated text and/or physical book as material for a game or puzzle. | P5: I mean it does, I do think of it as like there's a game element to it. Like I, I like reading because there's, it's more of a creative exercise for me to read it because it's hard to interpret basically. So, and it's like very open to interpretation as well. And in Molly's Feed like l said it felt like I was solving a puzzle because it wasn't making sense otherwise. |
| Reading Experience | Explicit or implicit (physical or psychological) reactions, emotions, difficulties, and responses to reading the generated text, or the text in a physical form specifically. Reading methods, strategies, or habits. Reading socially, reading out loud, or sharing the experience. Interpretation and sense-making, interpretive strategies. | … But Molly's Feed was like getting really difficult, whereas Victorian I could happily read for an hour and could probably, yea, have continued with that { P5: yea}. But Molly's Feed I, if I had just read it without reading it for a specific reason, as in, you know, your workshop, I would definitely not have spend an hour on it {P7: chuckle} probably not even 15 minutes {chuckle}. I got, yea I got very, as you can tell I really did not like Molly's Feed {L: okay} very much {chuckle}{P7: chuckle}, but Victorian I enjoyed quite a lot. Yea but I did both of them in one sitting. |

I worked the data through a second iteration phase to gradually develop themes, the middle analysis, and developed the following codes: *Reading and interpretation/sensemaking strategy, Referring to human intention in the programming or technical design, What is a generated novel, The impact or the function of authorial paratext for the reader, The impact or the function of authorial paratext for the reader, Author is human, Comparing valuing of reader interpretation vs. author's intention, Balancing or judging Human-Computer creation, Enjoyment Reading, Reading interactively or ergodically for enjoyment, Identifying artifacts or traces of the technical process in the text,* and *Improving the physical book*. In the Nvivo analysis software, I hierarchically structured some of these codes under others as I wrote reflexive memos and developed towards the conceptual model. For example, *Identifying artifacts or traces of the technical process in the text* and *Referring to human intention in the programming or technical design* were structured under the inductively developed *Code or Project Workings*. These codes are seen in *Table 15: Middle Analysis Codes*, where the first column documents the initial sensitizing constructs and early analysis phase codes which I revisited to in order to develop the middle analysis codes. The names of these newly developed codes are shown in the second column. Finally, the third column contains the inclusion criteria for the coded excerpts, and the last column shows an example excerpt from the workshop transcript. Hierarchical structuring is not represented in the table in favor of visual clarity.

Table 15: Middle analysis codes

| **Early analysis codes used for development** | **Developed Code Name** | **Description** | **Example** |
|---|---|---|---|
| Creativity, Framing, Reading experience | Reading and interpretation/ sensemaking strategy | Reference to the reader creatively interpreting the generated text on their own, versus being aware of the authorial or editor's/critic's paratext to help with interpretation. | it's like layers { P1: mmm}. Like you could read the thing without thinking about who created it, why they created it and find your own meaning to, like interpret it however you want. But I feel like knowing this is interesting because maybe it's better than anything you can come up with on your own. Like it's better than a framing that you can make for yourself. |
| Author related, Code or project workings | Referring to human intention in the programming or technical design. | Explicit or implicit references to an author or creator or programmer intentionally following a technical process that can be traced or notes in the generated text. | … But there are like repeating themes and repeating sentences even sometimes like pages apart { P4: mmm} So, I mean even if I did think of it as keywords it wasn't clear enough to me what they were, or like what the people who were creating it what they were going for with. Like I couldn't see a connection between makeup and Russians for example, like I, I couldn't figure out any structure to the keywords if there were keywords. So I wasn't, I didn't think it was intentional. Or if it was intentional I don't know what the intention is. |

| | | | |
|---|---|---|---|
| *Identified during analysis* | What is a generated novel | Grappling with the question of what a generated novel is or where it belongs, or analogies. Experts with text matching the search term "name on" for the phase "put my/their name on a book". | … it's quite an impressive sort of facsimile of a novel. It's almost like the equivalent of like a cardboard cut-out to a human, right? And so {P7: mmm} something, something you can imagine something like this would be a good addition to those, for those books which are just as decoration on bookshelves. So like you know in, in hotel lobbies… |
| Author related, Paratext | The impact or the function of authorial paratext for the reader. | Authorial intentions. Benefits of knowing authorial intention. Different interpretations with or without knowing the intentions. Presence, absence, and amount of authorial intention which should be presented in the physical book. Creator's progress page (Piece D). | I got more enjoyment from {chuckle} from understanding how they had come to be {L: okay} than from the actual content. |
| Author related, Paratext | Author is human | Realizations of the author being 'real', or comments on the author being human. | P7: I, I also Googled the, because I, I saw Janelle Shane and though, hmm that sounds like someone's come up with a vaguely Jane Austen sounding name, and no I, turned out they're both real people {group laughter}. |
| Reading experience, Author related | Comparing valuing of reader interpretatio | Reader's uninfluenced interpretation, author's intention. Comparing them to discuss which is | *L: Is that that important for, to enjoy the reading if, if you did, that there is some intentionality here with how the themes or the style |

| | | | |
|---|---|---|---|
| | n vs. author's intention | best for reading experience. | or whatever was structured? Or was it not?<br>P5: I think for me it's not important at all. Especially with AI art the fun is more interpreting it than figuring out the intention. {inaudible} yea. |
| Creativity, Code or project workings | Balancing or judging Human-Computer creation | Discussing the (differing levels) of involvement of machine, computer, AI, or code in creating the project in comparison to the human author. Who or where creativity lies. Comparing differing levels of involvement and creativity in the two generated texts. | … I think you could get a lot more out of it if you had a healthier balance between like the AI agency and the human agency that was involved in that curation { P1: mmm}, and you could come up with a nice happy medium where you'd use some of the really like inventiveness of the AI, but like the coherence and the narrative that a human could provide to sort of mold it into something that was more enjoyable to read for an hour. |
| *Identified during analysis* | Enjoyment Reading | A clear example of a participant reporting enjoying reading. The experience or process is described as fun, interesting. The reader explicitly or implicitly bring their creativity to the interpretative framing. | P5: Because a lot of it is, I mean, there's some really good sentences in between that I really liked so I feel like the language is very…I dunno, I just wanna like keep reading and not all parts of it might be interesting, but I feel like there's a lot of like seeds for creativity in Victorian. So like it's kind of, like I would read it. I mean I read a lot of it with just the first sentence, because that was fun {chuckle}. Like the, it, the universally acknowledged truths, like reading just the first |

| | | | |
|---|---|---|---|
| | | | sentences for each of the chapters was a lot of fun. |
| Creativity, Reading experience | Reading interactively or ergodically for enjoyment | Reading the generated text in an interactive way, choosing to read ergodically (with non-trivial effort) for enjoyment. | P5: Yea, I thought the same. Like the empty pages and ending things midway felt inviting of collaboration and creativity from me. I thought that was nice. |
| Generated text qualities or content, Code or project workings | Identifying artifacts or traces of the technical process in the text. | Explicitly reasoning or speculating about the code workings or technical processes to explain text qualities. Text qualities are artifacts of the technical process. These observations are not always accurate. *Excluding* references to an author or creator or programmer intentionally following a technical process. | I mean 110 jumps out as one, but there are multiple pages which just say, 'I know you know that I made those mistakes maybe once or twice one hundred eighteen OTWOLFFaceOff and I know that you know that I made those mistakes' {group laughter}{ P4: yea}, it just goes on. It's like it's got stuck. Like each, each repetition of that prompts it to go back again. So, it it does make you think about the process as well as just what you've ended up in front of you. |
| Physical book, Paratext | Improving the physical book | Suggestions or discussions about how to improve the physical book with the inclusion, exclusion, and placement of paratextual pieces. | P6: Just kind of like I guess the author's intention like the, the bit where they are describing what they did and why. Yea, I mean not necessarily like the whole entirety of the code, but just kind of like the auth-, I guess the author's statement would be nice to have at the end. But not necessarily at the beginning because then it would possibly sort of distract from, from what was going on {L: okay}. |

Some of the sensitizing constructs and codes were developed into two primary themes: *Latent Links Between the Generated Text and Technical Aspects*, and *Reader Creativity for Framing*. These describe the interplay between several codes and capture complex yet generalized trends from the workshop data. They were able to be made into distinct themes in later workings of the conceptual model, and I note that the development of the model was a challenging but useful means of crystalizing the primary themes.

### 5.7.1. Primary Theme 1: Latent Links Between the Generated Text and Technical Aspects

The primary theme *Latent Links Between the Generated Text and Technical Aspects* captures the latent links that participants seemed to make between the generated text and the technical aspects and paratextual pieces (such as Piece B, the code) when they were discussing the generated novels. A detailed expansion of this theme has been made in section *5.6.1.1. Latent links to technical aspects* where it is part of answering the first PhD project research question. Tracing this primary theme's development, the initial kernel advances from the deductive sensitizing constructs *Author related* and *Generated text qualities or content*, and the inductively developed *Code or project workings*. It then progresses with the more complex codes *Identifying artifacts or traces of the technical process in the text* and *Referring to human intention in the programming or technical design*. Although not necessarily captured by the code names, these codes collectively work to capture the apparent trend of participants speculating or not confirming their assumptions about how exactly the text processing algorithm functions. The relationship between this theme and other concepts developed from the workshop analysis will be discussed in section *5.8. Conceptual model*.

### 5.7.2. Primary Theme 2: Reader Creativity for Framing

The primary theme *Reader Creativity for Framing* captures the participants' ergodic reading and interpretation strategy where they creatively frame the generated text with a narrative vignette. A detailed expansion of this theme has been made in section where it answers the second PhD project research question. This primary theme develops from the deductive sensitizing constructs *Reading Experience* and *Creativity,* plus the inductively developed code *Framing*. The more complex code *Reading interactively or ergodically for enjoyment* advanced the codes towards the theme development. Interestingly, although there were very few coded excerpts between either of the generated novels and codes *Puzzle/game/play, Enjoyment reading,* and *Reading interactively or ergodically for enjoyment*, participant AH's positive response to a

question about whether or not there was something game-like or ludic about the generated novels explains that the game-like aspect describes the enjoyable processes of solving the interpretation challenge through creative framing:

> *P5: I mean it does, I do think of it as like there's a game element to it. Like I, I like reading because there's, it's more of a creative exercise for me to read it because it's hard to interpret basically. So, and it's like very open to interpretation as well. And in Molly's Feed like I said it felt like I was solving a puzzle because it wasn't making sense otherwise.* (p. 15)

Similarly, participant AH's response to my request for them to elaborate on their interaction with the generated text gives a more detailed account of an ergodic approach to reading, responding to, and framing *Molly's Feed* with characters in a makeup narrative vignette. Note that P5 refers to specific passages and reports sharing their framing with participant P2 before the workshop:

> **A: Oh, yea. That happened a lot more in Molly's Feed actually {chuckle}. Because I realized it that more social media element I found it very easy to sort of, in my imagination put faces to the sentences {group chuckle/acknowledgement} to who it was writing these things. So for example, P2 and I actually agreed about this – the makeup saga throughout the {group laughter}. It's when she was saying about putting on mascara on her bottom eyelashes thinking it will look different but it never does {P2: chuckle}, it looks like spiders on my eyes. I was like ha! The struggle is real {group laughter}. And then {L: what page was that?} on page 3. But then she talks, whoever it is, I imagined it was a woman, just mentions a younger girl actually, someone who is just learning to use makeup essentially. Then on page 5 again goes – didn't put mascara on my bottom lashes today. And all I've written there in capital letters is, the saga continues {group laughter} so, you know, very similar to, you know if I saw somebody posting that, you know, on my facebook or twitter feed or whatever, I'd be like haha that's an update about that. It's just one of those things I found really easy to, to interact with. The same with, when P7 was talking about the jets and the Russian or are they not Russian, and they repeat that somewhere earlier in the pages. I just started writing next the that, bots. I was like oh gosh, I am starting to think about this as if I would be reading that on a social media page and if I thought it was constructed. So yea it was, it was interesting. And then started making me think about how I felt sorry for Molly who was having to like learn about these people's mascara and bots and things like that.* (p. 9)

## 5.8. Conceptual model

While developing the conceptual model I iterated over the workshop data as I refined it and progressed through a total of 5 drafts. I do not consider the model to be complete, set, nor definitive as it is based on just one data source – the reading workshop. It is however a model that future confirmatory testing could be based on. The purpose of this model is not to offer a model reader perspective, but rather to better illustrate and synthesize some of the processes and concepts that I developed from my data.

As seen in *Figure 18: Conceptual Model Showing Different Potential Ways to Interpret A Printed Generated Novel*, the diagram shows an abstracted processes of reading and interpreting a generated novel. As based on the workshop analysis data, the model indicates three concepts which impact new readers' interpretation and reception of printed generated novels in book form: *Authorial intention*, *Latent links between the generated text and latent aspects*, and *Reader Creativity for Framing*.

### 5.8.1. Describing the conceptual model's layout

The generated text is seen at the center of the conceptual model where it is surrounded by the physical paratext and can only be accessed through it: the book's cover, chapters, pages, and so forth. At the bottom of the diagram two *Reader* icons represent two potential ways to read the generated text – with a reader's own creative framing or without. Both ways use a continuous reading loop where the reader parses and then interprets the meaning of the text. The *Reader* to the left of the diagram represents a potential reader who does not create their own creative framing with which to interpret the generated text. The *Reader* to the right of the diagram takes a fictional stance towards the text as they parse it and approach their reading ergodically as they traverse the text; working to (re)read it non-linearly as they familiarize themselves with its style, patterns and quirks, content, and so forth. Ideally, the reader's ergodic approach helps them to develop their own narrative vignette framing with which to interpret the printed generated text through. Note that the concept of creativity (as captured by *Reader Creativity for Framing*) is located outside of the generated novel and situated with the reader, and not with the *Project Paratext* or *Wider Paratext*.

The paratextual pieces referred to during the workshop are seen at the top of the conceptual model where they are split into the *Project Paratext* and *Wider Paratext*. The former encompasses project pieces B, C, G (the Code, the Code repository, and the Input) which could be considered as the more technical pieces. These are latently connected to the reader's parsing and interpreting loop in order to represent how a reader might identify artifacts or other traces of the technical process in the generated text, and how a reader might also potentially

not verify their speculations and assumptions about the technical aspects. A dashed line is therefore used to represent a latent, potentially speculative link between interpreting the generated text and the *Project Paratext*. If a reader has accessed the *Project Paratext* they can choose to link to the *Wider Paratext* and, as described by participant P2 during the workshop discussion, they can read a broader range of paratext with intertextual links to the *Generated Novel* and *Project Paratext* such as *Other Works, Classic Novel/Hypotext,* and *Wikipedia*. Returning to the *Project Paratext*, the reader could also choose to access *Piece D,* the Creator's progress piece. This might feed into the reader's interpretation of the generated text in the form of *Authorial Intention*.

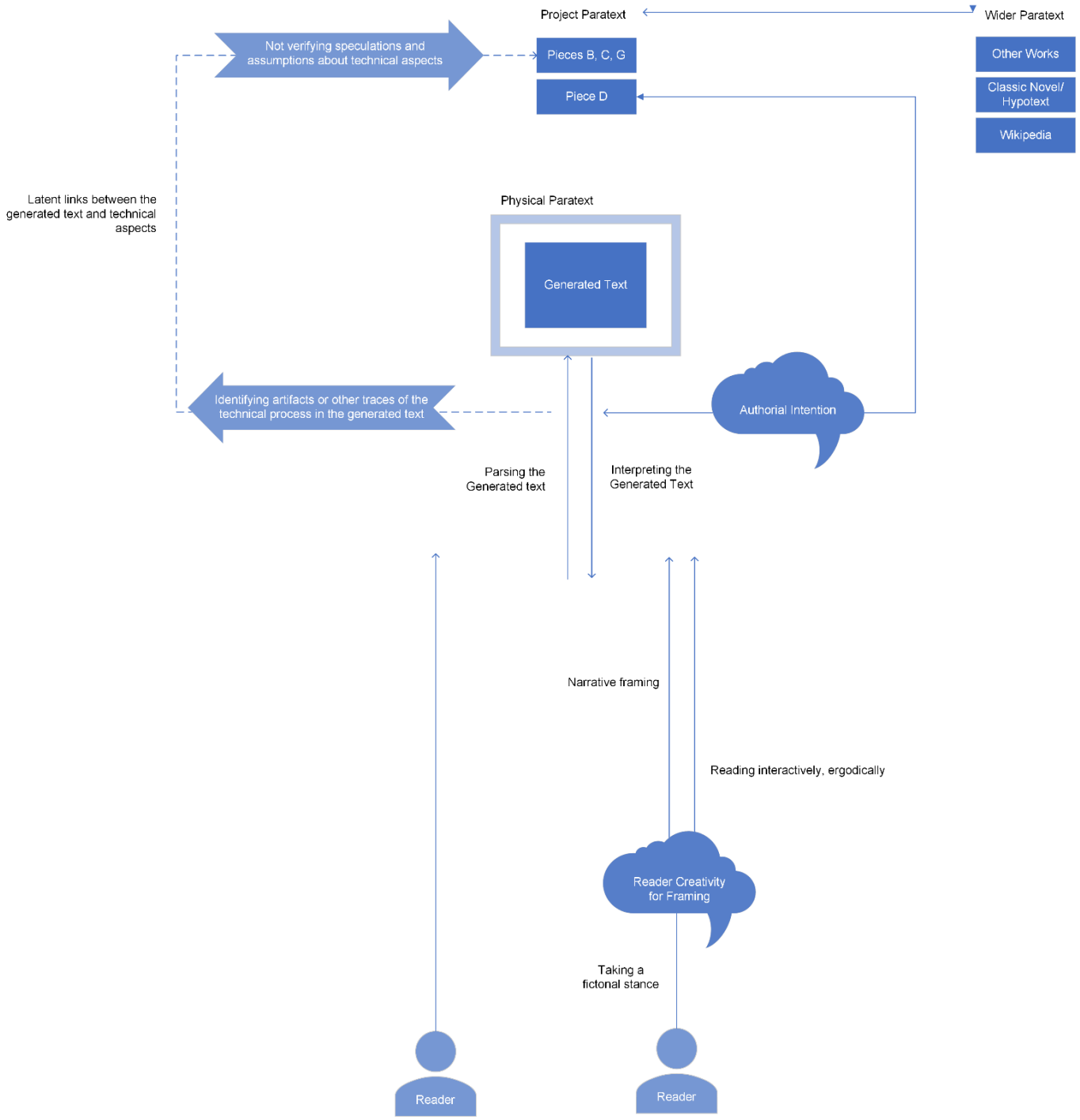
This conceptual model is intended to reflect what participant P5 explained during the workshop; that a reader can use and create several interpretations as they engage with a generated novel (such as their own framing and additionally reading along the grain of the author's intention). These are additional interpretive layers which help to make the reading experience more interesting:

> *P5: I think knowing how it was generated adds to the experience, but I think I find value in just reading jumbled text as well because it's fun to sort of, it's like, it's like a creative exercise for me to interpret it, and that's fun. Even if it is jumbled text. But knowing where, like how it's created makes it interesting in more ways. Like appreciating it from a meta level* (p. 11)

Thus, the model shows how a reader could interpret the printed generated novel through all three concepts: *Authorial intention*, *Latent links between the generated text and latent aspects*, and *Reader Creativity for framing*. This model is limited in that it does not capture any social aspects or social reading activities, although that falls outside the scope of this current study.

To comment on the proposed critic's framing piece introduced in section *5.6.1.2. Authorial intention and a proposed critic's framing*, if such a piece were included as part of a generated novel's physical paratext it could furnish both of the readers in the conceptual model with an additional paratextual lens with which to parse and interpret the generated text, thereby potentially improve their reception of the generated novel. This could be especially relevant and impactful to the *Reader* to the left of the diagram who might choose not to, or be unable to, develop their own creative framing. Finally, based on the reported enjoyment of workshop participants who described an ergodic approach to reading and interacting with the printed generated novel, I recommend that the proposed critic's framing piece should ideally aim to offer the reader a narrative vignette framing with an ergodic quality to it, where the reader is offered interpretive guidance but also some textual qualities, content, or patterns to recognize or discover - to narratively or creatively 'solve'.

Figure 18: Conceptual Model showing different potential ways to interpret a printed generated novel



## 5.9. Limitations

The generalizability of the workshop discussion results are limited in that the data describes only one instance of a reading group discussion and focuses on only two generated novels. The extent to which the survey study and the workshop study results are comparable is of course also limited because they differ in almost all respects, including in method and data collection. Perhaps the only similar element between them is that both studies ask participants to read *Molly's Feed* and *Victorian*. Unlike in the survey study where participants were asked to read a sample of approximately 1000 of the generated text, the workshop participants were given the complete NaNoGenMo texts of 50,000 words or more. Because both generated novels have a fairly homogenous *text*ural quality, comparisons of reader impressions which focus on the content of the generated text and its qualities can comfortably be made when accounting for linear and ergodic reading approaches.

The analysis presented in this chapter is influenced by my formal experiences around generative literature which includes reading academic literature, attending online Electronic Literature book launches and conferences presentations, and participating in Critical Code Studies Working Group sessions. In each of these, scholars and authors seem to put considerable emphasis on reading a project's code, to the point where I questioned whether the generated text itself actually was the (only) central element in a generated novel. While this expectation on my part and previous experiences have shaped my line of inquiry, I can also confidently report that despite my efforts to identify it, I was not able to show in this data that a generated novel's code is of particular conscious importance and explicit interest to most new readers when compared to other project pieces.

## 5.10. Chapter Conclusion

To link the results of this chapter to the survey study, the importance that the workshop participants give to authorial intention and their idea for a proposed critic's framing piece broadly align with the survey's ranking question results where the contextualizing pieces E, D and F (the NaNoGenMo page, the Creator's progress page, and the Media page) are consciously ranked by participants as the most important pieces, with the latter two being the closest in content and function to authorial intention and a critic's framing. Similarly, the more technical pieces, B, C, G (the Code, the Code repository, and the Input) were consciously less valued by the survey participants even though piece G was a sample of *Ulysses* which, for a supposedly lower value piece for understanding and interest, was referenced several times in the written reviews. This difference between seemingly conscious and unconscious judgements or

valuations amongst the survey participants is similar to the workshop analysis' *Latent Links Between the Generated Text and Technical Aspects* primary theme in the sense that possible or explicit connections to paratextual pieces are not fully realized by readers. This second comparison between studies is an initial thought and could benefit from further development. However, it will not be expanded upon further in this project because the research questions have been answered.

# Chapter 6 Discussion

This chapter opens by presenting an integration of answers to the research questions in section *6.1*. In section *6.2*, I take a firm stance by answering the primary research question by drawing from my data. Here, I develop a critical discussion of relevant theory based on the unexpected results relating to literary value in my studies. The critical discussion synthesizes relevant literary theory from reader response theory and narratological theory of postmodern narratives in relation to generated novels, and arrives at a theoretical explanation for the unexpected results. In section *6.2* I therefore contextualize my arguments within relevant literary theory areas and with my data. Finally in section *6.3 Revisiting Paratext*, I build on my data and reflect on my research process to arrive at and propose my own minimalist paratextual conceptualization (section 6.3.1), which calls for a refinement of Genette's conceptualization. My conceptualization is a research contribution. After reengaging with the discussion I began at the start of this research project in section *2.1 Paratext* and *6.2.5 Reflecting on Genette's paratextual conceptualization* based on my data, I conclude the chapter by discussing the implications of my data for using the paratextual model.

## 6.1. Integrated answers to research questions

Answers to the research questions have been detailed at the end of both study chapters. They are summarized again here with a focus on integrating the combined results from both studies where applicable, although it's borne in mind that not all aspects of the studies are comparable because of many major differences: they use different types of data (quantitative survey responses, many short-form text reviews, and one long-form text group-interview transcript) the data was collected about a year apart (spring 2021 and summer 2022), the studies use different analysis methods (statistical testing, Qualitative Content Analysis, Deductive Qualitative Analysis, and textual analysis when analyzing reviews and transcript excerpts), and they use different versions of the two generated novels (1000 work samples presented in a digital PDF file followed by webpages presenting the paratextual pieces, and the full 50,000 word works presented as print books with the digital paratextual pieces separated and remaining online). Indeed, one of the few concrete repeat elements between the two studies is that they both use *Molly's Feed* and *Victorian* as study items.

### 6.1.1. Answer to research question 1

Which paratextual elements have a central role in influencing the understanding and reception of a project?

In the survey study, pieces A, E, F, and D (Generated text, NaNoGenMo page, Media page, and Creator's progress page) were explicitly ranked as the most important for helping to understand the project and helping to make it interesting. Similarly in the workshop study, the generated text in physical book form is continually referred to and specific examples from it are described and quoted throughout the workshop. It is clearly treated as the central piece. Authorial intention is reported to be highly valued by the workshop participants, although this is difficult to show in the data; i.e. participants *say* they value authorial intention but this is difficult to *show* where they demonstrate this beyond fairly broad statements. For example, authorial intention was not discussed at length nor asked about. Nevertheless, this result fits comfortably with a high importance ranking for piece D, which functions as the primary location of authorial intention. Both piece F and the proposed critic's piece from the workshop would function as critical commentary, and both were considered valuable or important to participants in both studies. These three pieces, A, D, and F, and their workshop study counterparts were explicitly chosen or engaged with by the participants from both studies. Based on the rest of the survey results, though, the generated text, piece A, didn't actually significantly impact participant responses. Therefore, across both studies, while the generated text may be the central piece, the most impactful paratextual pieces are critic's commentary or framing, and those conveying authorial intention.

### 6.1.2. Answer to research question 2

Which reader skills and experience affect the influence of these paratextual elements?

Focusing on the survey study, from a statistical perspective previous knowledge or familiarity with the classic work which bears an intertextual relationship with the generated novel project could not be shown to impact literary value judgements. However, several reviews did assume that knowledge or understanding about programming, or knowledge or familiarity with *Ulysses* would have improved their impression of the project. This suggests that at least some participants expected paratext to impact reception when it's understood or is familiar. And yet, this could not be shown empirically in the specific cases of the survey study. While it was not tested quantitively, the ability and willingness to interpret and frame the generated text creatively or narratively is a participant skill that enabled a better reception of the generated text. These two results have interesting implications for generated texts because the former perhaps reveals reader assumptions about the imagined value of unknown or ununderstood paratexts, and the later shows that reception can be improved with trained, creative readers rather than with 'better' texts. Of course, the same could be said for any literary works.

### 6.1.3. Answer to research question 3

Which technical and cultural characteristics of a generated novel affect the influence of these paratextual elements?

In the survey study, participants' cultural beliefs and expectations about literature and literariness appear to be deeply set and difficult to shift. The workshop study echoed and expanded on this where it seems that the generated nature of the projects, despite each having differing levels of human authorial involvement, violates the ability for the projects to be confidently accepted as literature. Both studies did have participants who expressed a more liberal acceptance. For example, the survey study contained reviews which happily referred to the generated novel as a combination of literature and computing. The workshop study had participants who either negotiated for AI Literature as a separate category (thus not literature proper) or expressed an enthusiastic but still cautious suggestion that AI literature might be a type of literature. However, all the workshop examples and many of the survey reviews refer to literature with a (limiting?) qualifier.

## 6.2. Answer to primary research question as drawn from the data and knowledge and synthesis of relevant theory

In this section I engage with my primary research question to launch into a broader discussion linking to relevant literary theory.

My primary research question asks, in what ways does paratext influence potential readers' interpretation and reception of a generated novel? Based on the quantitative survey study results, in terms of reception the presence of the paratext positively impacted readers' perceived quality of creativity, technical, enjoyment, interest, and understanding. This is also reflected in the workshop study qualitative data, where participants agreed that a contextualizing and framing piece written by a knowledgeable critic/editor is a paratextual element that could have the greatest impact on a new readers' understanding and positive reception of a generated novel as a physical, printed book. Thus, there is some clear evidence suggesting that paratext can positively impact NaNoGenMo novel reader reception.

In terms of interpretation, the qualitative results of the workshop study indicated that for some readers, an ergodic reading strategy and working to develop a creative framing for the generated novel is what contributed to an interpretation and an enjoyable reading experience.

However, the studies' unexpected results around literary value are not easily explained by data analysis alone. But they can be interpreted through knowledge and synthesization of

assumptions and theories about literary works; namely, from reader response theory and poststructuralist perspectives on narratology. In this section I work to contextualize the unexpected results by developing an argument which takes stock of why judgements of literary value might have been the only quality that was virtually unimpacted by the generated novels' project paratext.

To contextualize my research with relevant theory, I identify some complications and mismatched assumptions which are encountered when applying a reader response lens to generated novels and to my research project, and I then advance to identifying some complications in narratology. Finally, I arrive at scholarship in postclassical narratology that focuses on postmodern literature which I connect to generated novels, and from that vantage point I argue for the unique epistemic position that generated novels can occupy within an emerging literary epistemology characterized by the digital turn. The main point I will make here is that generated novels do not easily fit within current literary epistemes because they are theoretically able to occupy a mode of engaging with the narrative event that is theorized to be impossible for literature: the mode of representing the event. I link this back to the unexpected results by proposing and theoretically evidencing that generated novels belong to a different literary episteme than what actual readers and literary theories have typically been prepared to account for.

### 6.2.1. Iserian reader response

I begin my discussion with reader response theory because it's a literary theory framework that focuses on readers and not exclusively on the text. It also makes clear several assumptions about literary works. By mapping out and critically engaging with some of reader response theorist Wolfgang Iser's core points, I pinpoint exactly where they mismatch and present complications with my research and focus. This then enables me to navigate towards ideas in post-classical narratology which comfortably fit with and enable a better conceptual understanding of the question of literariness and the generated novel. Both my PhD project and Isearian reader response theory focus on the reader and text, and this is why I chose to critically engage with Iser in this chapter. I don't engage with other prominent reader response theorists such as Stanley Fish (known for readers and interpretive communities) and Norman Holland (known for readers and psychoanalysis) because my study designs and data collection focused on how individual readers respond to text, and not on social or psychological data collection and frameworks.

One of the core complications with several reader response theories (and Iser's is no exception) is that they focus on the conceptual, model reader and not on actual living readers as my research does. To understand some of the complications and mismatched assumptions with

applying a reader response lens to generated novels, I being by discussing digital media researcher Jim Pope's 2010 participant-based reading study which focuses on another (then) emerging form of Electronic Literature – digital hypertext literature. Pope 2010 surveys and interviews actual readers of an emerging Electronic Literature form, and to the best of my knowledge it is the closest related work to mine within that field. Of course, there are still many differences between our studies. For example, a difference in the Electronic Literature forms, sample sizes, methodology, and focus (Pope does not engage with the concept of paratext).

Despite diverging greatly from traditional print novels, Pope (2010) is nevertheless able to successfully apply concepts from Wolfgang Iser's reader response theory to hypertext literature because several assumptions hold: narrative, meaning, and a traditional conception of authorship can be taken for granted in a work of hypertext literature – but not in generated novels. However, a theoretical mismatch is present in Pope's work. Pope refers to aspects of Iserian reader response theory when interpreting their study data which was collected from observations of actual readers. Yet, Iserian reader response explicitly focuses on the model reader – specifically, the implied reader. Unlike the actual living readers in my own and in Pope's participant reading studies, an Iserian reader is not an actual flesh-and-blood reader. Iser explains that "…the implied reader as a concept has his roots firmly planted in the structure of the text; he is a construct and in no way to be identified with any real reader." (1978, p. 34). Pope doesn't explain their reasoning for using Iserian concepts developed for a implied reader in order to discuss assumptions about the behaviour of several actual living readers. For example, Pope states that their study's post-reading questionnaires (which are presumably quantitative) are based on (among other scholars) the theoretical reading models of Iser (1978) (2010, p. 7). But they don't specify exactly which aspects of Iser's work they drawn on, nor the rationale for why their data collecting questionnaire is partly designed from theory based on the implied reader. While they do also reference work from researchers David S. Miall and Don Kuiken (p. 88) who are known for their empirical study of literary readers (and therefore a logical foundation from which to design actual reader data collection[57]), Pope has not explained why and how they were able to combine implied reader and actual reader scholarship and research, and what the implications for this might be. Nevertheless, Pope applies implied reader concepts to explain their study data. This is possible for Pope to do because several assumptions made by Iserian reader response theory hold true for the hypertext works in Pope's study – but they do not hold for generated novels. The implications of this is that I can't easily use reader response theory to interpret my own data without excessive omissions, exceptions,

[57] In my own survey design, I am careful to draw on research that is based on empirical studies of actual readers, such as Miall and Kuiken (1999), Henrickson (2019), McGregor et al. (2016), Koolen et al. (2020), Kuijpers et al. (2014), and Lamb et al. (2018).

and changes to theory. And this is not necessarily a problem; reader response theory and the model reader are simply not the right theoretical fit for my study data and for understanding the reading of generated novels more generally.

For example, Pope draws on Iser's consistency building and the concept of gaps and blanks to explain their participants' reported reading experience. Pope suggests that the greater interpretive challenges that their study's readers faced may be because the readers are encountering multimedia aspects ("...gaps and blanks...of many different kinds" (p. 83)) of the hypertextual works, rather than the lesser interpretive challenges of print and film. As already discussed in *Chapter 5 Workshop study*, it's clear from my qualitative data that the interpretive challenge for the generated novel readers is very high; certainly much more than is expected when reading a typical novel. Pope reasons that their participants' interpretation challenges (i.e. ultimately not being able to "...reveal the underlying story") stems from an inability to fill Iserian gaps and blanks and build narrative consistency (2010, p. 83). However, for several reasons I am arguably unable to similarly draw on Iser to describe my data. Firstly, because the emerging generated novel form (in this case namely *Molly's Feed* and *Victorian*), has a relatively low level of coherence, there isn't much to be gained in terms of analysis and insights about the form through the lens of gaps and blanks, especially when the readers in my *Chapter 5* workshop study instead described finding patterns and interesting or meaningful passages. Some readers worked to interpret the generated text through creative narrative framing in areas of the generated text, rather than consistently building a meaning across all of the body of the text. My core point here is that generated novel readers don't appear to focus on elements of the text that can be described as gaps and blanks (for there are far too many for that to be productive). Rather, those 'void' areas might be better described as traversed-over rather than filled or connected with narrative sense. Indeed, it would instead align better with the workshop study data to say that creative frames and patterns are developed or connected by readers. But equally, some participant readings appear to be more akin to ergodically searching for points of interest much like needles in a textual haystack. In essence, then, my data indicates a very different reading strategy and experience than what Iser and Pope appear to be referring to.

Indeed, secondly, as discussed in section *1.3 The generated novel*, one cannot assume that narrative and meaning are components of a generated novel. Conversely, the certainty of a narrative and an intended meaning is a given for both Iser and the hypertext works that Pope focuses on. Throughout *The Act of Reading* (1978) and in other sources, Iser also takes for granted that a work has characters, an author in the typical sense, and presumably hasn't been created through aleatory composition or creative constraints. For example, in *The Implied Reader* (1974) Iser writes by taking for granted that meaning, potential meaning, and narrative are discovered, and this suggests that for Iser these are already present in the text and

somewhere in the space between the text and the implied reader. This is in contrast to what my data suggests; that narrative and meaning can be (and perhaps need to be) developed. Recall however that not all actual readers of generated novels do this. Although it's unclear whether Iser means actual readers or implied readers, they do pair the assumption of meaning and narrative with the condition that the readers of a novel are "…forced to take an active part in the composition of the novel's meaning" and that this active participation is fundamental to the novel (p. xii). While Iser of course doesn't write about generated novels, I agree that active participation is also fundamental to a more enjoyable and satisfying reading of a generated novel. But based on my data from the workshop study, I have described this active participation as ergodic reading strategies and creative framing because of the considerable additional strategies and effort required from actual readers to do that. Considerably more that is, than Iser meant and could have possibly anticipated with regards to the generated novel form.

When introducing *The Implied Reader*, Iser marks 'discovery' as the dominant theme of that book, where "The reader discovers the meaning of the text", and also marks discovery as a form of "esthetic [sic] pleasure" in reading (p. xiii, 1974). Comparing this to my own studies, discovery is not a code or theme that was developed from my data. Although as Iser further discusses this theme and links it to consistency building, they highlight the implied reader's discovery of their own "…faculties of perception… [and their] tendency to link things together in consistent patterns, and indeed the whole thought process…" where the novel "…deliberately reveals the component parts of its own narrative techniques …[and]…the reader is forced to discover the hitherto unconscious expectations that underlie all his perceptions" (p. xiv). Interestingly, Iser's suggestion that a reader's unconscious expectations and realization of their own faculties of perception being discovered is seen in the qualitative data from both of my studies. For example, in *Chapter 4 Reading experiment and survey study* some of the written reviews describe expectations about novel writing not being met, or experiencing tiredness from reading (coded under *Reading experience* as described in *Table 14: Finalized categories codebook*). However, this agreement between Iser and my data cannot be sustained. Iser's quote stands as one of many examples where Iser assumes that meaning and narrative are there to be discovered. But, in the sense that Iser and Pope intend it, these simply cannot be expected from a generated novel. It's therefore a source of complication because I cannot make these same assumptions. Similarly, Iser's focus on the conceptual implied reader is tricky for me to work with and reconcile methodologically because, practically speaking, it's the polar opposite type of reader and method. These complications renders reader response theory difficult to confidently apply to my data.

To further expand on the complications I encounter with the implied reader, Iser stresses that the implied reader (as well as several other types of model readers surveyed by Iser in section

*Readers and the concept of the implied reader* (1978, pp. 27 – 38)) doesn't refer to living readers like the participants in my and in Pope's studies. Although they're "…drawn from specific groups of real, existing readers", they are "…primarily conceived as heuristic constructs" (p. 30). Therefore the type of reader that Iserian reader response focuses on doesn't align with the actual, living readers that my data has been collected from. My PhD project focuses on understanding the impact of paratext on readers' interpretation and reception, and it aims to fill the gap on a lack of research that produces data which evidences actual reader responses and reading strategies of generated novels. Indeed, scholarship in Electronic Literature which does not focus on actual living readers already arguably approaches works with an underlying model reader framework, such as Carter (2020) and Marino (2009). But my PhD project set out to do something else; to focus on under-studied actual readers and not on model readers. Therefore, taking an implied reader model simply does not comfortably fit with my studies of approximately 300 actual readers. Further, I ran a large online experimental study which is not only very different in scale and aim from Pope, but in a different methodological paradigm from Iser.

What Iser does focus on are specific, actual texts. They reiterate again in *The Act of Reading* (1978) that their focus is on texts (p. xi) and not on actual readers, and neither is it on a philosophically inflected theory through which to analyze a text. With regards to the latter, Iser warns that this would reduce the text to a secondary role in favour of prioritizing theory – something that they emphatically stress that they do not want to do (p. xii). Thus, actual literary texts are the focus for Iser, whereas my focus is on actual readers and the impact of paratext. Indeed, my research doesn't focus on a specific generated novel itself (*Molly's Feed* or *Victorian*) and it's not where my PhD project's major contributions lie – they lie in better understanding the impact of paratext on actual readers of generated novels. Thus, Iser and I have different focuses and different goals. To put it more plainly, the implied reader and the studied, actual reader are two different things which call for different methodologies and analyses.

Iser's implied reader model ceases to function upon encountering the generated novel; the model assumes that the text is enough to guide the reader, but my data suggests that the generated novel readers in my studies can read ergodically and create their own creative framing, or refuse to do so entirely. My data also indicates that reading reception significantly improves with paratext, and yet Iser neglects to consider paratext even though they are deeply engaged in theorizing about readers and print books. Indeed, without considering paratext one might wonder whether Iser is also theorizing a model text in the sense that it is detached from physical print books produced by publishers.

Finally, in *The Act of Reading* (1978) Iser appears to not engage with nor even consider the concept of paratext. Where Genette (1997b) goes into great detail to explain and taxonomize

authorial and publisher paratext and its intention to steer reader interpretation, a discussion or even an indirect reference to the potential impact of paratext on reading is surprisingly absent from Iser (1978). This absence is also apparent in *Prospecting: From reader response to literary anthropology* (1989) where an interview with Iser and Norman Holland (a literary theorist who takes an empirical and psychoanalytical approach to reader-response) shows that despite both scholars' focus on literary works (we can safely assume that they were referring to printed and published works), neither of them seems to acknowledge the presence nor the potential impact of paratext on Iser's implied reader nor on Holland's empirical reader. On the one hand the lack of consideration for paratext is surprising because the publishing industry's role in producing literary books and and paratext for print media are of course a reality for Iser and Holland. But on the other hand it should be noted that these sources predate Genette's published works in English on the architext and intertextuality. Therefore, one might speculate that paratextual impact on readers and reading may not have been a point of interest or relevance for Iser's concept of the implied reader at the time.

Drawing this discussion section to a close, it's important to bear in mind that perhaps one of the biggest differences between Iser's work and my own is that Iser is focusing on a well-established form where several assumptions can safely be made. I, on the other hand, focus on an emerging form and work from within an episteme characterized by the digital, pattern, data, and human-computer interaction. Therefore, unlike Iser I must carry out foundational groundwork to better understand the form and how it is read (which I do through a reading experiment, a group interview, and data collection and analysis). Whereas Iser's stable, establish form does not require this groundwork. Indeed, Iser focuses on literary works with well-established (physical) forms, such as novels and poems available in print, and begins publishing about reader response in English in the 1970s; before born-digital media, digital multimedia, digital interaction, and before the digital user experience became more prolific and relevant.

### 6.2.2. Narratology and Postmodern Literature

In order to conceptually solve one of complications I identify in Iserian reader response theory, I ask, can there be a literary work without a narrative? The reasoning here being that, if the generated novel form isn't expected to be able to inherently have a narrative, that means that it could potentially be able to be a type of literary work and belong to literature. More plainly: if all literature doesn't have to have narrative, then generated novels could be considered as a type of literature.

In *Towards A Postmodern Theory of Narrative* (1996), literary theorist Andrew Gibson offers narratology an alternative by drawing heavily on concepts from continental philosophy.

Focusing on postmodern works in their chapter “Narrative and the Event”, Gibson’s path through previous scholarship is guided by their question “how and in what contexts might it be possible to think of narrative in terms of the event?” (p.195). I am not bound to this question, so I read Gibson with the strategy of answering ‘is it possible to think of literary works in terms of not having narratives?’. Gibson’s question therefore offers an answer: “…in terms of the event”. In this section I focus on a selection of Gibson’s theoretical argumentation and not on their examples from postmodern literary works.

Gibson moves towards discussing the event by opening their chapter with a discussion about two contrasting forms of time: chronos and aion. Chronos functions as a measure of events and can been seen as a temporal aspect of causation. This is the form of time used to construct chronological narratives. Aion can be seen as cyclical time; “it is the dimension of surface effects” (p. 179). Aionic time is “…time as difference, from time as a multiplicity in which the elements [or, the events] ceaselessly vary and alter in relation to others; a time to be thought in terms of fission” (p. 184). I propose that a generated novel’s absence of narrative, and therefore an absence of chronologically enforced control, can be conceptualized as fitting within a “…plastic space” (p. 187) of aionic time, where the textural qualities of a generated text exhibit aionic chain reactions of continuous textual fission events, alternations, ruptures, and surprisal. Where chronos is paired with narrative, Gibson aligns ‘the event’ with aionic time as its counterpart. Where narratives are structured and controlled by forms of chronological time, events are unstructured emergences of effects and varying intensities. Within the context of a generated novel’s text, I would describe these effects collectively as textural qualities.

Compared to the narrative time of chronos, the concept of aionic time is of course much more readily applicable to describing the patterns, unexpected variations, tropes, and absence of sustained logical causation generally seen in generated novel texts. This is the first of connections I make where Gibson’s description of postmodern literary text characteristics happen to also fit the generated novel surprisingly well. Surprising, of course, because Gibson wasn’t referring to generative literature, nor, to the best of my knowledge, to works of aleatory composition. By marking this connection I am actively building towards arguing for the potential literary status of generated novels.

Conceptually, generated novels can be received by readers as being structured by aionic time and the event, rather than by chronological time and the narrative. Gibson critiques narratology scholarship for only thinking within the confines of chronos. This criticism points to a lack of exploration in time and narrative rhythm, which seems to have trapped narratology in a local maximum where the possibility for developing new ways of thinking plateaued. Gibson points out that the field therefore even lacks the terms and capacity to ask questions with which to describe other forms of time and to describe postmodern writing (p. 184), or indeed, I point out,

generative literature. I reason that if narratology itself has been confined to thinking through chronos, then it's reasonable to assume that the typical reader (such as those in my studies) is expecting to be able to interpret a literary text through a regime of chronology. Of course, with generated novels this regime is nowhere to be found; it can however be imposed by the reader.

Binding narrative to chronos, Gibson draws from continental philosopher and literary theorist Jean-François Lyotard to explain that narrative

> ...is constructed on a founding difference: a dissymetry between beginning and end, initial and final situations. 'Telling a story' is itself an introduction and elaboration of that dissymetry, and an ordering of it in terms of succession. Narrative as diachrony does not disturb or transgress the linguistic order, but rather confirms it in its irreversibility. So, too, in its orientation towards an end, its gradual alignment of dissymmetrical features, its final ordering, it pacifies difference, puts it in place within a system...Thus narrative neutralises the event in ceaselessly recuperating the other into the same (Gibson,1996, p. 187).

For Gibson, then, causal sequences and narrative arcs are a form of temporally enforced control. Of "...homogenizing connections..." (p. 185). The narrative text is presented as neutralized, coordinated, and programmed within its closed narrative system (p. 182). In this system, I add, destabilizations and ruptures are managed and events are controlled; they do not emerge. Conceptually, I note that this contrasts with generated novel texts because they do not have an authored narrative. Instead, control and development of narrative framing falls within the purview of the reader. Based on the insights I developed from my study data, a reader might choose to read ergodically and identify and temporally sequence selected events, then creatively develop and coordinate a narrativization of the generated text. An example of this is the tracing of recurring themes and creative narrative framing performed by readers during my reading workshop study.

Taking from philosopher Gilles Deleuze's work on the metaphysics of the event and the poststructuralist's interest in narrative, Gibson gives three Deleuzian modes of engaging with or simulating the event: representing, narrating, and writing the event. Because Gibson uses postmodern works for their examples, I make the assumption that these three modes are intended to describe or categorize literary works or effects. Gibson states that the first mode is, strictly speaking, an impossible one where "...there can be no representing of the event" (p. 200). For Gibson, in any text the event is always registered or reported with a distance, "...mediated or muffled by the process of registration" (p. 195). Deleuze seems to bypass this impossibility of representing the event by thinking of narrative as a simulacra of the event (p.

195), but Gibson's point nevertheless stands – the event cannot be directly represented in narrative and in the postmodern works that Gibson engages with.

### 6.2.3. Synthesizing theory

But what about generated novels which don't have a narrative? Are they able to represent the event(s)? I have already established that the emerging form's textural qualities fit very comfortably with Gibson's descriptions of aionic time and the event. And now I propose to develop this further and take Gibson literally; to conceptualize generated novel texts as actual representations of events. I propose that creative generated text can be conceived of as the textual event itself, where the generated text is the material output of the aionic event of processed textual patterns, differences, and ruptures directly output at some level of the system event. Even though a typical NaNoGenMo project will go through some post-processing and paratextual wrapping in order to present it as a generated novel, one of the main aesthetic points is to preserve the generated textual quality rather than to edit it away. Arguably, this is also a preservation of the 'eventness' - of the event traces[58]. The implications of 'the generated text as event' is that the emerging form is therefore theoretically able to belong to Gibson's impossible, first mode of engaging with the event. The reason that Gibson mentions this first type of 'impossible' mode is because the mode or category must logically exist. So, just as narrating and writing the event must exist as modes, so too it must follow that the mode of representing the event must exist as a possible category, even if Gibson was not able to identify literary works which belong to it at the time of writing. Indeed, Gibson is able to identify the works of abstract expressionist painter Mark Rothko as constituting an art of the event (p. 187), which indicates that creative works can belong to the mode of representing the event. So it's reasonable to say that Gibson themselves saw types of works that could belong to the first mode - just not text-based works of literature. Therefore, it's more accurate to say that the mode of representing the event in literature is not an impossible mode per se, but rather one that (for Gibson at least) undiscovered forms could potentially belong to.

Of course, in this section I have been building up the case that the generated novel is a type of work that fits remarkably comfortably into the first of the three modes that Gibson uses to categorize literary works. Logically, then, this therefore suggests that generated novels could therefore be a form of literature because it easily fits into Gibson's category or mode of representing the event. Here I am making the argument that, theoretically, the generated novel could be seen as an emerging class of literary work or object, and that its difficulty in being

[58] It is possible that in order to fully appreciate this argument one needs to experience creating a generated novel.

recognized as such is because it belongs to an other literary episteme. I mean episteme in the Foucauldian sense where it configures the frameworks, rules, and norms that structure what knowledge and which questions are considered to be possible or legitimate within a given historical period; the episteme as the "conditions of possibility" at a particular time and place (May, 2006, p. 56). I propose that this emerging literary episteme is characterized by the digital turn, the absence of narrative, and perhaps the ergodic and creative framing reading strategies that readers can perform. Indeed, consider the unexpected results from the survey study where literary value was virtually unimpacted by paratext. And consider the workshop study data where a participant described generated novels as distinctly 'AI literature', and another who specified that AI literature could be a type of literature (section *5.6.3. Answering research question 3*). This might indicate a reader's difficulty in having access to suitable categories, terms, concepts, literary frameworks, theories, and epistemic rules with which to approach and think about generative literature.

This idea has already been echoed and referenced earlier by Gibson, where they complain that even narratology can lacks the terms and capacity to ask questions with which to describe other forms of time and to describe postmodern writing (p. 184). This inability to even consider new questions is arguably indicative of epistemic differences. Indeed, I reason that the best theoretical explanation as to why I am able to offer an emerging literary form to fit with Gibson's 'impossible' but logically derived category or mode of representing the event is precisely because of an in-progress shift in literary epistemes from 1996 (the publication year of Gibson's work) to the present. Of course, epistemes do not suddenly shift overnight; the shift can be gradual, and more than one can be in relevant use at the same time. For example, one might wish to conceptually group ergodicity in hypertext literature, postmodern features, and the data and distant reading processes associated with the digital turn as elements which are all moving towards a distinct literary episteme – if not already in one.

My proposition of there being an emerging literary episteme that has the concepts and capacity to consider generative literature as literature is a reasonable proposition to make. Consider, for example, that fields which have traditionally stayed within a textual analysis and theory mode of research are able to balance both their traditional approaches, and also oscillate (although not without friction) between modes which belong to the digital turn. In the case of the Humanities this is seen with the Digital Humanities, where big data, patterns, testing, computational methods, experiments, and new forms to interpret are hallmarks of the field (Bod, 2013), and I add, signs of a developing digital episteme. Thus, I put forward the proposition that, based on my data, the possibility of accepting a of generated novel as having literary status is conceptually dependant on the literary episteme that the generated novel is received and interpreted within by readers. This can also theoretically serve to explain why many readers

expressed polarized opinions and negative responses in the *Chapter 4* survey study, and the negative responses expressed in the *Chapter 5 Workshop Study*. To borrow from Gibson's chapter 'Narrative and monstrosity', Gibson engages with continental philosopher Michel Foucault to explain that that which falls outside of a dominant episteme can be seen as unnatural and be rejected as monstrous:

> "For Foucault, the monstrous is that which is exiled by the normative judgments within a given episteme. …It is 'denatured', 'unnatural'…it is also a kind of treachery to social norms…In other words, monstrosity is epistemic illegitimacy understood as outrage in or against nature." (p. 238)

Certainly, for some of the readers in my studies the generated novels violated too many expectations or norms, were seen as unnatural in comparison to typical authored novels, and ultimately appear to be outside of the literary episteme within which some readers were situated. Indeed, Gibson's figurative use of the spatial term 'exiled' echoes back to the relocating of literary value to technical value theme that I developed from the written reviews survey study data in section *4.5.2.3.3.2 Novel B both conditions*. This theme describes how some readers appear to reject the idea that generated novels could have literary value, and that value is located outside of whatever the reader understands, perhaps, as the dominant literary episteme. Thus, the theoretical reason that I have built up to explain why readers' judgements of literary value might have been the only quality that was virtually unimpacted by the generated novels' project paratext is because some readers might be working within a literary episteme that is not easily compatible with generated novels.

## 6.3. Revisiting Paratext

This section revisits the critical discussion which began in section *2.1 Paratext*. Here, I offer my minimalist concetpualization of the paratextual model based on my experience with this project's research processes and based on my data. Throughout this research project I continue to alternate between naming paratext a theory, concept, conceptualization, model, and an idea. I have not settled on which of these is the best term to use because researching the nuanced meaning behind terminology and terminological development is outside the scope of this project. I have instead chosen to use my research time to study paratext in action and to better understand how the concept might apply to real-world data, the research questions, and to free myself to focus on developing my own ideas about refining Genette's conceptualization.

Finally, I reflect on Genette's paratextual conceptualization based on my data, and I discuss the implications of my data for using the paratextual model.

### 6.3.1. My minimalist paratextual conceptualization

Because of the critiques about Genette's paratext that I have discussed in sections *2.1.2 Criticism of paratext* and *2.1.5 Paratext extensions and reworkings* at the beginning of this research project, I considered that I would need to offer my own major reformulation or revamping of Genette's paratext. But I was less sure about this need after analyzing the results in *Chapter 4 Reading experiment and survey study*. After the analysis and discussion presented in *Chapter 5* I concluded that the paratextual model is a lean, transferable concept which is usable and valuable when it is treated as a simple means of describing, and organizing or segmenting a particular work or media form. By lean, I mean that paratext as an idea is rather thin in terms of theoretical apparatus and complications. But I argue that this is its greatest strength because it renders the concept of paratext straightforward to transfer and apply to current and future emerging new media forms. Indeed, as an organizational and sense-making concept (and as I have demonstrated in my two studies), paratext lends itself extremely well as a means of rationalizing or motivating the delineation or operationalization of a work into distinct pieces, or groups of pieces, which have different functions, strategies, intentions, and interpretations. For example, in section *4.4.5. Ranking the pieces* not only did I use paratextual theory to conceptualize the NaNoGenMo pages into pieces A-G, I also made sense of their top of the list and bottom of the list groupings by considering the similarities that the pieces shared in terms of their paratextual function. Therefore, my position is that revamping by adding complexity to Genette's initial conceptualization of paratext should be avoided as much as possible because it is already a highly useable concept.

In order to refine Genette's paratextual conceptualization for my own, I remove Genette's restriction that only authorial or publisher paratexts are legitimate. This is a specification that prevents the concept from being easily applied to works and forms that have proliferated since, for example, the web era. Since the web, elements produced by readers, fans, and so forth are perhaps easier to create, find, link, and attribute now than they were before 1987 when Genette's book on paratext (1997b) was originally first published in its native French, titled *Seuils*.

My refining of the paratextual conceptualization is a research contribution. My conceptualization approaches paratext as a conceptual tool that can be used to describe a specific set of relationships, functions, and actors. Such as an individual work or a form of media that is composed of one or more central texts, and supplementary or periphery elements which fall in a range of physical or conceptual proximity to the text. In my conceptualization, there is at least one producer (the creator, publisher, and in some cases marketer) who is

presumed to want to steer reader interpretation to some degree, and one or more interpreting readers (or similar: viewer, player, etc).

My conceptualization considers the following to constitute a paratext: any distinct or reasonably definable element that is part of, linked, or otherwise related to a central text(s), regardless of how and why it was produced. So, while authorial and publisher approval is not a requirement in my simpler paratextual conceptualization, a sense of definable 'entity-ness' is. For example, I would avoid considering a reader's nebulous sense of distrust of algorithmic automation as a paratext not because it isn't impactful or relevant to understanding their interpretation of a work, but because it is simply too difficult to pin-down and account for within the paratextual conceptualization. It is not a specific piece that can be pointed to, defined, and its relationships and creators traced. I would however consider specific items which don't have an easy to define (human) producer or reception shaping aims as paratext: such as metadata, digital receipts, system statuses, and other byproduct artifacts of technical infrastructure. Whether or not such a level of paratextual detail would useful or meaningful is a disciplinary, methodological, and research question issue, and not a complication or limitation of the paratextual model itself.

In the name of good research practice, I recommend that it may be pertinent to investigate or test (as I have done in *Chapter 4 Reading experiment and survey study*) which element is treated as the central text or texts by readers, and if interpretation and reception is meaningfully impacted by the presence of certain paratextual elements or readers' backgrounds. But, I stress, not taking an assumed central text for granted and doing the work to study, and to determine, and to locate where function and impact might be located across a work's or a forms' elements doesn't invalidate or complicate Genette's paratextual conceptualization. This is simply the act of putting the model into practice by using paratext as a conceptual tool to structure, operationalize, and study a produced work and its audience impressions. The difficulty here may be in delineating one paratextual element from another, or the text from the paratext, but I argue that this is a question of disciplinary method or operationalization rather than a question that necessarily challenges the paratextual model itself.

Indeed, determining how to delineate, segment, and structure elements as part of a paratextual model is something that is already done by researchers, although it often isn't formalized as a step in the process of using paratext as part of a study. My minimalist paratextual conceptualization therefore foregrounds this. My recognition and formalization of the paratextual model as a conceptual tool or method of segmenting and analyzing works is an original research contribution. I demonstrate that this recognition and formalization can be applied to quantitative and qualitative studies in my study Chapters *4* and *5*. This segmentation and analysis can also be seen in other research as well, even if it isn't conventionally recognized

as such. For example, logically speaking, identifying the central text and paratextual pieces is actually implied in Genette's conceptualization, because doing so is a prerequisite for describing a print novel through a paratextual lens. Indeed, Gray (2010) also demonstrates this step when they describe a Disney film marketing campaign (previously discussed in section *2.1.3.2. Franchising and transmedia storytelling)*. Although Gray's example film might ostensibly appear to be the central text in a campaign as far as consumer consumption is concerned, Gray's closer analysis of the marketing elements challenges this centrality from the producer or marketing viewpoint. While at first this example may seem to challenge mine and Genette's conceptualizations because there does not seem to be a fixed element at the center of this constellation of paratextual pieces, Gray is nonetheless able to identify the impact and function of numerous marketing elements in relation to the producer, consumers/audiences, and names what the latter may consider to be the central text. Thus, Gray ultimately uses Genette's conceptualization to structure, operationalize, and successfully describe the Disney film campaign through the lens of paratextual theory unproblematically. The core of my argumentation here hinges on treating the paratextual model as a conceptual tool which can be used to structure, describe, and study: producers, a work or form, readers, interpretation, and reception. Based on my experience of successfully carrying out this projects' research process, I argue that the paratextual model should be used as a descriptive tool, and not as an inflexible, prescriptive regime that keeps running into the same problems when new media doesn't comfortably fit into Genette's *example* of the print book.

My minimalist refining of the paratextual model doesn't aim to be able to capture and describe all complexity. Afterall, a theory or concept exists as a generalization, and by definition cannot and should not aim to capture a full breadth of nuances and exceptions. Instead, I propose that it is more useful to consider that paratextual theory can only describe limited aspects of the reader, the work, and its constellation of elements, and it cannot account for all the facets of interpretation and reception. I propose that seeking to describe a form or process through a paratextual lens should also consider that an additional theory or theories may be needed to fully account for and explain the phenomena or subject of study, rather than risk over-engineering a bloated and tedious revamping of the paratextual model which can end up being impractical as a conceptual research tool. Therefore, if a particular work or form doesn't comfortably resolve through a paratextual lens, then perhaps the answer is simply to find another concept that is better suited to describing that work. This first point can be illustrated with an example from my workshop study results. In section *5.6.1. Answering research question 1*, examples of the latent links to technical aspects might not clearly connect to identified paratextual elements; but they do not necessarily have to be considered as complications or short-comings of the paratextual model. I propose that the 'complications' may instead be considered as indications of theories or established concepts which interplay with paratext at

the site of the generated novel. For example, an ergodic text and complex sociotechnical relationships such as folk theories of AI (introduced in section *5.6.3.1*) might be ever-present confounding variables when studying generated novels. This is because of the multifaceted nature of the generated novel as an emerging medium, and not because mine or Genette's paratextual conceptualization is necessarily inadequate or brittle; especially because it already is a rather simple, lean, and flexible idea about how different elements and extensions or connections of media works function for producers and readers. An example of established concepts or theories interplaying with paratext can be seen in de Bruin-Molé (2018). Digital media scholar Megen de Bruin-Molé focuses on and critiques transmedia story strategies in Disney's *Star Wars* franchise, but alongside the concept of transmedia also introduces Genette's concept of paratext in order to continually refer to and discuss specific elements, relationships, and audiences (2018).

### 6.3.2. Reflecting on Genette's paratextual conceptualization

During the workshop study, participants suggested that an expert written paratextual piece positioned at the back of the printed generated novel (the proposed critic's piece) would be a welcome addition to helping them interpret, understand, and enjoy engaging with the generated novel through several 'meta' layers. If such a piece were present in a self-published or especially a commercially available printed generated novel, then it would surely count as a publisher-approved paratext. Thus, this reader suggested proposed piece does not pose a challenge to Genette's paratextual conceptualization - it fits comfortably with Genette's vision of paratext.

Genette only considers authorial and publisher pieces to be legitimate paratexts, which is not how the term has come to be used in wider media studies research, including in my own use of the term and in the scholarship I have reviewed and discussed in section *2.1 Paratext*. However, I don't find this to be an irreconcilable difference. Genette (1997b) focuses entirely on the print book, and by only acknowledging author and publisher materials Genette is essentially decreasing the scope of their own work within a specific and highly commodified print medium, which makes sense to do in the context of writing an academic work. What media studies scholars have gained by instead accepting paratextual pieces regardless of who made them is of course an optional distinction between author/publisher paratexts, and between other paratexts. Therefore for my own minimalist paratextual conceptualization I have reframed Genette's restricted focus on 'official' paratexts as an optional descriptive category rather than a prescriptive restriction.

In introducing my minimalist conceptualization I propose that we can successfully treat paratext as an instrument or tool through which to structure and analyze media and contexts (as

I have demonstrated in my studies), rather than to be frustrated with its inadequacies when treated as a full-blown conceptual framework or totalizing theory in a way that Genette arguably didn't develop it for. Genette's intention is plainly seen in the introduction to *Paratexts: Thresholds of interpretation* (1997):

> "The approach we will take in studying each of these elements, or rather each of these types of elements, is to consider a certain number of features that, in concert, allow us to define the status of a paratextual message, whatever it may be. These features basically describe a paratextual message's spatial, temporal, substantial, pragmatic, and functional characteristics. More concretely: defining a paratextual element consists of determining its location (the question where?); the date of its appearance and, if need be, its disappearance (when?); its mode of existence, verbal or other (how?); the characteristics of its situation of communication - its sender and addressee (from whom? to whom?); and the functions that its message aims to fulfill (to do what?). This questionnaire is a little simplistic, but because it almost entirely defines the method employed in the rest of this book" (Genette, 1997, p. 4)

Instead *Paratexts* carefully details which elements exist for print books, where they are located, contextualizing explanatory examples, and why they are relevant to shaping interpretation. So rather than laying extensive theoretical groundwork and expounding a framework, Genette focuses on explaining the taxonomization and demonstrating procedure for identifying and motivating the paratextual value of a range of elements. This is clearly demonstrated in, for example, 'The cover and its appendages' (pp. 23-32).

By the end of the *Chapter 5 Workshop study*, several of the participants seemed to prioritize the (imagined) authorial paratext. This is in-line with Genette's focus on authorial and publisher paratexts. Brookey and Gray (2017) points out that Genette overestimated the amount of control that the author has over the paratext around their published book. While this is valid in general, coincidentally NaNoGenMo projects are self-published so this happens to fit with Genette's view rather well to an extent.

### 6.3.3. Implications of my data for using the paratextual model

The implications of my data for using the paratextual model to study generative literature is that paratext is versatile; it can be applied to both a model reader (as Genette implicitly does) and to an actual reader approach (as I do).The survey study's qualitative data showed that the presence of the NaNoGenMo project paratext (pieces B to G) together had a significant positive impact on actual readers' reception. The data also showed that the absence of the generated

text itself (piece A) had virtually no significant impact on actual readers' reception. Thus, the implications of these results and of using the paratextual model to study and test actual living reader responses is evidence that actual reader studies can supplement and expand on the understanding of how generative literature functions for readers.

While Pope (2010) studies hypertext literature with actual readers, they don't critically discuss and expand on the implications of model reader and actual reader responses. My data suggests that this is valuable to do because there are significant differences and assumptions which can hold in one approach but not the other. For example, statistically speaking my quantitative results in section *4.4.3 The impact of removing the generated text* suggest that a generated novel's text (piece A) doesn't positively impact reader reception; although some actual readers assume that it would be, as discussed in section *4.5 Qualitative analysis of survey reviews*. Further, because my results suggest that the presence of paratext can have a statistically significant positive impact on readers of NaNoGenMo novels, it's possible that this might also hold true for other emerging forms of generative literature and potentially other types of Electronic Literature more generally. Indeed, while the conventional assumption is that paratext might work to steer reader interpretation, my results build on this by indicating that paratext can contribute a measurable positive impact.

Further, these implications serve to outline the limits of theorizing about generated novels from a model reader perspective. In addition to my results which indicate that paratext is impactful, I also show that it's a practical and effective way of structuring and operationalizing NaNoGenMo novels for actual reader studies. Indeed, the versatility of paratextual theory can be appreciated when considering its structuralist origins and its ability to nevertheless have been extended into and arguably easily applied to (digital) media by more contemporary scholars as a means with which to study actual readers and 'actual' media. By 'actual' I mean that a work exists as a published object and is related to and embedded with peripheral, impactful paratextual elements and contexts. However, the versatility of the paratextual model often appears to be underappreciated. It's not convention for researchers to acknowledge paratext as a structuring approach or conceptual tool for segmenting and analyzing different aspects of a work, or for systematically thinking in more fine-grained terms about potential impact on readers.

In the initial stages of this PhD research project, I expected to encounter several problems with Genette's paratextual conceptualization and that I would conclude the project by offering a major theoretical framework revamping. However, this is no longer the case because I have instead come to realize how versatile and useful a minimalist paratextual conceptualization is. Through my study-based research process, what has surfaced is paratextual theory's utility as a conceptual tool in being able to easily discuss, predict, operationalize and test assumptions and understanding of emerging digital forms. In fact, although it's not explicitly recognized as

such, this is exactly what other new media researchers (discussed in section *2.1 Paratext*) like Švelch (2017) and van Dijk (2014) do to segment and understand videogame media and web-based electronic literature. Indeed, Švelch (2017), Henrickson (2019), and McGregor et al. (2016) use paratextual elements as a means of structuring, relating, and segmenting media to carry out their participant studies (although only Švelch explicitly recognizes related and segmented elements as paratext in their studies). Through my research process, I also tested paratextual elements in participant studies and, based on my results, i was surprised to find the concept of paratext unproblematic to theorize through, to practically apply to my study designs, and to interpret the qualitative data where readers discuss different aspects and elements of generated novels. Through this success, then, I'm able to articulate and develop my minimalist paratextual conceptualization.

To reflect on reader response theory: paratext, and specifically my minimalist conceptualization, is the conceptual tool that can be applied to actual living reader studies of digital emerging forms, such as NaNoGenMo generated novels. The paratextual model can also function as a supplement to frameworks which have been developed to focus on the model reader, such as the reader implied by Genette and the implied reader in reader response.

The model reader approach is of course useful, but my data (which was generated from applying a paratextual model to design the survey study) outlines a clear limitation of model readers - they cannot be tested and evidence new insights. For example, what the readers in my study believed was the most important or impactful element (piece A, the generated text), actually wasn't statistically significantly impactful. A model reader approach wouldn't have been able to question and to evidence this. In terms of where such insights may be applicable, paratext and an actual reader approach could be useful to digital publishing areas that works with emerging forms, and perhaps to generative tool design research as well. An actual reader approach is also useful to Electronic Literature because it demonstrates the different kinds of data and insights that can be gained with an actual living reader study, thus offering a supplement to existing research that focuses on the model reader.

# Chapter 7 Conclusion

This final chapter concludes the PhD research project by summarizing the ground covered, the analysis narrative, and the insights developed from *Chapter 2 Background and literature review* to *Chapter 6 Discussion*. The chapter begins with a summary of the research project's data analysis narrative from *Chapter 2* to *Chapter 5 Workshop study*, and then revisits the research design steps which underpin it. Then my research contributions are taken stock of. Finally, the research project's limitations are explained, and I chart possible future research directions.

## 7.1. Summary of analysis narrative

### 7.1.1. Literature Review

The paratextual theory literature review (*Chapter 2 Background and literature review*) formed the theoretical underpinning upon which I constructed my research questions. These were designed in a way that would enable me to empirically test how paratext functions to impact reader interpretation and reception so that I could draw my own conclusions how paratext works conceptually.

### 7.1.2. Quantitative survey

In *Chapter 4 Reading experiment and survey study*, I began the survey's statistical analysis by focusing on the impact that paratext had on readers' perception of literary, creative, and technical value. I also investigated the influence of paratext on perceived enjoyability, interest, and understandability. The results showed that the presence of the paratext did make a difference across all of these values – except for literary. In fact, the only case which had a significant difference in literary value judgments was when *Molly's Feed*'s generated text was presented on its own without paratext. These were unexpected results, especially because I expected a similarly culturally loaded concept – creativity – to be valued similarly to literary value. Next, I performed the same statistical analysis again to test the impact that the generated text alone had on readers' perceptions. But unexpectedly, the absence of the generated text from the project doesn't make a difference to reader's value judgements; it didn't appear to have any impact.

Because *Molly's Feed* and *Victorian* both have a hypotextual relationship with an esteemed literary work, it was unexpected that perceptions of literariness weren't influenced by the presence of the paratext. I wondered whether this was because there might be a group of readers who were not familiar with Joyce's *Ulysses* or Austen's *Pride and Prejudice*. And further,

whether it was possible that being familiar with the classic works led to perceiving the generated novel project as being more literary. But another statistical analysis showed that there wasn't any evidence for this – perceptions of literary value remained almost immutable.

Next, I wondered how important the other paratextual pieces or their function might be to readers. Based on the survey's ranking results, the pieces which worked to explain why the generated novels exist (the NaNoGenMo challenge page detailing the aim, the press article giving a readers' review, and the creator's progress page) were the most important for helping readers to understand the generated novel projects and to find them interesting. Notably, although the generated text wasn't considered to be important for understanding, readers prioritized it for helping to make the project interesting. Meanwhile, readers found the more technical of the paratextual pieces (the code, the code repository, and the input data) to be the least important overall.

### 7.1.3. Qualitative survey

My analysis then progressed onto the survey written reviews. In addition to answering the research questions I also wanted to better understand the unexpected literary value results. Why did the presence of the paratext with the generated text have no significant impact on participants' literary value judgements? And what was distinctive about the review responses to *Molly's Feed* when participants only read the generated text without the paratext? Using Qualitative Content Analysis, I coded the review data for categories and found that categories *Character*, *Art*, *Game-puzzle-interaction*, *Text Quality*, and *Reading experience* were all more frequent in reviews which were written by readers who only read the generated text. There were also differences in category frequencies between skill groups. But when readers were shown the generated text and the paratext together, categories *Joyce and Ulysses, Programming and algorithms, Authorship and writing, Literature and literary, and Creativity* were all more frequent in those reviews. Category frequency differences between skill groups were present here as well. Therefore, while skill differences were not able to be shown in the quantitative survey data, skill differences could be shown in the qualitative review data.

I wanted to have a deeper understanding of these differences, so I continued my qualitative analysis and identified that some of the reviews which read *Molly's Feed* highlighted the intertextual or hypotextual relationship between *Molly's Feed* and *Ulysses*. These reviews either described how the relationship between the two works impacted their impressions, or they assumed that the classic works would have an impact on value – namely interest and creativity. This was interesting because the quantitative survey analysis showed that this wasn't the case for literary value – familiarity with *Ulysses* did not meaningfully impact literary value ratings of *Molly's Feed*. In some of the reviews where readers read both the generated text and the

paratext, I identified a pattern and developing theme: literary value relocated to technical value. Here, some reviews rejected the idea that the generated novel project could have literary value. These reviews then seemed to instead (re)locate the project's value in technical or computational areas. While there were reviews which referred to the project in a way which didn't reject literariness or which accepted it, this wasn't a distinct, complex theme.

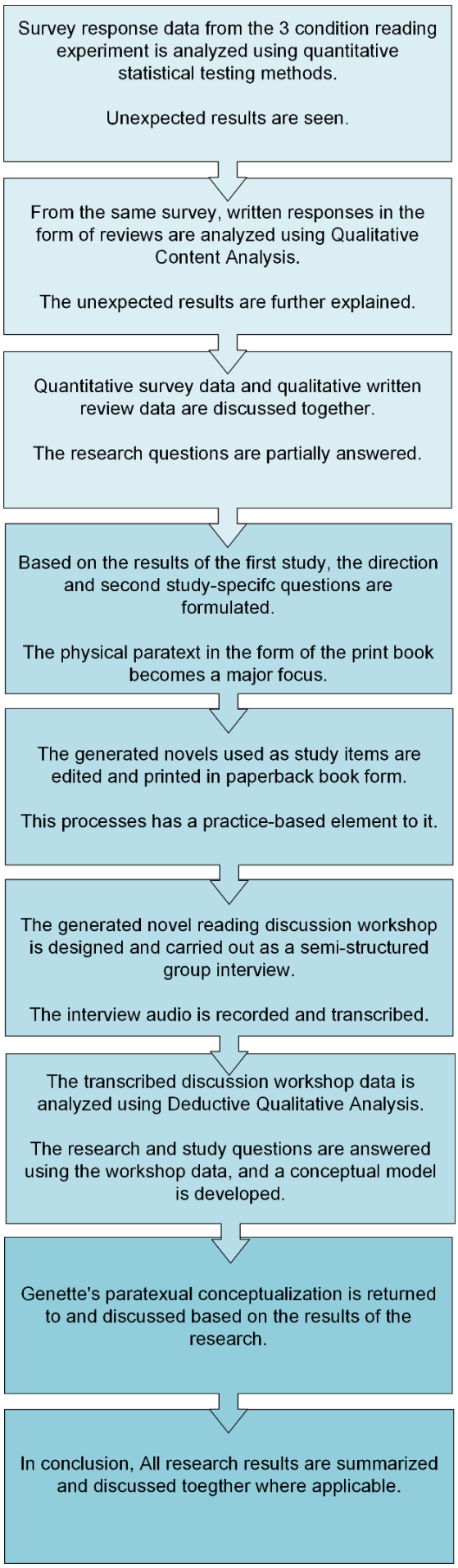
### 7.1.4. Reading workshop

Based on the survey study results, I wondered whether the medium that the generated novel is presented in would impact value and reception more positively, especially literary value. I also wondered in what ways would a physical, bookish paratext and formal reading activities impact literary value perceptions. I assumed that the physical paratext and the reading activities would lead to a broader acceptance of the generated novels' literary value, but my Deductive Qualitative Analysis showed that this wasn't clearly the case. My analysis did show that authorial intention, and latently the technical workings of the project, had a considerable role in impacting the understanding and reception of a printed generated novel. I linked insights about participants' interactive reading strategies to the concept of ergodic reading, which was able to explain some participants' ability and willingness to interpret and frame the generated text creatively or narratively. I identified this as a reader skill that enabled a better reception of the generated text, and this developed into a concrete recommendation for a proposed critic's piece which could be included in a printed generated novel and function paratextually to positively impact reception. Yet, analysis also showed that broader cultural expectations about authorship, and novels and books appeared to affect the ability of printed generated novels to be accepted as having literary value. Despite different levels of human authorial involvement, intention, and design, it seemed that the generated nature of the projects violated the ability for them to be able to be confidently considered along literature or literary lines.

*Generated Text and Technical Aspects* and *Reader Creativity for Framing*. These themes enabled me to develop a conceptual model which represented an abstracted processes of reading and interpreting a generated novel. Finally, this analysis development culminated in the crystallization of my own paratextual conceptualization.

*Figure: Research steps Overview* from *Chapter 3 Research Framework* is reproduced here for convenience to illustrative the research design steps which were performed as part of the analysis.

Figure 1: Research steps overview



## 7.2. Contributions

My theoretical refinement of the paratextual conceptualization is a major novel contribution to theory because rather than advocating for more complexity as other paratextual theory reformulations suggest, I argue that the concept of paratext is most useful to researchers when it is approached as a lean, generalizable model which enables it to be used as a conceptual tool to structure and study emerging media forms. This is detailed in section *6.2.4. My minimalist paratextual conceptualization*.

To the best of my knowledge, this research project is the first to conduct actual reader studies of generated novels with a large reader sample size. It's also the first to use mixed methods to empirically test and measure the impact of paratext on the interpretation and reception of generative literature works in the literary studies and media studies disciplinary space. This therefore has a methodological contribution aspect to it.

Additional novel research contributions yielded from this project include the collected data from both studies, offering the term folk theories of AI as a more specific alternative to folk theories of algorithms, insights about ergodic reading and creative framing interpretation strategies which are seen with an enjoyable reading experience and positive reception, a conceptual model which describes an abstracted processes of reading and interpreting a generated novel based on links with paratextual elements and other influences, and a research-based recommendation about which paratextual pieces are the most impactful to develop and include in print books versions of generated novels.

Finally, I have made a novel contribution in the form of contributing novels; I intend to donate a spare print copy of *Molly's Feed* and *Victorian* to a library or temporary library display. This will be done in an effort to stay aligned with the open-source and publicly available spirit of NaNoGenMo where the generated novels originated. If no library is receptive to the idea, then I shall do as one of the workshop study participants suggested and leave a copy of *Molly's Feed* on a bus.

## 7.3. Limitations and future work

Since the start of my Master's project focusing on NaNoGenMo works in 2018, and this PhD research project which began in late 2019, the text generation technology landscape has changed considerably. When I began the project there were the periodical travesty generator and predictive keyboard 'fake conference papers' being reported about, quirky Twitterbot

making site Cheap Bots, Done Quick![1] Was still working, and posts claiming to have 'forced an AI to watch 100 hours of Olive Garden Commercials'[2] abounded on my (now abandoned) Twitter feed. OpenAI's GPT models (the first one being released in 2018) were barely registering in public discussion despite news articles reporting concerns that these models were 'too dangerous to release to the public yet'. I therefore judged it to be relevant in terms of text qualities and technological relevance to include Shane's *Victorian* as one of my study items. This generated novel was created using a GPT-2 model (released in early 2019) that was fine-tuned on a crowd-sourced dataset. As can be seen in the workshop discussion excerpts, by the time the workshop study took place in summer 2022 there had been more public exposure and recognition of a neural network style of generated text. A few months after the workshop study OpenAI's ChatGPT was released in late 2022, and the rest is history.

The data collected from the 2021 and 2022 studies describes a sociotechnical point in time just before Large Language Models (LLMs) became freely available in English and were marketed and presented as writing productivity or information tools and features. As tools or features wrapped in interfaces which are accompanied by explanatory (or perhaps in future, regulatory) information, marketing materials, prompt writing tips and tutorials, specific language models and fine-tuning data, quality and safety evaluation reports, and a reader/user who perhaps sculpts and interprets output, I propose that emerging LLM products can be studied through the lens of paratextual theory. Indeed, this may be an exceptionally useful approach to identifying and measuring elements which impact reader/user perceptions, and which have been published by producers and other actors who undoubtedly aim to impact reader/user understanding and interpretation. Thus, if pursued this could translate into a transferable disciplinary contribution made from the literary studies' and media studies' paratextual theory, to the industry field of formalized user experience research and the study of emerging technology tools and their users.

I stress that generated text isn't only produced using LLMs, nor indeed only Neural Networks or other form of Machine Learning. Generated text has been around for a much longer time both in practical applications and as part of creative works[59]. The number of entries participating in NaNoGenMo has fallen over the last few years, and this might be because an easier to use or more (ostensibly) coherent text generation style has come to be increasingly accessible through LLMs (although I haven't studied this specific topic so I cannot confirm it). I don't know if the NaNoGenMo challenge will last in its current form, but I do believe that it's a treasure trove archive of generative literature works. I believe that generative literature will continue to unfurl

[59] See Tkacz (2019) for examples.

as an emerging form as technology and people's awareness of creative generated text develops. I certainly intend for my own creative pursuits to expand in the generative literature area based on what I have learnt from this research project and the ideas that have emerged from it. But based on the data that I have collected as part of this project, it seems unlikely that generated text technologies will be used to produce popular commercial novels aimed at book lovers (or at least an algorithmic involvement would not be marketed as such), precisely because of the cultural expectations about authorship and books that I have discussed *in Chapter 5 Workshop Study* – an algorithmic aspect is unlikely to be acceptable in the context of novels and books. Unless of course reader opinions change.

As explained in previous chapters, in this research project I have named paratext a theory, concept, conceptualization, model, and an idea. Determining the best terminological option is another piece of research in itself, and I therefore mark it as an aspect of potential theoretical development for future work.

## 7.4. Final remarks

To playfully reference back to Genette's focus on 'officialness' in terms of only considering author and publisher paratext to be legitimate, an amusing complication is the question of whether the printed generated novels used in the workshop study are in fact 'official' in any capacity. I am not an author, nor do I represent a legitimate publishing house. Indeed, one might argue that the only legitimate thing about this project is my approved use of research resources to arrange for printing, official university research ethics approval for the study, and my creative execution of the Creative Commons Attribution ShareAlike 4.0 license. It's true that universities are also in the business of publishing (research), but if I did not have the prestigious and legitimizing academic elements around my work, I wonder if Genette would view it as a sort of rogue paratext operation bordering on paratextual hijacking.

# References


Aarseth, E.J. (1997) *Cybertext: perspectives on ergodic literature*. Baltimore, Md: Johns Hopkins University Press.

Agafonova, Y., Tikhonov, A. and Yamshchikov, I. P. (2020) 'Paranoid Transformer: Reading narrative of madness as computational approach to creativity', *Future Internet*, 12(11), p. 182. doi: 10.3390/fi12110182.

Barker, M. (2017) 'Speaking of "paratexts": A theoretical revisitation', *The Journal of Fandom Studies*, 5(3), pp. 235–249. doi: 10.1386/jfs.5.3.235_1.

Bingham, A.J. (2023) 'From Data Management to Actionable Findings: A Five-Phase Process of Qualitative Data Analysis', *International Journal of Qualitative Methods*, 22, pp. 1-11. Available at: https://doi.org/10.1177/16094069231183620.

Braun, V. and Clarke, V. (2024) 'Supporting best practice in reflexive thematic analysis reporting in *Palliative Medicine* : A review of published research and introduction to the *Reflexive Thematic Analysis Reporting Guidelines* (RTARG)', *Palliative Medicine*, pp. 608-616. Available at: https://doi.org/10.1177/02692163241234800.

Braun, V., Clarke, V. and Hayfield, N. (2022) '"A starting point for your journey, not a map": Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis', *Qualitative Research in Psychology* [Preprint]. Available at: https://www.tandfonline.com/doi/abs/10.1080/14780887.2019.1670765 (Accessed: 17 May 2024).

Braun, V. and Clarke, V. (2021) 'A not-so-scary theory chapter: Conceptually locating reflexive thematic analysis', in *Thematic Analysis: A Practical Guide*. SAGE, pp. 155–194.

Brookey, R. and Gray, J. (2017) '"Not merely para": continuing steps in paratextual research', *Critical Studies in Media Communication*, 34(2), pp. 101–110. doi: 10.1080/15295036.2017.1312472.

Busselle, R. and Bilandzic, H. (2009) 'Measuring narrative engagement', *Media Psychology*, 12(4), pp. 321–347. doi: 10.1080/15213260903287259.

Carter, R.A. (2020) 'Tweeting the cosmos: On the bot poetry of The Ephemerides', *Convergence*, 26(4), pp. 990–1006. Available at: https://doi.org/10.1177/1354856519837796.

Colombatto, C. and Fleming, S.M. (2024) 'Folk psychological attributions of consciousness to large language models', *Neuroscience of Consciousness*, 2024(1), pp. 1–5. Available at: https://doi.org/10.1093/nc/niae013.

Culler, J. (2010) 'The Closeness of Close Reading', *ADE Bulletin*, (149), pp. 20–25. Available at: https://doi.org/10.1632/ade.149.20.

de Bruin-Molé, M. (2018) '"Does it Come with a Spear?" Commodity Activism, Plastic Representation, and Transmedia Story Strategies in Disney's Star Wars: Forces of Destiny', *Film Criticism*, 42(2). Available at: https://doi.org/10.3998/fc.13761232.0042.205.

Drisko, J.W. and Maschi, T. (2015) 'Qualitative Content Analysis', in J.W Drisko and T. Maschi (eds) *Content Analysis*. Oxford University Press, pp. 81-120. Available at: https://doi.org/10.1093/acprof:oso/9780190215491.003.0004.

Fife, S.T., D'Aniello, C., Eggleston, D., Smith, J., and Sanders, D. (2023) 'Refining the Meta-Theory of Common Factors in Couple and Family Therapy: a Deductive Qualitative Analysis Study', *Contemporary Family Therapy*, 45(1), pp. 117–130. Available at: https://doi.org/10.1007/s10591-022-09648-3.

Fife, S.T. and Gossner, J.D. (2024) 'Deductive Qualitative Analysis: Evaluating, Expanding, and Refining Theory', *International Journal of Qualitative Methods*, 23, pp. 1-12. Available at: https://doi.org/10.1177/16094069241244856.

Flores, L. (2020) 'MITH Digital Dialogues with Leonardo Flores'. University of Maryland, 10 March. Available at: https://www.youtube.com/watch?v=It2seaB5x-U&feature=youtu.be (Accessed: 20 March 2020).

Genette, G. (1997a) *Palimpsests: literature in the second degree*. University of Nebraska Press.

Genette, G. (1997b) *Paratexts: thresholds of interpretation*. Cambridge University Press.

Gibson, A. (1996) *Towards a postmodern theory of narrative*. Edinburgh: Edinburgh University Press.

Gilgun, J.F. (2019) 'Deductive Qualitative Analysis and Grounded Theory: Sensitizing Concepts and Hypothesis-Testing', in *The SAGE Handbook of Current Developments in Grounded Theory*. SAGE Publications, pp. 107–122. Available at: https://doi.org/10.4135/9781526436061.

Gray, J. (2010) *Show sold separately: Promos, spoilers, and other media paratexts*. New York University Press. Available at: http://ebookcentral.proquest.com/lib/soton-ebooks/detail.action?docID=865489 (Accessed: 7 April 2020).

Henrickson, L. (2018) 'Computer-Generated fiction in a literary lineage: Breaking the hermeneutic contract', *Logos*, 29(2–3), pp. 54–63. doi: 10.1163/18784712-02902007.

Henrickson, L. (2019) *Towards a new sociology of the text: The hermeneutics of algorithmic authorship*. Loughborough University. Available at: https://repository.lboro.ac.uk/collections/Doctoral_Thesis/4663709 (Accessed: 30 October 2019).

Henrickson, L. (2021) *Reading computer-generated texts*. 1st edn. Cambridge University Press. doi: 10.1017/9781108906463.

Koolen, C. *et al.* (2020) 'Literary quality in the eye of the Dutch reader: The national reader survey', *Poetics*, 79, pp. 1-13. doi: 10.1016/j.poetic.2020.101439.

Kuijpers, M. M. *et al.* (2014) 'Exploring absorbing reading experiences: Developing and validating a self-report scale to measure story world absorption', *Scientific Study of Literature*, 4(1), pp. 89–122. doi: 10.1075/ssol.4.1.05kui.

Iser, W. (1974) *The implied reader : patterns of communication in prose fiction from Bunyan to Beckett*. Baltimore : Johns Hopkins University Press.

Iser, W. (1978) *The act of reading : a theory of aesthetic response*. London: Routledge and Kegan Paul.

Iser, W. (1989) *Prospecting : from reader response to literary anthropology*. Baltimore: Johns Hopkins University Press.

Lamb, C., Brown, D. G. and Clarke, C. L. A. (2018) 'Evaluating computational creativity: An interdisciplinary tutorial', *ACM Computing Surveys*, 51(2), pp.1-34. doi: 10.1145/3167476.

Leitch, V. B. (ed.) (2001) 'Stanley E. Fish' in *The Norton Anthology of Theory and Criticism*. 1st ed. Norton & Company, pp. 2067–2070.

Littrell, Z. (2017) *Novels 'Written' for NaNoGenMo*, *Book Riot*. Available at: https://bookriot.com/novels-written-nanogenmo/ (Accessed: 13 June 2021).

Livingstone, S. (2019) 'Reception Studies', in *The Blackwell Encyclopedia of Sociology*. John Wiley & Sons, pp. 1–3. doi: 10.1002/9781405165518.wbeosr031.pub3.

Livingstone, S. and Das, R. (2013) 'Interpretation/Reception', in Moy, P. (ed.) *Oxford Bibliographies Online: Communication*. Oxford University Press. Available at: http://www.oxfordbibliographies.com/ (Accessed: 12 January 2021).

Margolis, J. and Fisher, A. (2002) *Unlocking the Clubhouse: Women in computing*. Cambridge, Mass: MIT Press.

Marino, M.C. (2009) 'Disrupting Heteronormative Codes: When Cylons in Slash Goggles Ogle Anna Kournikova', in *Proceedings of the Digital Arts and Culture Conference 2009*. *Digital Arts and Culture Conference*, University of California, Irvine. Available at: https://escholarship.org/uc/item/09q9m0kn (Accessed: 31 May 2025).

Marino, M.C. (2020) 'Introduction', in *Critical Code Studies: initial methods*. Cambridge, Massachusetts: The MIT Press (Software studies), pp. 1–36.

McGregor, S., Purver, M. and Wiggins, G. (2016) 'Process based evaluation of computer generated poetry', in *Proceedings of the INLG 2016 Workshop on Computational Creativity in Natural Language Generation*, pp. 51–60. doi: 10.18653/v1/W16-5508.

Miall, D. S. and Kuiken, D. (1999) 'What is literariness? Three components of literary reading', *Discourse Processes*, 28(2), pp. 121–138. doi: 10.1080/01638539909545076.

Montfort, N. (2014) *New Novel Machines: Nanowatt and World Clock*. technical TROPE-14-01. Cambridge, Massachusetts: Massachusetts Institute of Technology, pp. 1-8. Available at: https://dspace.mit.edu/handle/1721.1/92422 (Accessed: 25 April 2019).

Norman, G. (2010) 'Likert scales, levels of measurement and the "laws" of statistics', *Advances in Health Sciences Education*, 15(5), pp. 625–632. Available at: https://doi.org/10.1007/s10459-010-9222-y.

Pihur, V., Datta, Susmita and Datta, Somnath (2009) 'RankAggreg, an R package for weighted rank aggregation', *BMC Bioinformatics*, 10(1), pp. 62-72. Available at: https://doi.org/10.1186/1471-2105-10-62.

Pope, J. (2010) 'Where Do We Go From Here? Readers' Responses to Interactive Fiction: Narrative Structures, Reading Pleasure and the Impact of Interface Design', *Convergence: The International Journal of Research into New Media Technologies*, 16(1), pp. 75–94. Available at: https://doi.org/10.1177/1354856509348774.

Pressman, J. (2020) *Bookishness: loving books in a digital age*. New York: Columbia University Press (Literature now).

Rettberg, S. (2019) 'Combinatory Poetics', in *Electronic Literature*. Polity Press. pp. 20–53.

Riffe, D. *et al.* (2019) *Analyzing Media Messages: Using Quantitative Content Analysis in Research*. 4th edn. New York: Routledge. Available at: https://doi.org/10.4324/9780429464287.

Schoonenboom, J. and Johnson, R.B. (2017) 'How to Construct a Mixed Methods Research Design', *Kolner Zeitschrift Fur Soziologie Und Sozialpsychologie*, 69(Suppl 2), pp. 107–131. Available at: https://doi.org/10.1007/s11577-017-0454-1.

Stewart, G. (2010) 'The paratexts of Inanimate Alice: Thresholds, genre expectations and status', *Convergence: The International Journal of Research into New Media Technologies*, 16(1), pp. 57–74. doi: 10.1177/1354856509347709.

Švelch, J. (2017) *Paratexts to non-linear media texts: Paratextuality in video game culture*. Thesis. Charles University. Available at: https://www.researchgate.net/publication/319932075_Paratexts_to_Non-Linear_Media_Texts_Paratextuality_in_Video_Game_Culture.

May, T. (2006) 'Archaeological histories of who we are', in *The Philosophy of Foucault*. Oxford, UK: Taylor & Francis Group, pp. 36–71. Available at: http://ebookcentral.proquest.com/lib/soton-ebooks/detail.action?docID=1886853.

Tkacz, L. (2019). *Creative text generation: A NaNoGenMo 2018 study* (Unpublished master's thesis). University of Southampton, UK.

Tkacz, L. (2023). *What Has AI Ever Done For Us? A Dip Into Generative Literature*. Hampshire Writer's Society event, November 14 2023, Winchester, UK.

Van Dijk, Y. (2014). The margins of bookishness: Paratexts in digital literature. In Desrochers, N., & Apollon, D. (Eds.). *Examining paratextual theory and its applications in digital culture*. (pp. 24-45). IGI Global.

Van Stegeren, J. and Theune, M. (2019) 'Narrative generation in the wild: Methods from NaNoGenMo', in *ACL 2019 Proceedings of the Second Storytelling Workshop*. *StoryNLP Workshop, ACL 2019*, Florence, Italy, pp. 65–74. doi: 10.18653/v1/W19-3407.

Winter, J.F.C. de and Dodou, D. (2010) 'Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon (Addendum added October 2012)', *Practical Assessment, Research, and Evaluation*, 15(1). Pp. 1-16. Available at: https://doi.org/10.7275/bj1p-ts64.

Ytre-Arne, B. and Moe, H. (2021) 'Folk theories of algorithms: Understanding digital irritation', *Media, Culture & Society*, 43(5), pp. 807–824. Available at: https://doi.org/10.1177/0163443720972314.

# Appendix A

Table A1: Transtextual Relationship Terms and Their Definitions

| Term | Definition |
|---|---|
| Intertextuality | A relationship of copresence between two or more texts. For example: quoting, reference, plagiarism, and allusion (Genette, 1997a, pp. 1-2). |
| Paratextuality | The relationship of the text and print elements such as the title, preface, foreword, notes, blurb, book cover, official and unofficial commentary. "…the 'foretext' of various rough drafts, outlines, and projects of a work can also function as a paratext" (Genette, 1997a, p. 3). |
| Metatextuality | The relationship between the text and commentary which can be critical and can be without citation - "It unites a given text to another, of which it speaks without necessarily citing it (without summoning it), in fact sometimes even without naming it" (Genette, 1997a, p. 4). |
| Hypertextuality | A hypotext is an earlier, pre-existing text from which a later text, the hypertext, is explicitly or implicitly derived, or 'grafted', although not as commentary. Hypertextuality therefore describes the relationship between the hypo- and hypertext (Genette, 1997a, p. 5). |
| Architextuality | Where a single text emerges from general categories such as literary genre, discourse types, modes of delivery, and form (novel, poem, etc.). The can be explicitly or implicitly signaled, but it is not typically stated in the text (Genette, 1997a, pp. 1, 4). |

Table A2: Survey Questions Grouped by Type

| Survey | Subquestion | Response | Source | Measure |
|---|---|---|---|---|
| B1, B2, B3, S1, S2,<br>S3 | A1. The project is literary.<br>A7. The project is thought-provoking.<br>A8. The project has multiple layers of meaning. | Likert | A1, A8 Koolen et al. (2020)<br>A7, my own question | Literary value |
| B1, B2, B3, S1, S2,<br>S3 | A2. The project is creative.<br>A10. The project is surprising.<br>A11. The project is meaningful.<br>A4. The project is unimaginative. | Likert | A2, McGregor et al. (2016)<br>A10, Lamb et al. (2018), Miall and Kuiken (1999)<br>A11, McGregor et al. (2016)<br>A4, my own question | Creative value<br>A4 is control question to be compared to A2 response |
| | A3. The project is technical.<br>A12. The project is understandable.<br>A13. The project is cleverly designed. | Likert | A3, A13, my own question<br>A12, Henrickson (2019), Lamb et al. (2018) | Technical value |
| B1, B2, | A5. The project is interesting. | Likert | A5, | Overall value |

| | | | | |
|---|---|---|---|---|
| B3, S1, S2,<br>S3 | A6. The project is enjoyable.<br><br>A9. The project is boring. | | Henrickson (2019)<br><br>A6, Busselle and Bilandzic (2009), Kuijpers (2014)<br><br>A9, my own question | A9 is control question to be compared to A5 response |
| B1, B2, B3, S1, S2,<br>S3 | A14. I am a literary reader.<br><br>A15. I am a creative person.<br><br>A16. I am a technical person.<br><br>A17. I am an unimaginative person. | Likert | A14, Koolen et al. (2020)<br><br>A15, A16, A17, my own questions mirroring A14. | Self-perception<br><br>A17 is control question to be compared to A15 response |
| B1, B2,<br>B3 | A18. When I first read the project, I realized right away that it was inspired by a monologue made by Molly Bloom, who is a character in James Joyce's book Ulysses.<br><br>A19. I have read James Joyce's book Ulysses, or I have engaged with an adaptation (for example, I have seen a film based on the book).<br><br>A20. I am familiar with James Joyce's book Ulysses (for example, I know some details about the plot or the book's cultural status). | Likert | A18, A19, A20, my own questions as motivated by links to literary reference, Miall and Kuiken (1999) | Recognition of and familiarity with the project's hypotext |

| | | | | |
|---|---|---|---|---|
| S1, S2,<br>S3 | A18. When I first read the project, I realized right away that it contained the first line from Jane Austen's book *Pride and Prejudice*.<br><br>A19. I have read Jane Austen's book Pride and Prejudice, or I have engaged with an adaptation (for example, I have seen a film based on the book).<br><br>A20. I am familiar with Jane Austen's book Pride and Prejudice (for example, I know some details about the plot or the book's cultural status). | Likert | A18, A19, A20, my own questions as motivated by links to literary reference, Miall and Kuiken (1999) | Recognition of and familiarity with the project's hypotext |
| B2, B3, S2,<br>S3 | A21. Overall, I feel like I understood how the code (piece B) works.<br><br>A22. I feel like I understood what the creator of the project intended the code (piece B) to do.<br><br>A23. I feel like I understood the purpose of the code (piece B) within the project.<br><br>A24. I feel like I understood the purpose of the code repository (piece C) within the project.<br><br>A25. I feel like I understood the purpose of the input (piece G) within the project. | Likert | A21, A22, A23, A24, A25, my own questions | Self-perceived comprehension |
| B1, B2, | Q6. What did you think of the | Free | Q6, my own | Overall value |

| | | | | |
|---|---|---|---|---|
| B3, S1, S2,<br><br>S3 | project? Please write a short review explaining your personal opinion. | Response | question | |
| B1, B2, B3, S1, S2,<br><br>S3 | Q71. The names of the pieces you read earlier are shown below. The survey would like to know your personal opinion about which of these were the most important for helping to **understand** the project. Please rank the pieces in order of importance.<br><br>Q72. The names of the pieces you read earlier are shown below. The survey would like to know your personal opinion about which of these were the most important for helping to make the project **interesting**. Please rank the pieces in order of importance. | Ranking | Q71, Q72, my own question | Piece ranking |
| B1, B2, B3, S1, S2,<br><br>S3 | A26. Before taking this survey, I had read works which are part of the NaNoGenMo challenge.<br><br>A27. Before taking this survey, I had read works of computer generated poetry.<br><br>A28. Before taking this survey, I had read works of Electronic Literature.<br><br>A29. Before taking this survey, I had read works of Experimental Literature.<br><br>A30. Before taking this survey, I had read works | Likert | A26, A27, A28, A29, A30, A31, my own questions based on forms which have links to creative text generation | Reading background |

| | | | | |
|---|---|---|---|---|
| | which were made using the AI Dungeon text adventure generator.<br><br>A31. Before taking this survey, I had read works which were made using the Oulipo's constrained writing techniques. | | | |
| B1, B2, B3, S1, S2,<br><br>S3 | A32. Before taking this survey, I had created or studied computer generated poetry works.<br><br>A33. Before taking this survey, I had created or studied Experimental Literature works.<br><br>A34. Before taking this survey, I had created or studied NaNoGenMo works.<br><br>A35. Before taking this survey, I had created or studied works made with the AI Dungeon text adventure generator.<br><br>A36. Before taking this survey, I had created or studied works which use the Oulipo's constrained writing techniques.<br><br>A37. Before taking this survey, I had created or studied Electronic Literature works. | Likert | A32, A33, A34, A35, A36, A37, my own questions based on forms which have links to creative text generation | work with a form |
| B1, B2, B3, S1, S2, | A38. Before taking this survey, I was aware of the GitHub platform. | Likert | A38, A39, A40, A41, A42, my own questions | Awareness of a technology or platform |

| | | | | |
|---|---|---|---|---|
| S3 | A39. Before taking this survey, I was aware of the Python programming language.<br><br>A40. Before taking this survey, I was aware of Machine Learning.<br><br>A41. Before taking this survey, I was aware of Neural Networks.<br><br>A42. Before taking this survey, I was aware of the GPT2 or the GPT3 text generation model, which was made by the company OpenAI. | | based on technology or platforms used by both projects | |
| B1, B2, B3, S1, S2,<br><br>S3 | A43. Before taking this survey, I had used the GitHub platform.<br><br>A44. Before taking this survey, I had used the Python programming language.<br><br>A45. Before taking this survey, I had used Machine Learning.<br><br>A46.Before taking this survey, I had used Neural Networks.<br><br>A47. Before taking this survey, I had used the GPT2 or the GPT3 text generation model, which was made by the company OpenAI. | Likert | A43, A44, A45, A46, A47, my own questions based on technology or platforms used by both projects | Use of a technology or platform |

Table A3: Changes Made to Pieces

| NaNoGenMo Project | Piece | Changes Made |
|---|---|---|
| Shane (2019) | A generated text | Used the first 2 sections beginning with “It is a truth universally known’ from Shane’s victorian .txt file, and created a PDF. This is a common format in NaNoGenMo and makes Shane’s piece A similar in look to Bhatnagar’s piece A. Not all of the second section is included because of wordcount constraints. A title page reading ‘Victorian’ and Page numbers were also created.<br><br>The sample PDF has a total of 1064 words, excluding the title. |
| Shane (2019) | B code | Inserted some Javascript into the HTML file which disables the ability to click and follow links.<br><br>I emailed Shane asking for a copy of their code, and received the reply “Here's the framework I used! (Note the colab notebooks that let you do this for free without needing to program). https://GitHub.com/minimaxir/aitextgen”. The google colab notebooks look very similar to the less verbose jupyter notebooks which are in the aitextgen repo, by which I mean that the latter part of the notebook appears to use the same code and the same shakespeare plays input data. As Bhatanagar 2015 also uses a jupyter notebook, I have decided to use and alter the aitextgen training_hello_world.ipynb as the code example for my study website.<br><br>I forked the aitextgen repo and saved the page. I then replaced references to my repo with janelleshane and with novel-first-lines-dataset. I replaced jupyter notebook input and output with the input file (piece G) and with 'it is a truth universally acknowledged' text samples from the generated text (piece A) inserted into the notebook as example output. Also removed date, the topmost occurance of the code author's name, and altered heading to refer to 'training example' rather than to 'Hello World. I kept the code author’s name (Max Woolf) and the copyright type at the bottom of the page, but I removed the longer |

| | | |
|---|---|---|
| | | copyright statement to shorten the survey participant's reading time. |
| Shane (2019) | C repository | Inserted some Javascript into the HTML file which disables the ability to click and follow links.<br><br>Added training_example.ipynb to the repository, with a date of '16 months ago'.<br><br>Removed description and links to further readings on Shane's blog/website aiweirdness.com, in order to discourage the survey participants from trying to follow the links to read more and potentially forget about the study. For the same reason, I removed a link to a Google Forms which collected data for the input (piece G) - the link location was not described on the repo page anyway.<br><br>Removed all the text within the 2017: syll rnn heading because this page is very long and needs to be cut, plus I judge it to be less relevant to understanding the project than other sections of the text.<br><br>Removed a more detailed description of GPT2 because it is repeated in the creator's page (piece D).<br><br>Removed all text within the 'ancient', 'ponies', and 'potter' because the focus of this study is on 'victorian'. The first 4 output examples for 'victorian' were kept, and the rest removed in order to shorten the length of time it will take to read the webpage.<br><br>Reordered some text, and moved in some text from the creator's progress page to try to cut down on repetition in both, and separate different types of information a bit.<br><br>Added a file named victorian.pdf to the repo to reinforce that the PDF version of piece A that I created is part of the project. Adding a PDF of the generated text to the repo also mirror's Bhatnagar 2015's repo. |
| Shane (2019) | D creator's | Inserted some Javascript into the HTML file which disables |

| | | |
|---|---|---|
| | page | the ability to click and follow links.<br><br>Removed similar things and repetitive information which was already written or removed in the repo's text.<br><br>Added output examples and 2 extra paragraphs of explanatory text from Shane's writeup about the project on aiweirdness.com. This was done to describe the project from the creator's perspective further, and to add some information which was not already in the repo (piece C).<br><br>Removed the headings and examples from 'ancients', 'ponies', 'potter', and 'victorian' because the latter is already in the repo (piece C).<br><br>Removed text which further elaborated on GPT2 because that is already discussed in the repo. I am trying to move some of the more technical topics to the repo (because that is what Bhatnagar 2015 does - although Shane has much more text and description in both the repo and the creator's progress page).<br><br>I logged into GitHub and added my own comment to this issue page in order to balance the amount of comments in both NaNoGenMo projects: "I really like 'victorian' - the absurdity mixed with patches of older style language work well together. It reminds me of reading Victorian era novels where I've lost track of the rambling text but I can still pick out comical descriptions of things." My comment echoes what Shane has mentioned already about style and absurdity either on GitHub or on aiweirdness.com. In piece D I altered my comment to look like it was posted on the same day as the other posts and notifications on the page, and removed my avatar photo from the 'post a comment' section. My GitHub username is Dx9240, so survey participants are less likely to make the connection that the researcher is also the poster of that comment. |
| Shane (2019) | E nanogenmo page | Inserted some Javascript into the HTML file which disables the ability to click and follow links. |
| Shane (2019) | F media | Inserted some Javascript into the HTML file which disables the ability to click and follow links. |

| | | |
|---|---|---|
| | | Removed the '25 min read' note at the top of the page, because this sample no longer takes that long to read.<br><br>Removed descriptions of all NaNoGenMo works except for Shane 2019. Removed the 'ancients' generated text sample.<br><br>Total word count is about 301 words.<br><br>Removed ads and titles of other articles at the bottom of the webpage. This mirrors what I did for Bhatnagar's media piece as well. Also removed a link with a follow button showing 'More from Greg Kennedy' from the bottom of the page. |
| Shane (2019) | G input | Preserving the first and final lines, I skipped around the document to select and delete sections of lines, in order to cut down the input to about 1000 words. |
| Bhatnagar (2015) | A generated text | Shorted the PDF to create a sample text from pages 1-3, and the final page 164. The final page of the sample contains the last 2 lines from page 163. This was done in order to fill the page so that the Bhatnagar 2015 generated text sample looks similar and has a similar amount of text to the Shane 2019 generated text sample.<br><br>The sample PDF has a total of 1043 words, excluding the title.<br><br>Removed a link to "source code at https://GitHub.com/moonmilk/nanogenmo2015" so that the survey participants do not go to the repository during the study.<br><br>Added page numbers to the PDF to show that pages have been cut from the document.<br><br>Not all of the lines from the text are on the same page they were originally on, due to formatting processing when cutting pages and deleting source code messages. However each individual line is preserved just as one would preserve the lines of a poem. |

| | | |
|---|---|---|
| Bhatnagar (2015) | B code | Inserted some Javascript into the HTML file which disables the ability to click and follow links.<br><br>Shortened some of the output in the jupyter notebook to reduce the survey participant's reading time. |
| Bhatnagar (2015) | C repository | Inserted some Javascript into the HTML file which disables the ability to click and follow links. |
| Bhatnagar (2015) | D creator's page | Inserted some Javascript into the HTML file which disables the ability to click and follow links. |
| Bhatnagar (2015) | E nanogenmo page | Inserted some Javascript into the HTML file which disables the ability to click and follow links. |
| Bhatnagar (2015) | F media | Removed distracting advertisement, registration box, and undisplayed images which were between paragraphs. Text referring to the 2016 and 2017 NaNoGenMo projects was also removed in order to reduce the amount of time needed to read the page. I then reordered the remaining projects to appear in reverse chronological order so that *Molly's Feed* is the first project to appear in the text.<br><br>Removed the social media posting bar from the left margin of the page because the icons were not able to be disabled.<br><br>Inserted some Javascript into the HTML file which disables the ability to click and follow links.<br><br>I cut some additional text so that this piece is a similar size to Shane 2019's Media piece.<br><br>I also made sure that both media pieces had similar content - describing NaNoGenMo, the quality (gibberish) of most NaNoGenMo entries, making a connection to AI technology, and listing a handful of other generated text or NaNoGenMo examples.<br><br>Total word count is about 296 words. |

| | | |
|---|---|---|
| Bhatnagar (2015) | G input | Preserving the first and final lines, I skipped around the document to select and delete sections of lines, in order to cut down the input to about 1000 words. |

# Appendix B

Submitted as entry #58 to the 2016 NaNoGenMo challenge, *If on a winter's night a library card holder* is a generated novel project by GitHub user robincamille. The text below offers an overview of the project's paratexts as an example of how they may be structured and may function on GitHub. Where possible, each project paratext is paired with a piece name (piece A, B, C, D, E, F, or G) in order to show how *If on a winter's night* maps onto the online survey study's pieces.

The project is hosted on the GitHub platform, where it is tagged as being completed along with other NaNoGenMo 2016 projects[60]. robincamille's GitHub code repositories[61] bear titles which suggest that they have worked on several twitterbots, NaNoGenMo entries, and other language processing projects. Their repository (piece C) for *If on a winter's night*[62] contains a folder of 4 versions of generated text[63] (piece A), 1 input file containing tabular data on New York City public library addresses[64] (piece G), 1 input file containing tabular data on classic book titles and the location of their plain text[65] (piece G), and code files such as makebook.py[66] (piece B) which shows how the hardcoded text templates and input data are combined to create the generated text. In a very literal sense, the input data are hypotexts from which the generated text is algorithmically derived. The repository also contains a readme file[67] which explains that the generated novel is inspired by Italo Calvino's *If on a winter's night a traveler* (1979) - this cements an intertextual relationship between the postmodernist narrative and the generated novel.

Robincamille's readme goes beyond the original technical function of such files, which is to explain the purpose of the code, how to run it, and which data or software tool versions it requires. The readme additionally describes the combinatorial aesthetics of the generated novel by calculating how many unique possible novels can be generated, and includes an image of

---

[60] https://github.com/NaNoGenMo/2016/issues?page=4&q=is%3Aissue+is%3Aopen

[61] https://github.com/robincamille?tab=repositories

[62] https://github.com/robincamille/nanogenmo2016

[63] https://github.com/robincamille/nanogenmo2016/tree/master/outputs

[64] https://github.com/robincamille/nanogenmo2016/blob/master/nyc_public_libraries.tsv

[65] https://raw.githubusercontent.com/robincamille/nanogenmo2016/master/GITenberg_repos_list_2.tsv

[66] https://github.com/robincamille/nanogenmo2016/blob/master/makebook.py

[67] https://github.com/robincamille/nanogenmo2016/blob/master/README.md

one generated novel version produced as a print book. A story synopsis is also included either to paratextually help with contextualizing and reading the generated text, or to generate interest about it.

*If on a winter's night*'s issue page[68] (piece D) shows positive comments and comment 'likes' from other GitHub users. These could be considered as critical paratexts which may influence reader reception and value. Here, robincamille posts an image showing two pages from an open "...hardback book printed from one of the novel's many outputs"[69], which is seen to have a dark red cover, and a dark green and red marbled pattern inside the cover. The book is lying open on a dark brown wood surface which could be a table or desk. The attributes of the book object can be seen as self-referential; its red cover could be in reference to the red book which is described in every generated text version, which is in turn printed on the object's pages. Further, the book object might be lying on a desk in a library, which again may reflexively reference the settings described by the generated text (the book object is not, for example, presented in another setting, such as on grass or on a kitchen tablecloth). If this is recognized by readers, then the self-referential image works paratextually because it enables the connection between *If on a winter's night* to metafiction as a literary device, and to metafictional uses in literature. The paratext therefore expands the breadth of literary devices that the reader is offered. This print book image is not the same as the one seen in the readme file.

NaNoGenMo organizers have marked *If on a winter's night* as a project which has received press coverage in a 2017 article on the *Book Riot* website (piece F), which describes itself as a large editorial book site catering to diverse readers and genres. The article gives a brief overview of interesting NaNoGenMo novels from 2013 - 2017, and selects robincamille's as the pick for 2016 and recognizes it as an "...algorithmic pastiche" of Italo Calvino's *If on a winter's night a traveler*.[70] The article uses robincamille's full name and refers to them as a librarian and 'her'. This suggests that the article's writer may have looked outside of the GitHub platform to learn more about its creator on their personal website.

---

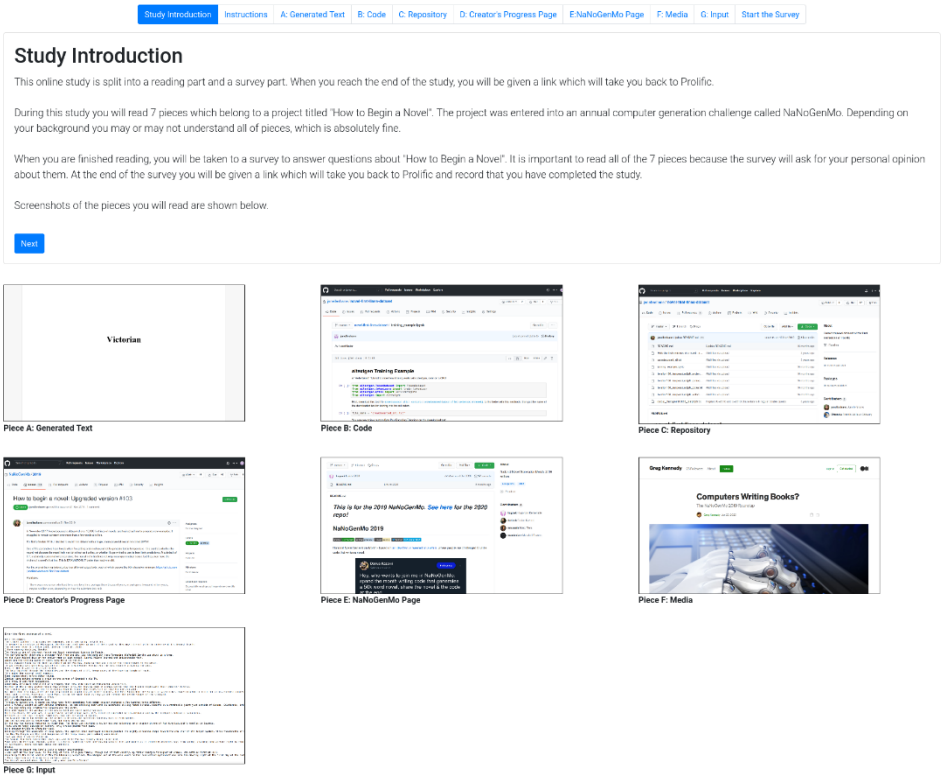
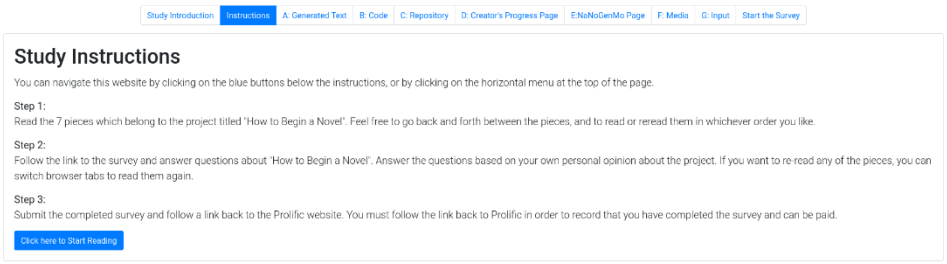
[68] https://github.com/NaNoGenMo/2016/issues/58

[69] https://github.com/NaNoGenMo/2016/issues/58#issuecomment-261819227

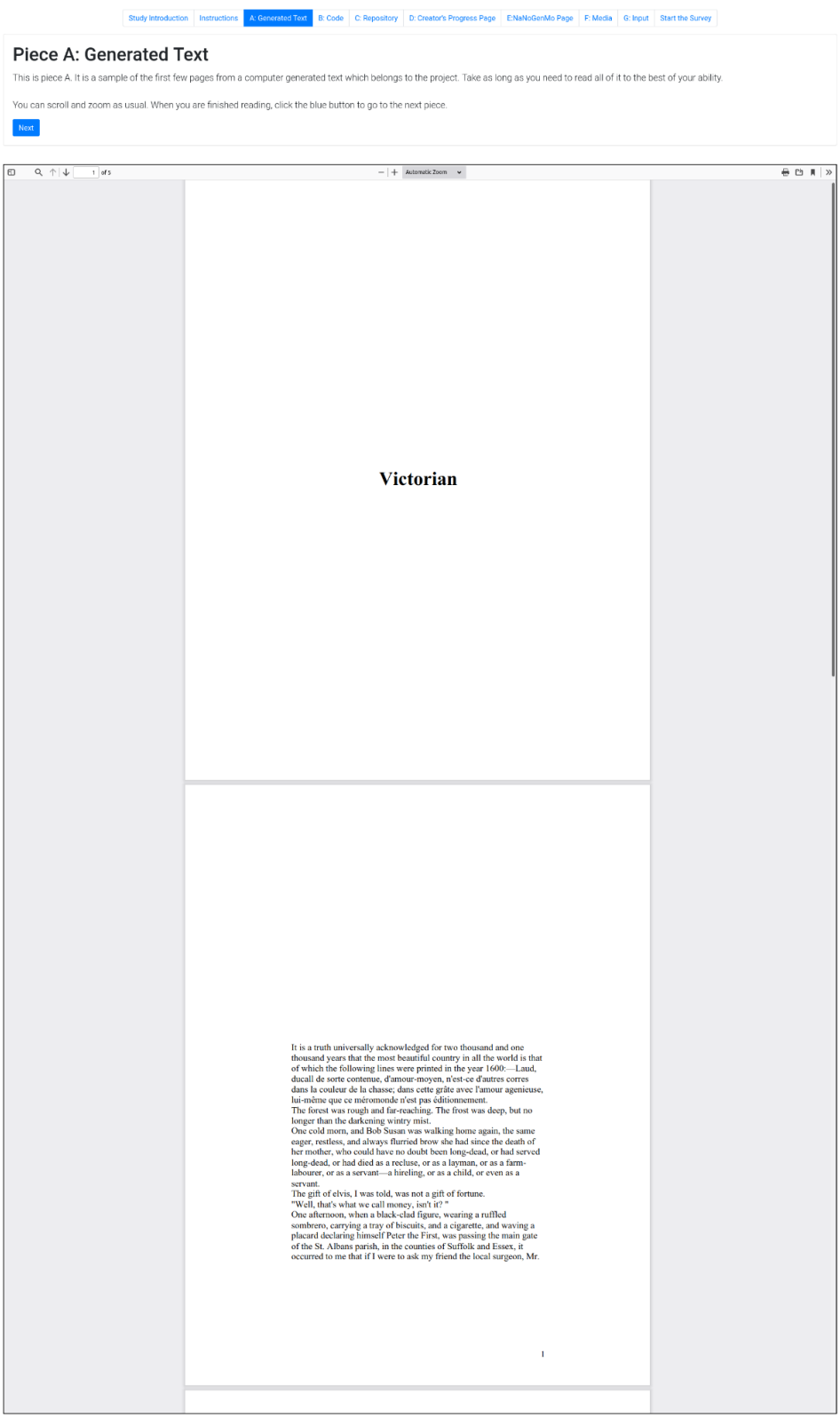
[70] https://bookriot.com/2017/11/27/novels-written-nanogenmo/

# Appendix C

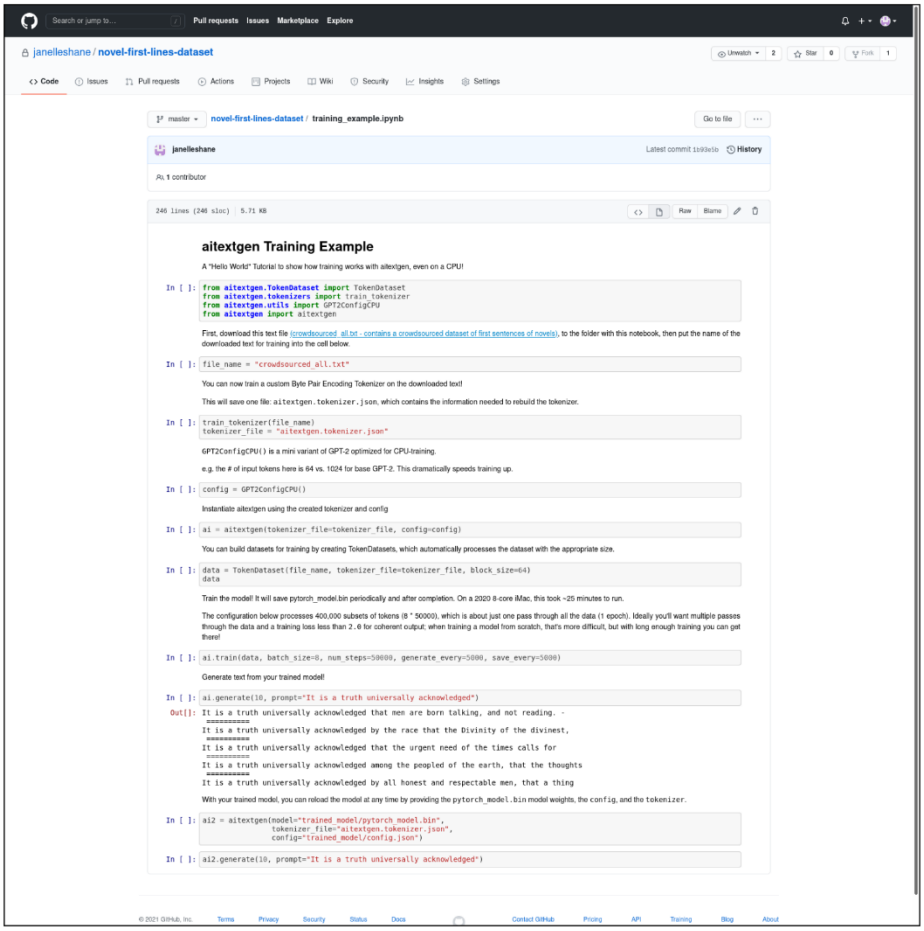
Table C1: S3 Reading Part Website Screenshots

| **Description and Link** | **Screenshot** |
|---|---|
| S3, Study Introduction page |  |
| S3, Study Instructions page |  |

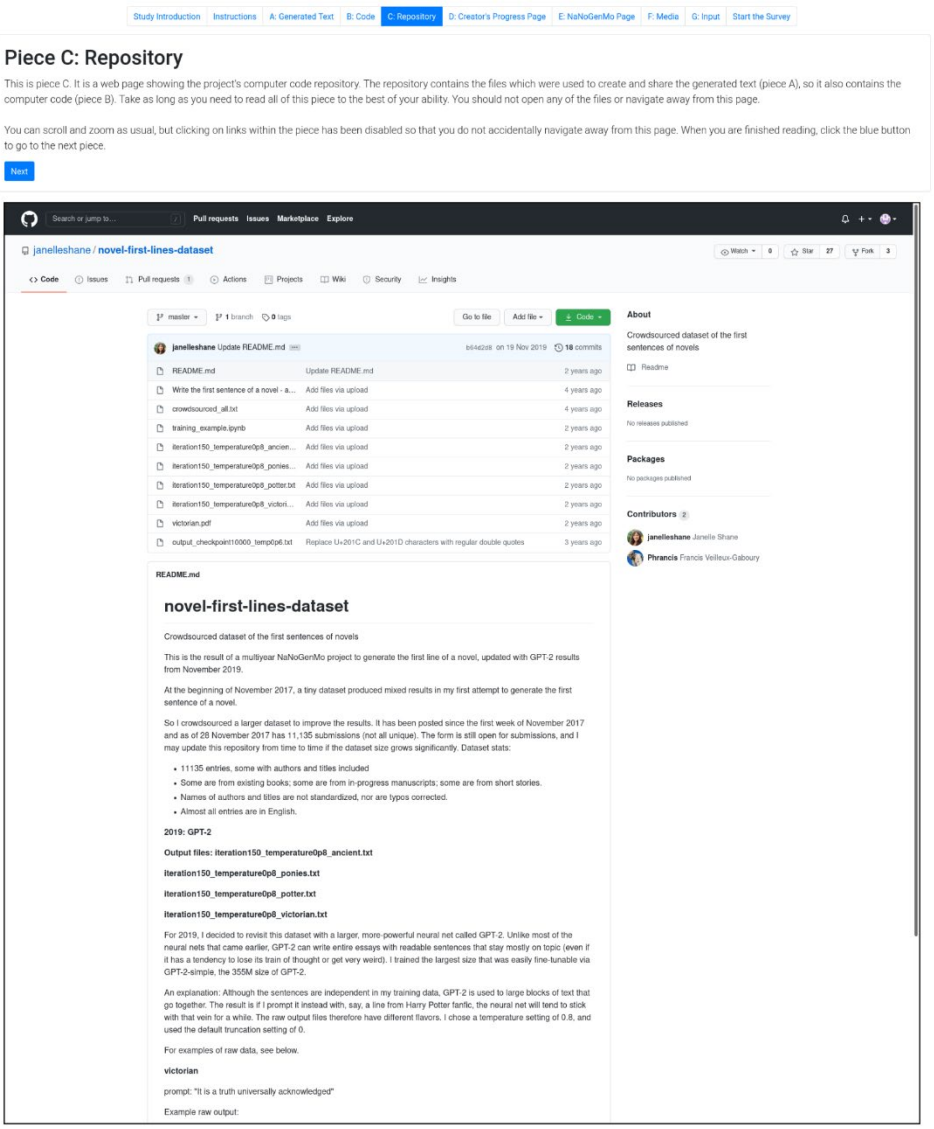
S3, Piece A: Generated Text



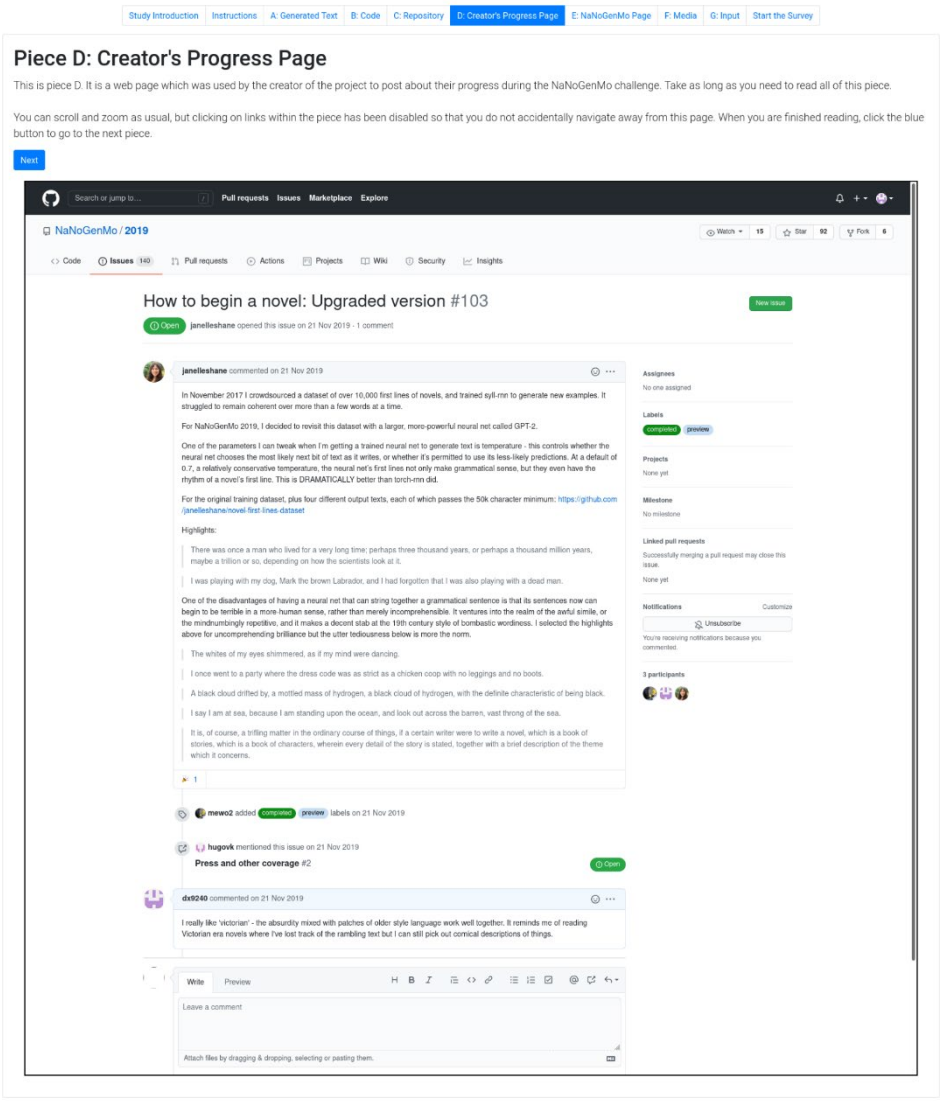
S3, Piece B, Code

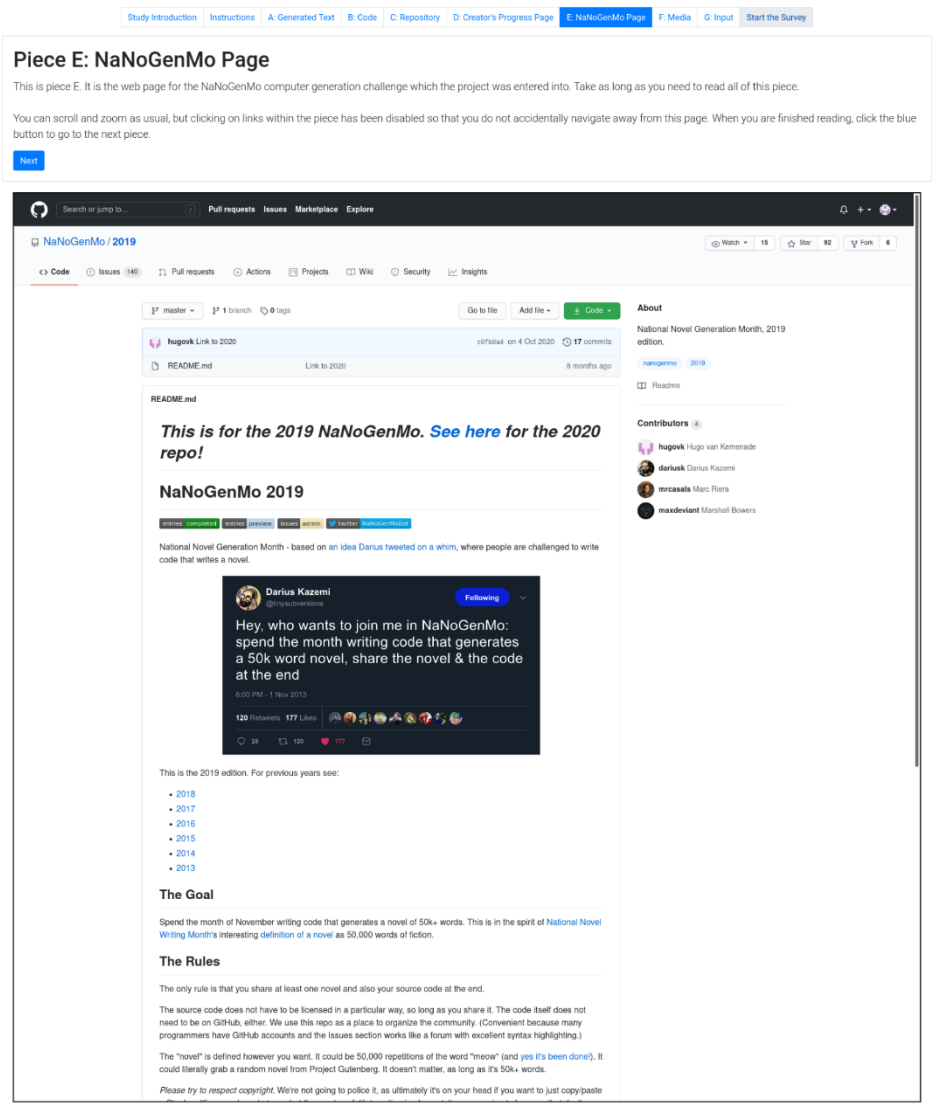


S3, Piece C: Repository



S3, Piece D: Creator's Progress Page



| | |
|---|---|
| S3, Piece E: NaNoGenMo Page |  |

S3, Piece F: Media

Study Introduction | Instructions | A: Generated Text | B: Code | C: Repository | D: Creator's Progress Page | E: NaNoGenMo Page | F: Media | G: Input | Start the Survey

## Piece F: Media

This is piece F. It is a sample of a press article discussing the project. Take as long as you need to read all of this piece.

You can scroll and zoom as usual, but clicking on links within the piece has been disabled so that you do not accidentally navigate away from this page. When you are finished reading, click the blue button to go to the next piece.

Next

Greg Kennedy 26 Followers About Follow Sign in Get started

# Computers Writing Books?

The NaNoGenMo 2019 Roundup

Greg Kennedy Jan 20, 2020


Each November, dozens of programmers, NLP enthusiasts, computer artists, and other folks gather together to participate in National Novel Generation Month. The goal, like NaNoWriMo, is to produce a novel; the difference is that, rather than writing the novel by hand, NaNoGenMo participants write a *computer program* which generates a novel. There are only two real requirements: By the end of November the program must output at least 50,000 words, and the source code must be made available.

Despite its origin as a half-joking tweet, NaNoGenMo has now run for six years, and it remains the highest-profile long-form computer writing venue around. As long-time participants know, NNGM entries aren't typically "readable" as novels. Despite advances in AI, the dream of a computer writing passable novel-length fiction in 2019 remained just a dream.

. . .

63

## Trends

It is impossible to have a discussion of AI in 2019 without mention of OpenAI's GPT-2 language model, a massive data set which released to much fanfare in February. GPT-2 so far has proven to be a flexible tool finding its way into all sorts of projects: complete essay writing prompts via Talk To Transformer, play text adventures with it in AI Dungeon, or use it to replace popular Twitter personalities with uncanny accuracy.

Unsurprisingly, GPT-2 featured heavily in this year's NaNoGenMo entries as well.

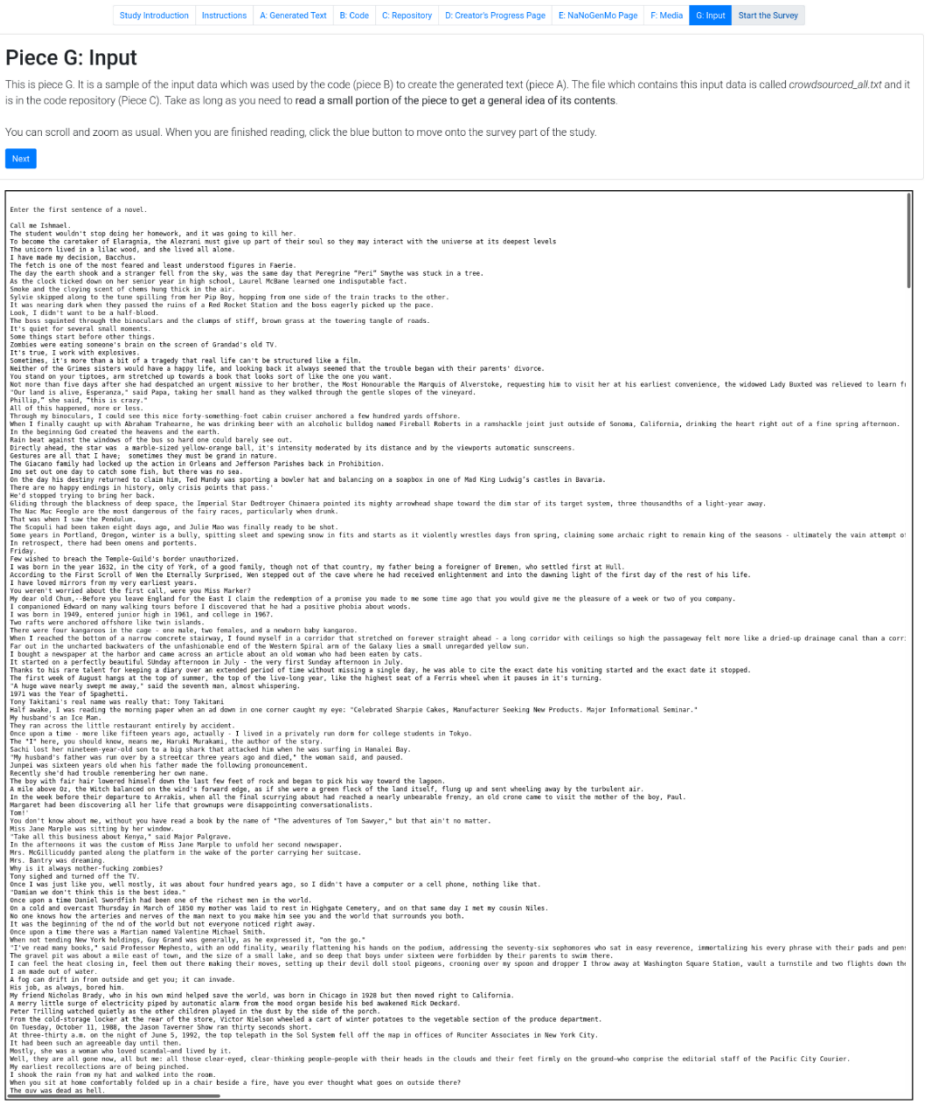
### "How to Begin a Novel" by Janelle Shane

For 2019, "How to Begin a Novel" employs GPT-2 trained on first lines of famous books, with a goal of producing new opening lines to books yet unwritten. This is an update to a previous attempt to generate novel openers using torch-rnn. The results are markedly improved but still often weird. Janelle produced four different sample outputs, each one using a different input data set ("ancient", "ponies", "potter", and "victorian") to create different genre styles.

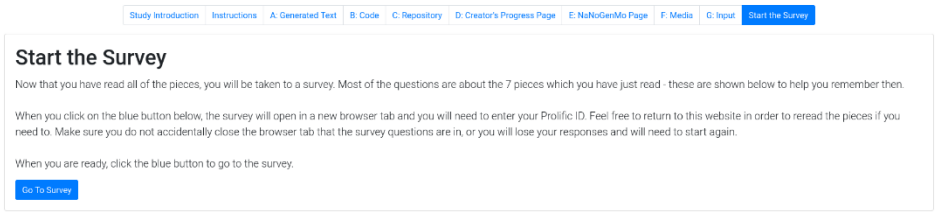
63

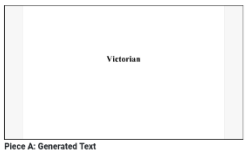
AI Nanogenmo Procedural Generation Programming Generative Art

S3, Piece G: Input



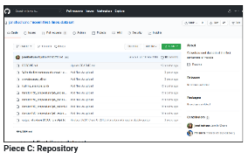
S3, Start the Survey page



Piece A: Generated Text


Piece B: Code


Piece C: Repository

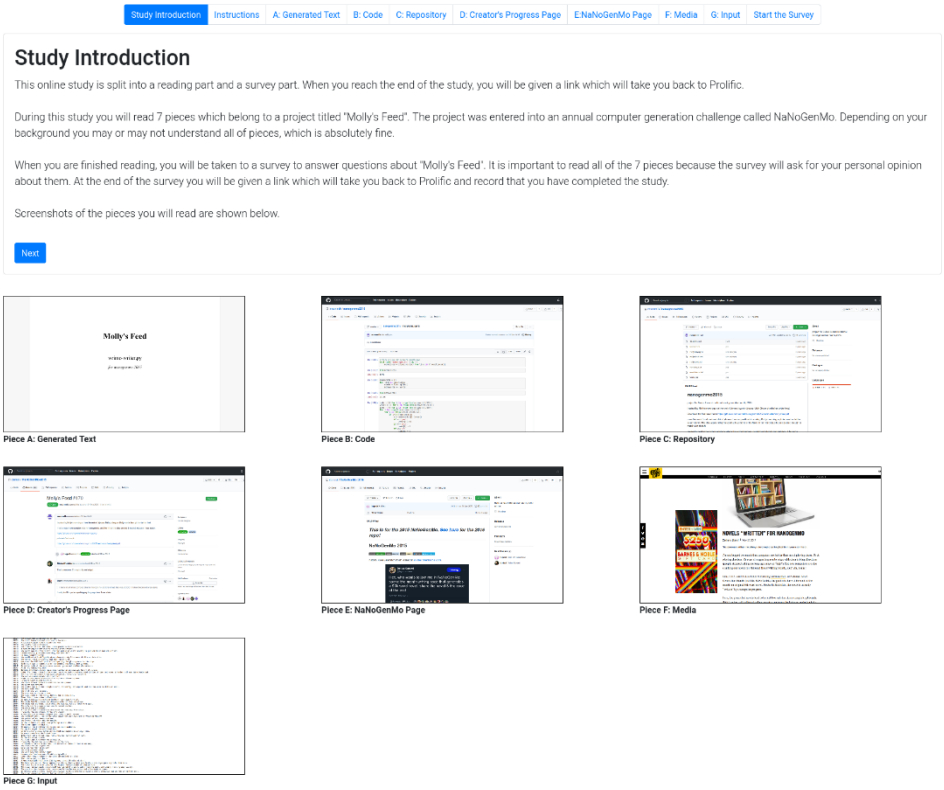
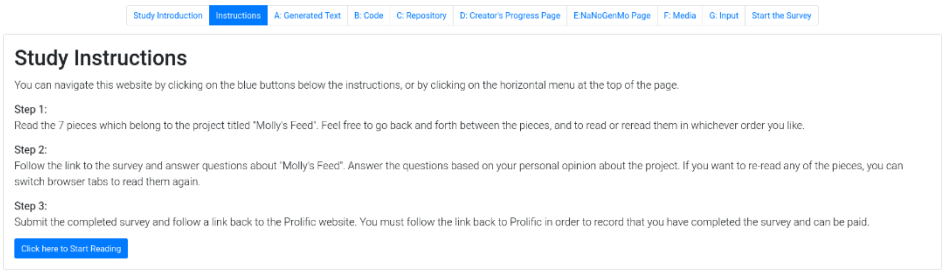

Piece D: Creator's Progress Page


Piece E: NaNoGenMo Page

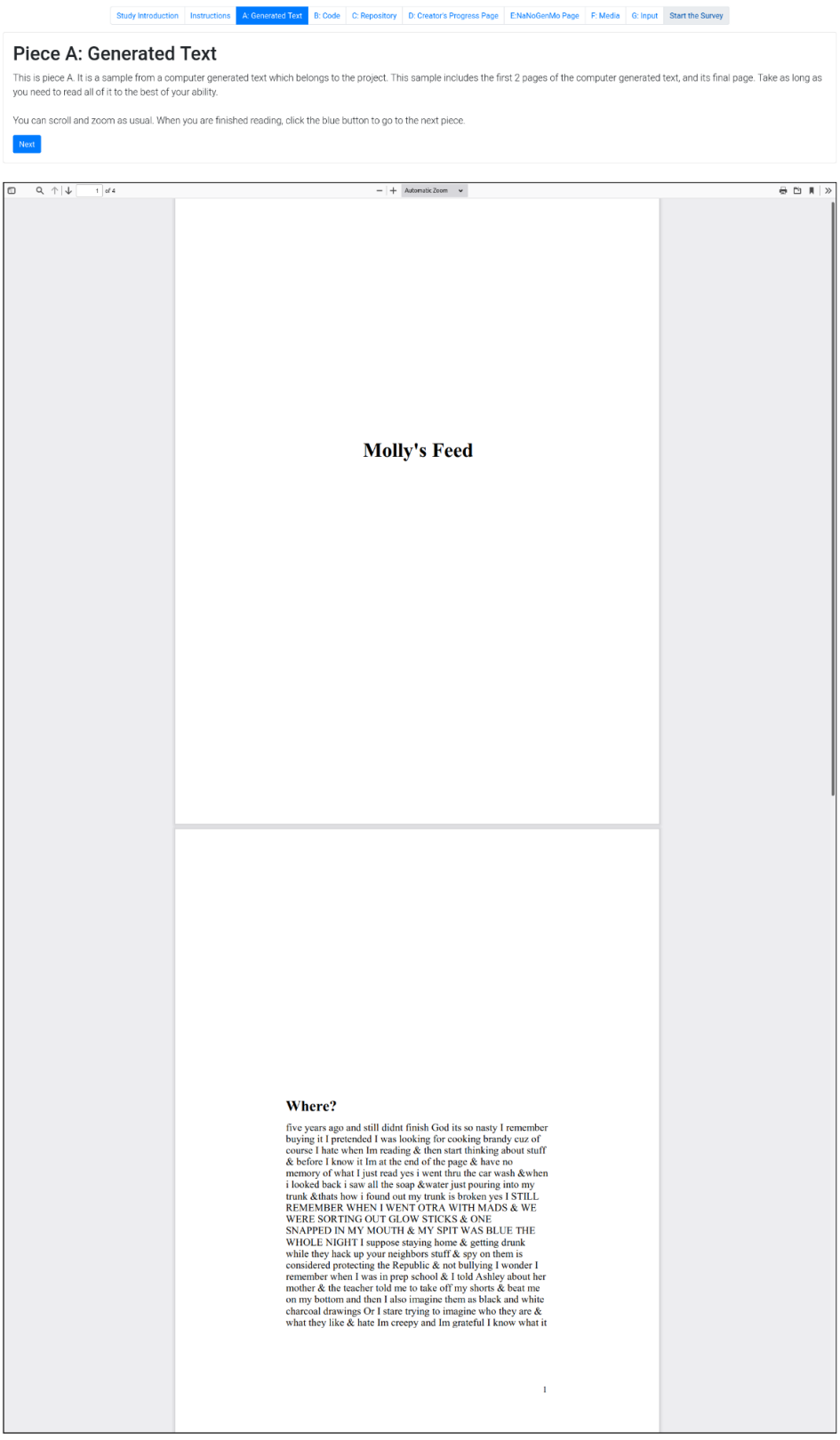
Piece F: Media

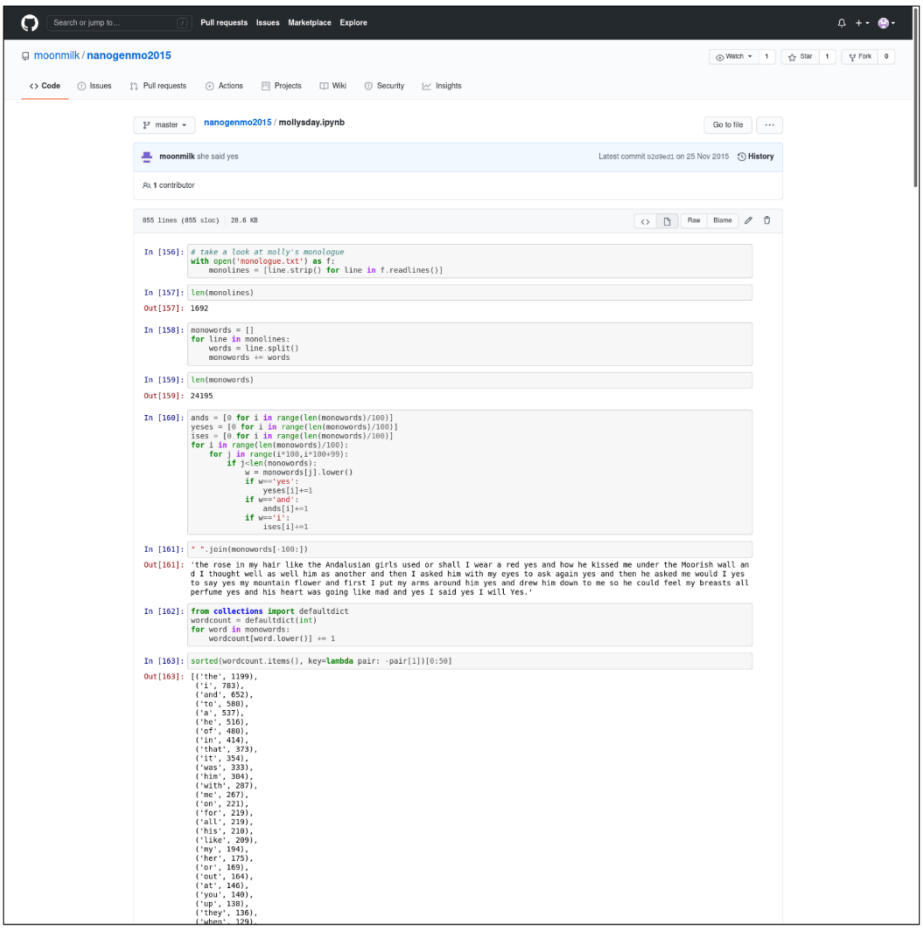

Piece G: Input

Table C2: B3 Reading Part Website Screenshots

| **Description and Link** | **Screenshot** |
|---|---|
| B3, Study Introduction page |  |
| B3, Instructions page |  |

B3, Piece A: Generated Text



| | |
|---|---|
| B3, Piece B: Code | Study Introduction \| Instructions \| A: Generated Text \| B: Code \| C: Repository \| D: Creator's Progress Page \| E:NaNoGenMo Page \| F: Media \| G: Input \| Start the Survey<br>**Piece B: Code**<br>This is piece B. It is a sample of the project's computer code which was used to create the generated text (piece A). Take as long as you need to read all of this piece to the best of your ability.<br>You can scroll and zoom as usual, but clicking on links within the piece has been disabled so that you do not accidentally navigate away from this page. When you are finished reading, click the blue button to go to the next piece.<br>Next<br>  |

Search or jump to... Pull requests Issues Marketplace Explore

moonmilk / nanogenmo2015

Watch 1 Star 1 Fork 0

Code Issues Pull requests Actions Projects Wiki Security Insights

master nanogenmo2015 / mollysday.ipynb Go to file

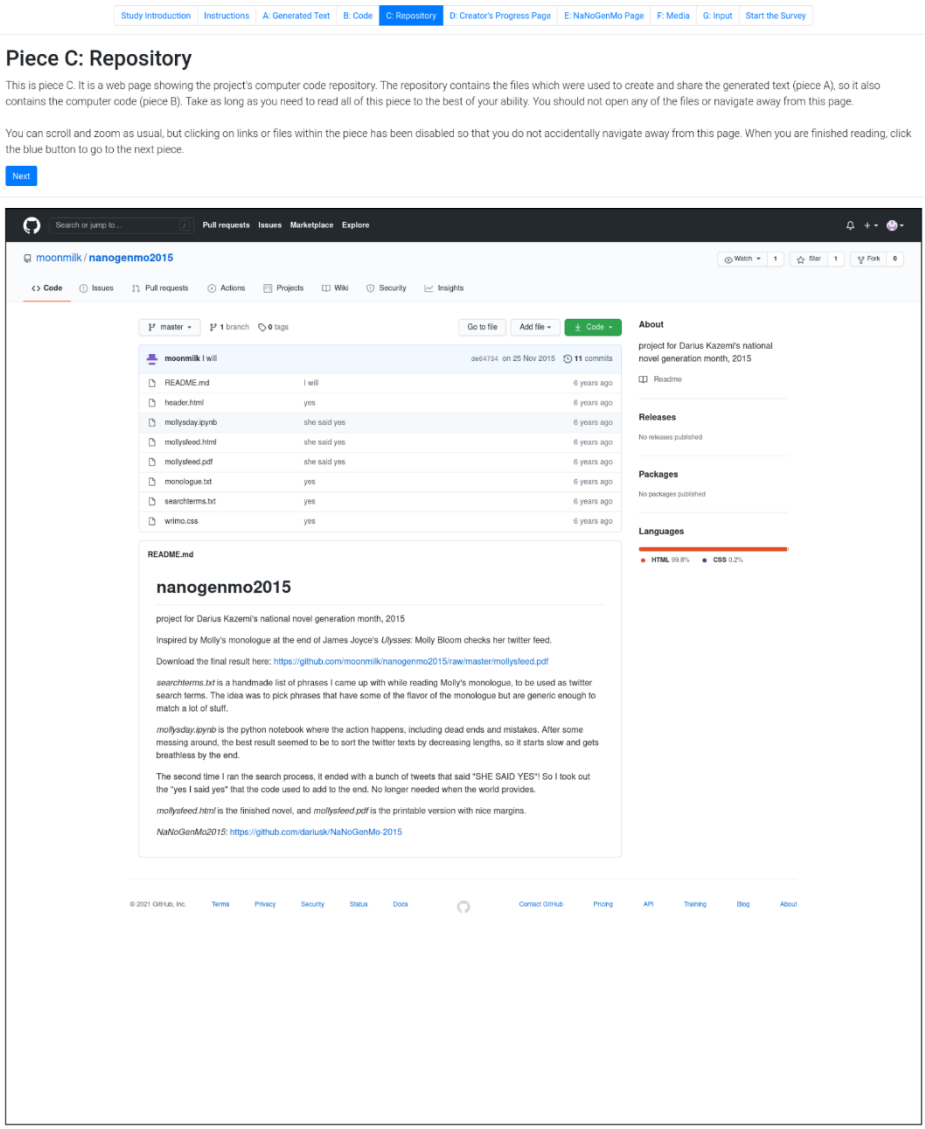
moonmilk she said yes Latest commit b2d9ed1 on 25 Nov 2015 History

1 contributor

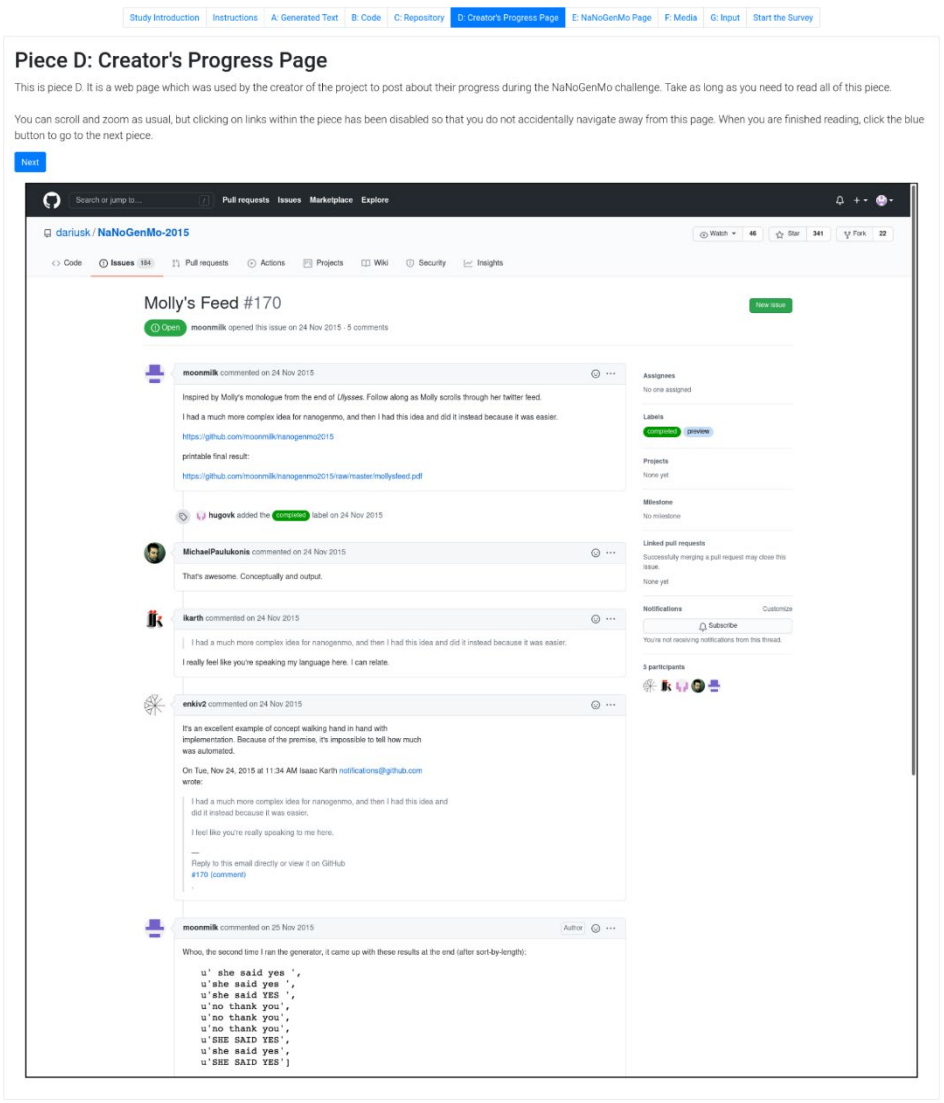
855 lines (855 sloc) 28.6 KB Raw Blame

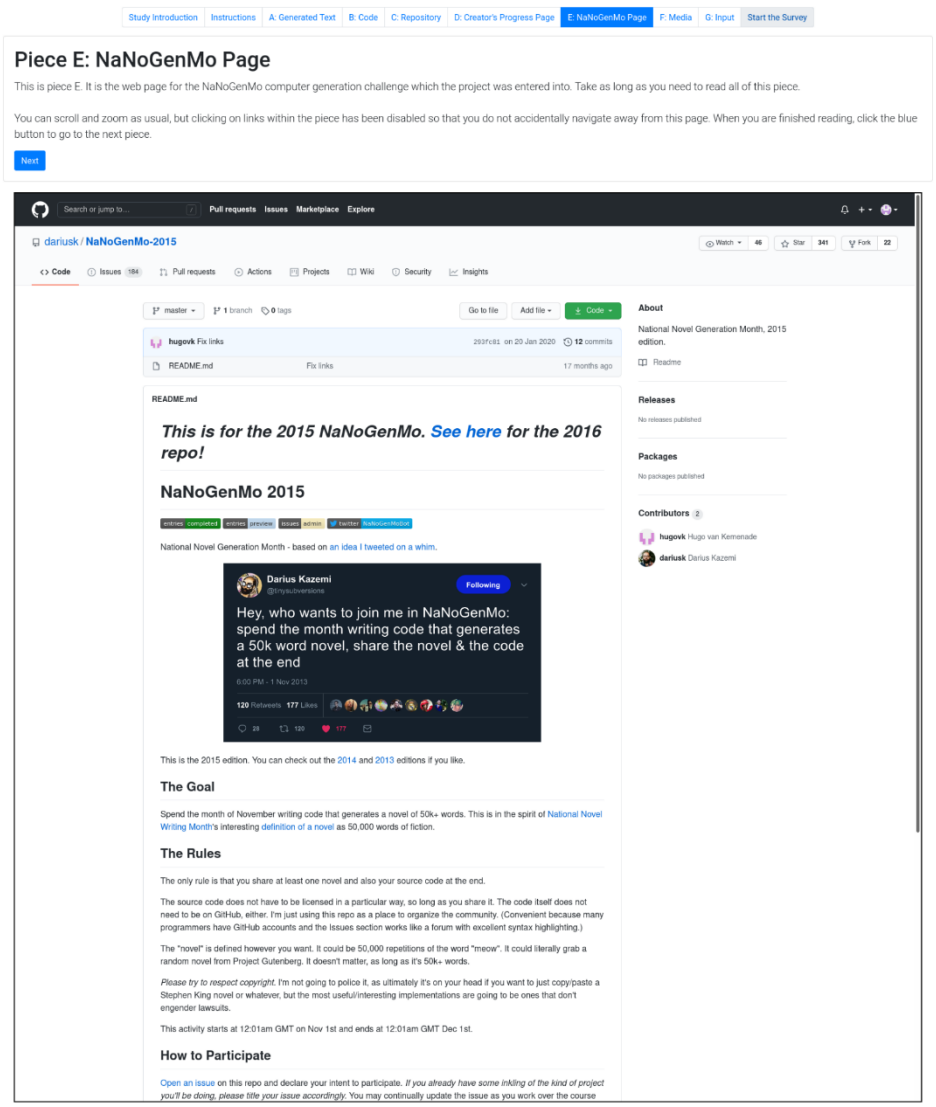
```
In [156]: # take a look at molly's monologue
          with open('monologue.txt') as f:
              monolines = [line.strip() for line in f.readlines()]

In [157]: len(monolines)
Out[157]: 1692

In [158]: monowords = []
          for line in monolines:
              words = line.split()
              monowords += words

In [159]: len(monowords)
Out[159]: 24195

In [160]: ands = [0 for i in range(len(monowords)/100)]
          yeses = [0 for i in range(len(monowords)/100)]
          ises = [0 for i in range(len(monowords)/100)]
          for i in range(len(monowords)/100):
              for j in range(i*100,i*100+99):
                  if j<len(monowords):
                      w = monowords[j].lower()
                      if w=='yes':
                          yeses[i]+=1
                      if w=='and':
                          ands[i]+=1
                      if w=='i':
                          ises[i]+=1

In [161]: " ".join(monowords[-100:])
Out[161]: 'the rose in my hair like the Andalusian girls used or shall I wear a red yes and how he kissed me under the Moorish wall an
          d I thought well as well him as another and then I asked him with my eyes to ask again yes and then he asked me would I yes
          to say yes my mountain flower and first I put my arms around him yes and drew him down to me so he could feel my breasts all
          perfume yes and his heart was going like mad and yes I said yes I will Yes.'

In [162]: from collections import defaultdict
          wordcount = defaultdict(int)
          for word in monowords:
              wordcount[word.lower()] += 1

In [163]: sorted(wordcount.items(), key=lambda pair: -pair[1])[0:50]
Out[163]: [('the', 1199),
           ('i', 783),
           ('and', 652),
           ('to', 580),
           ('a', 537),
           ('he', 516),
           ('of', 480),
           ('in', 414),
           ('that', 373),
           ('it', 354),
           ('was', 333),
           ('him', 304),
           ('with', 287),
           ('me', 267),
           ('on', 221),
           ('for', 219),
           ('all', 219),
           ('his', 210),
           ('like', 209),
           ('my', 194),
           ('her', 175),
           ('or', 169),
           ('out', 164),
           ('at', 146),
           ('you', 140),
           ('up', 138),
           ('they', 136),
           ('when', 129),
```

B3, Piece C: Repository



B3, Piece D: Creator's Progress Page



| | |
|---|---|
| B3, Piece E: NaNoGenMo Page |  |

B3, Piece F: Media

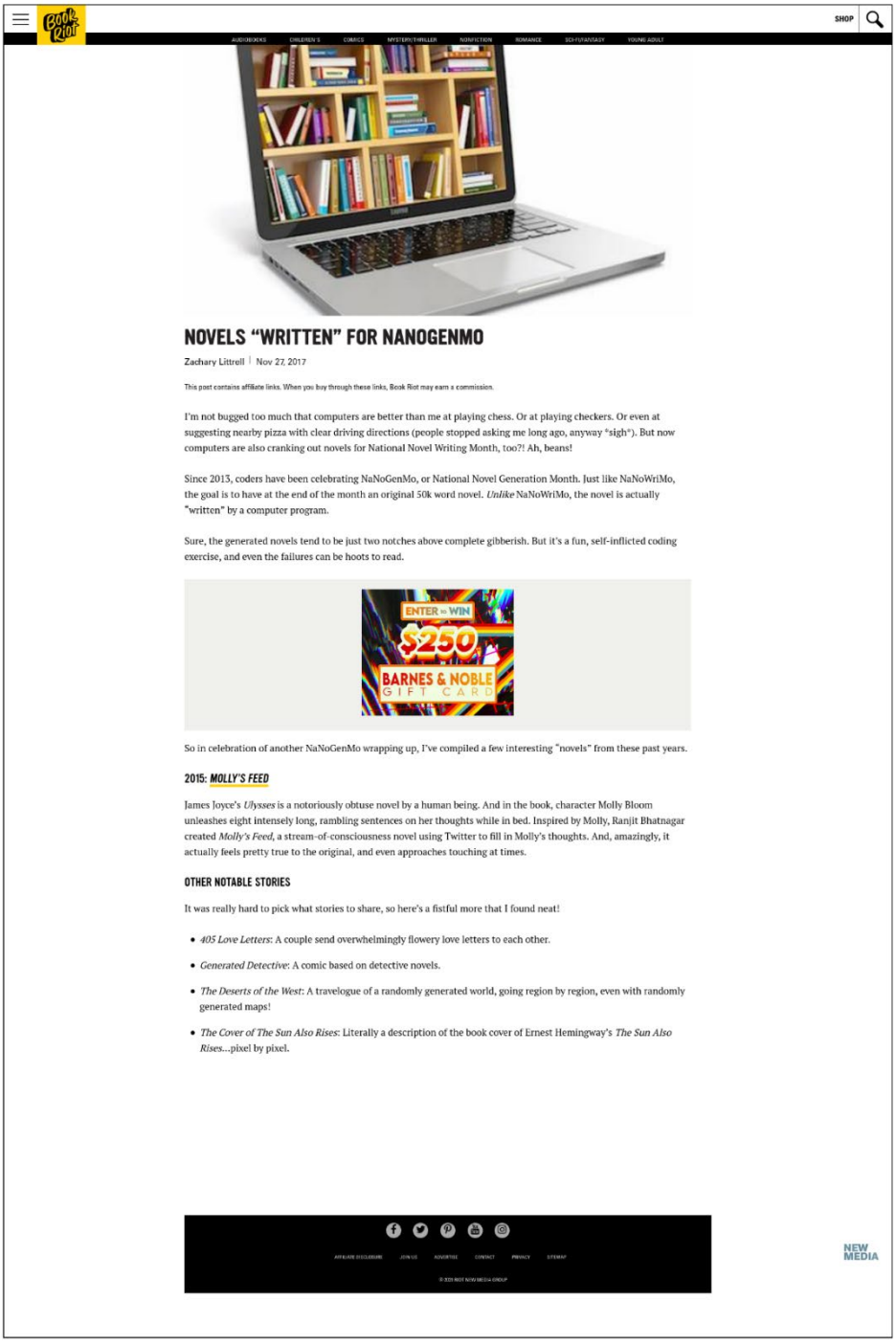
Study Introduction | Instructions | A: Generated Text | B: Code | C: Repository | D: Creator's Progress Page | E: NaNoGenMo Page | F: Media | G: Input | Start the Survey

## Piece F: Media

This is piece F. It is a sample of a press article discussing the project. Take as long as you need to read all of this piece.

You can scroll and zoom as usual, but clicking on links within the piece has been disabled so that you do not accidentally navigate away from this page. When you are finished reading, click the blue button to go to the next piece.

Next



Book Riot

SHOP

AUDIOBOOKS | CHILDREN'S | COMICS | MYSTERY/THRILLER | NONFICTION | ROMANCE | SCI-FI/FANTASY | YOUNG ADULT

### NOVELS "WRITTEN" FOR NANOGENMO

Zachary Littrell | Nov 27, 2017

This post contains affiliate links. When you buy through these links, Book Riot may earn a commission.

I'm not bugged too much that computers are better than me at playing chess. Or at playing checkers. Or even at suggesting nearby pizza with clear driving directions (people stopped asking me long ago, anyway *sigh*). But now computers are also cranking out novels for National Novel Writing Month, too?! Ah, beans!

Since 2013, coders have been celebrating NaNoGenMo, or National Novel Generation Month. Just like NaNoWriMo, the goal is to have at the end of the month an original 50k word novel. *Unlike* NaNoWriMo, the novel is actually "written" by a computer program.

Sure, the generated novels tend to be just two notches above complete gibberish. But it's a fun, self-inflicted coding exercise, and even the failures can be hoots to read.



So in celebration of another NaNoGenMo wrapping up, I've compiled a few interesting "novels" from these past years.

#### 2015: *MOLLY'S FEED*

James Joyce's *Ulysses* is a notoriously obtuse novel by a human being. And in the book, character Molly Bloom unleashes eight intensely long, rambling sentences on her thoughts while in bed. Inspired by Molly, Ranjit Bhatnagar created *Molly's Feed*, a stream-of-consciousness novel using Twitter to fill in Molly's thoughts. And, amazingly, it actually feels pretty true to the original, and even approaches touching at times.

#### OTHER NOTABLE STORIES

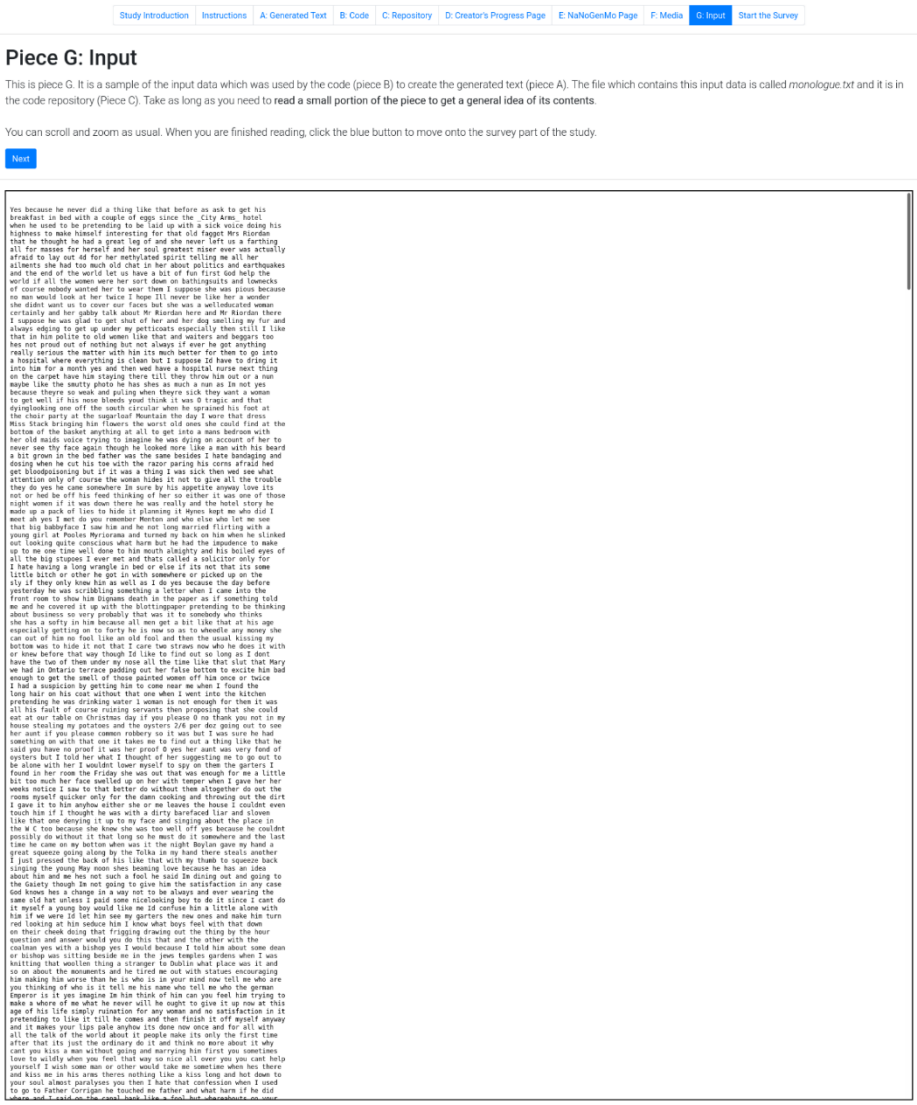
It was really hard to pick what stories to share, so here's a fistful more that I found neat!

- *405 Love Letters*: A couple send overwhelmingly flowery love letters to each other.
- *Generated Detective*: A comic based on detective novels.
- *The Deserts of the West*: A travelogue of a randomly generated world, going region by region, even with randomly generated maps!
- *The Cover of The Sun Also Rises*: Literally a description of the book cover of Ernest Hemingway's *The Sun Also Rises*...pixel by pixel.

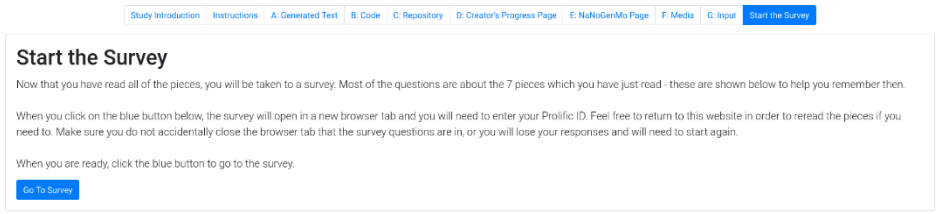


NEW MEDIA

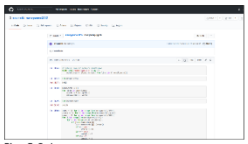
B3, Piece G: Input



B3, Start the Survey page




Piece A: Generated Text


Piece B: Code


Piece C: Repository


Piece D: Creator's Progress Page

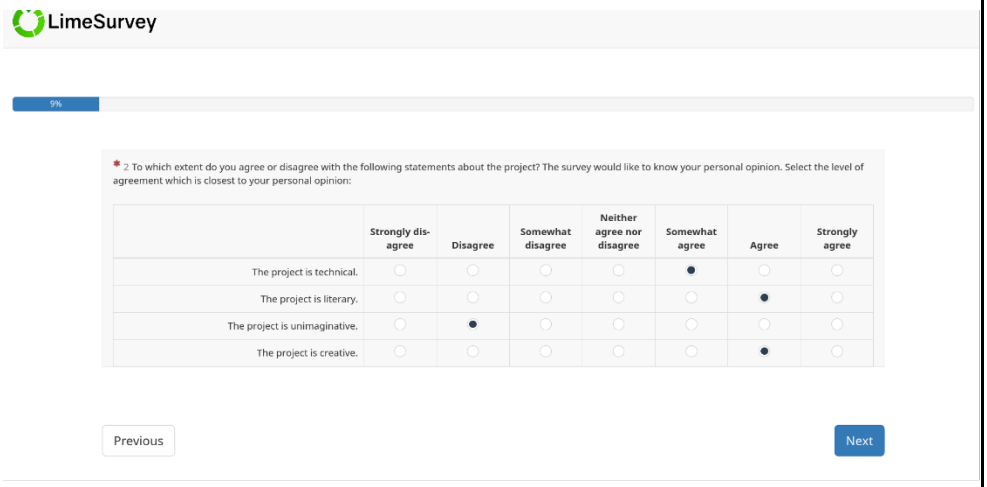
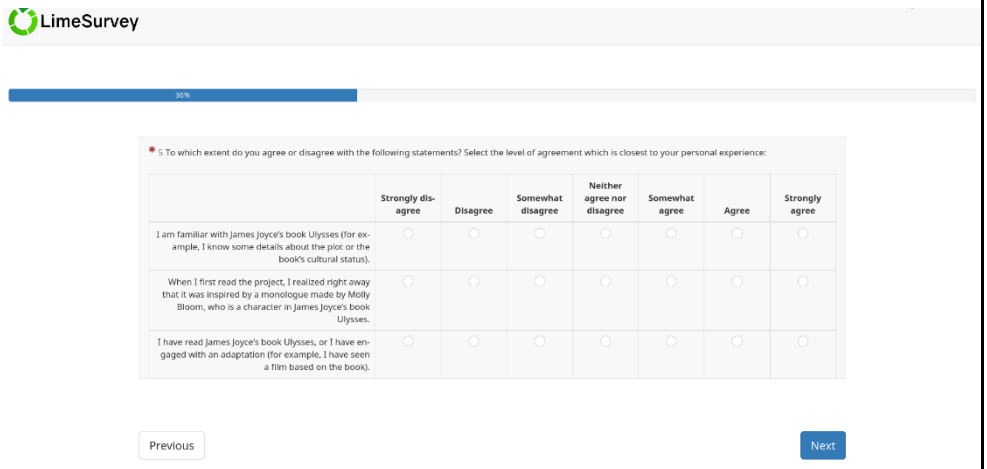

Piece E: NaNoGenMo Page

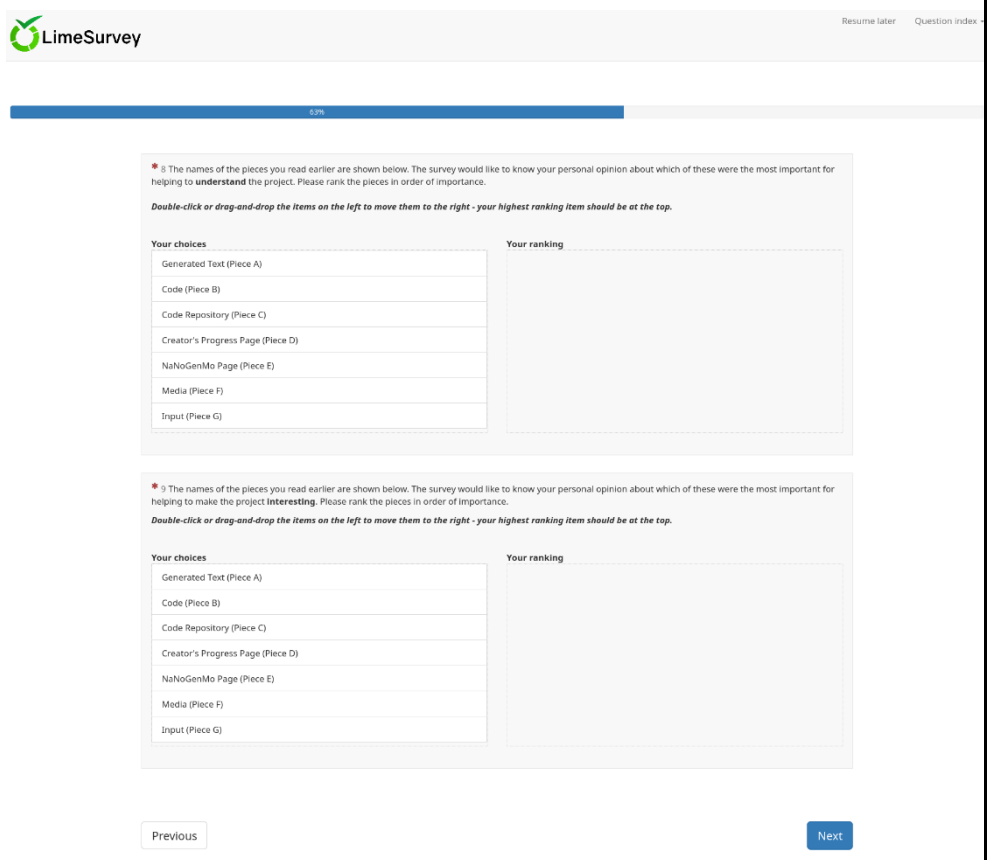

Piece F: Media


Piece G: Input

# Appendix D

Table D1: Survey Part Select Screenshots

| Description | Screenshot |
|---|---|
| B3, Questions A1, A2, A3, and A4 |   |
| B3, Questions A18, A19, and A20 |   |

| | |
|---|---|
| B3, Questions Q71 and Q72 |  |

# Appendix E

## 1. E Workshop instructions

The instructions below were printed on a piece of paper and included in the box of printed generated novels that the workshop study participants were given to read:

> Hello,
>
> Thank you for agreeing to participate in the research workshop. This box contains the two books that you should read beforehand. The workshop will begin with a group reading discussion and then move onto a creative making task. **Please bring both books with you to the workshop at the Winchester School of Art on 11 July from 13:30 - 17:30.**
>
> **Please try to read each book for about an hour.** The books are unusual, so it's perfectly fine if you find that reading in the conventional sense is challenging. Try to find a way to read and engage with them in any way that works for you. You can read the books any time before the workshop.
>
> The workshop is interested in your personal opinions about the books, so **make sure to add your own notes directly onto the pages of each book to document your reading, impressions, and opinions.** Please don't hesitate to write permanent comments, highlight passages, tape on notes, add bookmarks, or any other alterations that you would like onto the books. Do not worry about damaging them. The purpose of this is to help you find parts of the book that you may like to share, question, or discuss with the group during the workshop. It is also helpful for my research because I will be borrowing the books at the end of the workshop and will be using your notes as part of my data analysis.
>
> You will be given new copies of the books and art materials for the creative making task. More detailed instructions will be given during the workshop.
>
> You will have the opportunity to get all of your books back by collecting them from the Winchester School of Art campus when my research project is finished in 2023.
>
> **Please remember to bring the books with you to the workshop.**
>
> Do feel free to email me with any questions or concerns.
>
> All the Best,
> Lesia Tkacz

## 2. E Interview topic guide

**Openings**

- Which book did you read first?
- What were your assumptions about the book before you started reading it?
- Did you know anything for certain about the book before you got it?
- Did you find out anything about the book after reading it? For example by searching for the title online?

**Reading**

- What was your personal opinion of each book?
- Were you able to read the books?
- How did you manage to read or otherwise engage with the books?
- Did you develop a method of reading?
- Did you share the book or its contents with others? How did you do this?
- How did the books compare to each other for you? (in what ways are they distinguishable from each other, or are they both the same? Do they have distinct styles or familiar styles?)

**Literary**

- Was there anything in the books that reminded you or made you think of something else? Such as something else you have read or seen?
- In your personal opinion, how literary is either of the books? Can you give examples?
- Is one more literary than the other? Why? Can you give examples?
- Would you consider either of the works as a work of literature?

**Creativity**

- Is the book creative? Why or why not? Can you give examples?
- What is the source of this creativity? (the author, the machine? Both?)

**Value**

- In your opinion, what is the best way to engage with this book? On your own, sharing with another, as a book club? As a 5 minute flip through?
- In your personal opinion, where (if any) does the value of these works lie?
- Who might they be valuable to?

**GitHub vs book**

- You have read the printed book form of this project, but I have also shown it to you as a project made up of digital files where the text you read is in a PDF document, you can see the computer code, the text data that was input into the code, and the creator's progress or issue page. If I had sent you a link to the project instead of sending you the printed book, do you think that it would have changed your impression of the work? Why or Why not?

- If these books were to be reprinted as a second edition, is there anything that you feel should be added or removed from the print book that might change people's interpretation of it? Or is the copy you have just right?
- Is there anything that you think could be added or taken away to improve the book's reception?

**Authorship**

Did you know anything about the name that is on the cover of the book before or after reading the book? (are they talking about the author or about the computer?).