An exploration into the effectiveness of procedurally generated worldbuilding versus traditional handcrafted narratives in games



Dominic Cohen

COH16656067

16656067@lincoln.ac.uk

School of Computer Science

College of Science

University of Lincoln

Submitted in partial fulfilment of the requirements for the

Degree of BSc(Hons) Computer Science

*Supervisor:* Dr. Olivier Syzmanesyk

April 2022

Acknowledgements

Firstly, I want to thank somebody, and somebody else. Here is another thing.

Abstract

Games often use procedural generation to build digital environments and distribute handmade items, enemies and decorations within them - but rarely do they use it to produce the context for such placements. Most developers handcraft this part themselves, giving the generation system rules of what to place where, rather than letting it produce the content it places on its own. This study aims to compare these two techniques, examining player responses during gameplay. Players will play through a simple dungeon crawler multiple times, in which they will discover various worldbuilding elements placed in the level which they can interact with. Half of the playthroughs contain a handcrafted series of events, and the other half utilises a replacement grammar system to generate a narrative.

Table of Contents

[Introduction 1](#_Toc58407584)

[1.1 Some notes 1](#_Toc58407585)

[1.2 Testing some mathematics 1](#_Toc58407586)

[1.3 Undergraduate Project Report 1](#_Toc58407587)

[1.4 Referencing 1](#_Toc58407588)

[1.4.1 Ludography 2](#_Toc58407589)

[Literature Review 3](#_Toc58407590)

[2.1 Background 3](#_Toc58407591)

[2.2 Related Literature 3](#_Toc58407592)

[Methodology 4](#_Toc58407593)

[3.1 Project Management 5](#_Toc58407594)

[3.2 Software Development 5](#_Toc58407595)

[3.3 Toolsets and Machine Environments 5](#_Toc58407596)

[3.4 Research Methods 6](#_Toc58407597)

[Design, Development and Evaluation 7](#_Toc58407598)

[4.1 Software Development Projects 7](#_Toc58407599)

[4.2 Research Projects 8](#_Toc58407600)

[Conclusions 9](#_Toc58407601)

[Reflective Analysis 10](#_Toc58407602)

[References 11](#_Toc58407603)

List of Figures

3.1 A picture of the Brayford from Google Images. . . . . . . . . . . . . . 5

List of Tables

1. Here is a table. . . . . . . . . . . . 5

Chapter 1

# Introduction

## 1.1 Background

In games, procedural generation is a tool often used in the games industry to create digital worlds, generate and place unique items, and distribute resources and obstacles throughout the game for the player to discover - as an alternative to the time-consuming development of such artifacts on their own.

When creating more qualitative elements, such as item descriptions, developers often handcraft these elements themselves, placing them in the level after generation. The advantage to this is having absolute creative control on the background of the game - however meaningful worldbuilding is a time-consuming endeavour, one that grows exponentially alongside the size of the level.

If there was a way to procedurally generate narrative events within a game, and then distribute items and obstacles across the level that conveys that narrative effectively, then the developer only needs to specify what *can* happen rather than what *will*. Whilst this does inhibit the ability to tell specific stories, it does allow for unique events that “employ apophenia, the human tendency to perceive patterns, in their interpretations” (Grinblat and Bucklew, 2017, 2)

Qualitative Procedural Generation (QPG) is still being explored. Games like *Caves of Qud* and *Dwarf Fortress* have each explored methods of generating history and culture within their games, but both of these methods were developed as tertiary systems within an already developed product, limiting their ability to have a marked effect on the player. These sorts of systems should lend “depth and, crucially, meaning to all of the forms of procedural generation a player encounters in a playthrough” (Johnson M., 2016, 5)

This project aims to explore the effectiveness of these systems further by putting it into the foreground for the player to interact with. In addition, rather than placing historical events into the world at complete random, the project aims to distribute events in a way that makes sense within the context of the level itself - Whereby the beginning of the narrative is found near the beginning of the level, the end near the end etc. This is because good level design is driven by mechanics, “whose primary function is to leverage your mechanics to create a great experience” (GDC, 2018b) - and in order to incentivise engagement with this system, giving the player a piece of the puzzle early in the game serves as both a taster of the narrative and a tutorial of the game mechanics.

Naturally, once a system like this has been created within a game, it would need to be tested against a more traditional worldbuilding method. Using the Game Experience Questionnaire (GEQ), we can have players run through the game multiple times, with half the playthroughs using procedurally generated narratives, and half using traditional handcrafted worldbuilding. This way, we can compare the flow and immersion of the two techniques directly, as well as analysing the players ability to tell the difference between each method of narrative generation - deducing the effectiveness of the new system in the process.

## 1.2 Aims and Objectives

The project aims to develop a procedural narrative generation system that enriches the existing level design of a game through novel worldbuilding. In order to achieve this aim, the project must complete the following objectives:

* Produce a set of 20 generic entities and events that can be strung together and fed into a replacement grammar system.
* Develop a simple, playable, 5-10 minute dungeon crawler to house the project.
* Devise a means of interacting with and displaying 8 generated text snippets within the created dungeon crawler.
* Construct a series of 3 narratives complete with a blueprint for their distribution within the game world.
* Measure player responses to 3 instances of each generation technique using the Game Experience Questionnaire.
* Collect the data from player responses and identify any major patterns that form between the handcrafted and procedurally generated levels.

Many instances of these objectives, such as the 20 entities for the replacement grammar system or the number of handcrafted levels were based purely on what seemed achievable at the time of writing – in many cases these numerical elements were modified as the project progressed.

Here are two equations using the equation editor (1, 2):

(1)

(2)

And here is some text with some nice inline maths, (*x,y*) wow *γ* so cool *ρ*.

## 1.3 Undergraduate Project Report

Currently, this template is set up for use with undergraduate project reports. However, the template can be modified fairly easily to conform to, for example, an MComp project report.

## 1.4 Referencing

It is worth noting that the standard for referencing is Harvard.

### 1.4.1 Ludography

There is an optional ludography for Games Computing students. To cite games, you can cite like any other reference with Harvard styling.

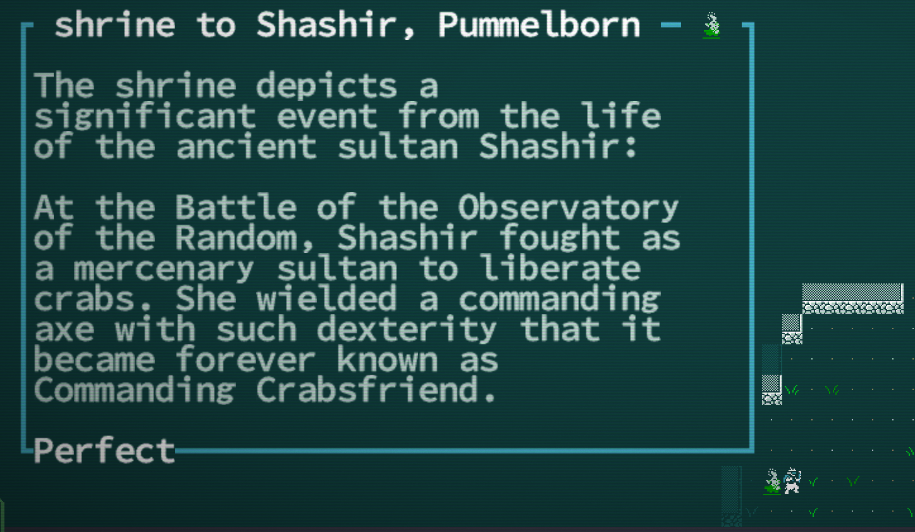
Chapter 2

# Literature Review

## 2.1 Background

The literature review is an essential requirement of any academic project. A comprehensive review of the literature will provide background to the project. This section establishes what you intended to do and shows the reader that what you have done is the result of academic study, rather than an unfounded whim. This section can use the literature review submitted as part of the Interim Report. If you want to add to it, you can, but it is not directly assessed again.

There are many ways to go about procedurally generating content for use in games or other media. Procedural content generation is essentially random generation (or as close as we can get with a computer) overlaid with a series of rules to curate the results. When designing these rules, you are aiming to create a “possibility space (all the kinds of artifacts [the generator] can generate) where most of the artifacts have the good properties, and few (or none) of the artifacts have the bad properties” (Compton, K., 2016). For example, a possible generative method for this project was a constraint solver, where you generate every possible variant of content, testing it against a series of tailored constraints until a desired outcome is reached. This is a viable method, but “too many choices will create a number of possible artifacts to search that is greater than the number of atoms in the universe”, which isn’t ideal.



**Fig. 1.** An example of a shrine depicting a historical event in *Caves of Qud*.

A system such as this relies on a decent replacement grammar system to generate text snippets from the given data - snippets that make sense within the context of both the content of the game and the rules of the English language. This makes it “easy to encode knowledge about a particular artifact and its structure and its options all in one piece of data” (Compton K., 2016). Tracery.io, a free generative text tool, is used by the people at Freehold Games (*Caves of Qud*) as well as a variety of existing projects both in academic and developer communities. Its “scalable complexity allows for a low barrier of entry and a simple JSON file format [that] supports a culture of sharing and remixing” (Compton K., 2015, 1). This means that it will both be easy to embed into the game engine and allow for easy modifications should we wish to alter the games themes or narrative scope.

The creators of the game *Caves of Qud* wrote a paper on the devising of a system to generate the biographies of historical ‘Sultans’ through using a replacement grammar system, “whose rules map sultan properties to text fragments for a variety of narrative circumstances” (Grinblat and Bucklew, 2017, 4). It then modifies the Sultans properties based on the randomly determined results of that event, before moving on to the next event - provided that the Sultan is still alive. The text snippets from these events - such as the one in figure 1 above - are then scattered across the game world, which the player can unearth through interacting with the environment. This system could easily be replicated on a smaller scale, mapping actions spanning a few hours/days rather than an individual's lifetime.

There are many ways to analyse and measure the game experience. Of those that are academically backed, two that stand out are the Player Experience of Need Satisfaction (PENS) and the Game Experience Questionnaire (GEQ). “The PENS questionnaire assesses player experience in the dimensions: Competence, Autonomy, Relatedness, Intuitive Controls and Presence/Immersion” (Johnson D., 2014, 2). Some of these dimensions are useful in comparing the effectiveness of our two systems, but there are ultimately a lot of dimensions of the questionnaire that present no real use to this project.

On the other hand, the GEQ is a much more concise self-report measure of the gameplay experience. Participants indicate their agreement on a number of questions using a 5 point Likert scale, the answers to which measure the player's sense of Competence, Immersion, Flow, Tension, Challenge, Negative affect and Positive affect. In addition, there is a shorter ‘In-game’ questionnaire, which “has an identical component structure” and is used for “assessing game experience at multiple intervals during a game session” (IJsselsteijn, W.A., 2013, 3). This makes it suitable for this project since we can compare the feel of multiple short playthroughs played in quick succession, rather than an extended test at the end of 6+ playthroughs.

## 2.2 Related Literature

As above. If you are undertaking an external project, you should also describe the client and outline the nature of their work or business and explain how the artefact will address the client’s needs.

Chapter 3

# Methodology

Here is a sentence, and you can see a nice picture in Figure 1.



Figure : A picture of the Brayford from Google Images.

A table showing some data is displayed here (Table 1). This doesn’t have to be a table, it could be in the form of a chart or some other form of data representation.

|  |  |  |
| --- | --- | --- |
| First name | Last name | Age |
| Bob | Bobbington | 24 |
| Beth | Wavies | 49 |
| Joe | Bloggs | 37 |
| Billy | Bob | 10 |

Table : Here is a table.

This section will cover a number of aspects of your project where appropriate. **Not all projects will require every section though**. The key thing is that you demonstrate critical awareness of all of the processes that you have employed in your work and that for all sections needed in your report you are presenting a justification for the methods you adopted and not just presenting a list of methods.

## 3.1 Project Management

This project is focused primarily on the design and development of a procedurally generated narrative, and the evaluation of that system within a game. However, in order to carry out this evaluation, a game must be created to house the procedurally generated content, preferably in such a way that the game can be played with or without the content allowing for players to compare and gauge whether or the game is improved by its inclusion.

Because of this, the project requires careful management, since many elements cannot be started without the completion of others. However, elements such as the game that house it need to be managed in such a way that they are only developed to the point of being functional – it is important that they are not overdeveloped at the expense of other tasks. A lean approach would cater to this rather well, reducing wastefulness and streamlining the projects progress as much as possible.

As for the actual day-to-day handling of the project, I intended to implement a simple waterfall model, outlining the requirements before developing the game, implementing the two narrative systems and finally testing and evaluating the merits of each. However, as mentioned previously the potential for the game aspect of the project to become an endless cycle of iterative development means that a bottleneck is far too likely, meaning that a waterfall model wouldn’t be very useful in determining how far along the project actually is.

Breaking the project down into smaller steps and managing those individually would better suit the project, especially if it allowed for the prioritisation of certain tasks over others. Creating a KANBAN board would allow for tasks to be added and ordered within a todo list, letting us visualise the project overall as tasks are moved across the board overtime. It could also allow for the use of a ‘back-burner’ section, allowing wasteful tasks to be pushed to the back to make way for more important elements, preventing the bottlenecks which would otherwise be likely in other methodologies.

Some awareness of project management should be demonstrated in all projects. This section should outline the nature of your project and the specific characteristics that need to be considered in determining what project management methodology you should use. You should identify the specific demands of your project in terms of project management and support your rationale for the selection of a methodology with appropriate and recent academic references. Questions which may be relevant here are:

1. What are the guiding principles and processes in managing your project?
2. What project management methods may be useful for this project?
3. How can you exploit their advantages for your project and mitigate their drawbacks?

## 3.2 Software Development

There should be a methodological analysis of software development approaches used in your project. It is important to note that what is NOT required here is a pedestrian account of popular software development methodologies or a simplistic review of their strengths and weaknesses.

Where relevant, you should give serious thought to the proper design of research and requirements capture approaches. This may include surveys, questionnaires and interviews.

There were two main elements of software development for this project: the system for worldbuilding generation itself, and the simple game that houses it. Naturally our focus is on the system itself, however in order to effectively test and compare such a system there needs to be a game to grant it context.

The pieces of software being developed are different enough to warrant different methodologies for each – the PCG system is using an unfamiliar tool created by a third party, whereas the game is using a well-established game engine to create something that fits a specific set of requirements.

As established, the game needed to be completed in as short a timeframe as possible – it needed to meet the bare minimum requirements for the implementation and testing of the PCG content, nothing more. As such, sticking to the KANBAN approach employed in the project overall was a suitable approach – it meant that since each required element of the game was visualised on the board, going off-task would be discouraged – any development on non-critical features of the game would visibly have no effect on the progress of the project.

Whilst the PCG system originally started out being developed in the waterfall method, it quickly proved to be an unreliable method for creating a system of such complexity. The generative texts were designed first on whiteboards, then put into the inbuilt editor on Tracery’s website before finally being converted over into Unity, where they could be tested on actual in-game items. This quickly proved ineffective, as the syntax between the websites javascript and the C# library in Unity were different enough that small changes to the script required a great deal of editing before testing could occur – an issue that quickly got out of hand as the history generation systems got more complex.

(Dom: put a diagram here of the plan – writing up tracery possibilities on the left, unity implementation in the centre and the tests as deliverables)

To solve this, an agile methodology was adopted – the javascript element of designing the system was cut entirely, in place of translating planned worldbuilding elements directly into Unity script. This allowed for a more iterative approach, where every addition/change to the system could be reviewed quickly and efficiently and any defects or design changes identified. Removing Javascript development also meant that the time to complete each iteration was much shorter, letting there be more iterations in a smaller timeframe and allowing the system and the game to be developed alongside each other within the KANBAN board.

## 3.3 Toolsets and Machine Environments

As mentioned, initially the primary element of the project – the procedural generation of historical events – was handled using Tracery, a Javascript library created by Dr Kate Compton that uses grammars to create interesting and varied texts. It’s a library that I have experience in before, as well as a library that pairs well with game engines for creating unique and interesting content. Time was spent looking at other options such as GPT-2, an AI that autocompletes text when given prompts. However, GPT-2 is known for quickly becoming nonsensical when writing entirely original content – “although its grammar and spelling are generally correct, it tends to stray off topic, and the text it produces lacks overall coherence”. (Vincent, 2019) <https://www.theverge.com/2019/2/14/18224704/ai-machine-learning-language-models-read-write-openai-gpt2>) This ultimately means that a system like this isn’t fit for purpose – the text players encounter in the game needs to relay a specific set of information through the historical content, as well as making sense – straying off topic would pull the player out of the experience. In addition, the use of an advanced AI trained on over 1 billion datasets is a tad overkill for a project of this scope – not to mention the time that would need to be spent integrating it into Unity and feeding it the necessary data to generate text of the correct tone and style.

There are some requirements of the procedurally generated history system that cannot be achieved with Tracery alone – with each event generated, elements such as the name and age of the subject must be saved and reused across each event, giving the impression of continuity. Tracery has methods to save data for use within its system, however there are no methods of preserving that between generations. The solution to this is to use Unity scripts to generate texts piece by piece and storing the important elements as class properties, allowing for them to be initialised and updated in-between each historical event. Each class would then represent a particular figure and their ‘story’, the pieces of which can be jumbled and distributed throughout a level in the game.

Toolsets refer to both software development and to project management, so the coverage should address both. This section will outline the tools for software development and project management process; it will make appropriate comparisons between tools available and argue for the most appropriate selection based on metrics, possibly a matrix diagram and other criteria. DO NOT justify the grounds for using specific toolsets and environments simply because you know them well or have developed skills already.

## 3.4 Research Methods

You should investigate the types of research methods necessary to validly answer the research questions that your project addresses. You should cite relevant sources to justify your choices.

Chapter 4

# Design, Development and Evaluation

This section of the report will vary significantly in both structure and content, depending on the type of project you are undertaking. For example, a Games design project may include a Game Design Document. However, it must be noted that if your project contains significant software development work, this should be presented in the structure expected of a formal development report. If your project involves an experimental evaluation – especially if that evaluation involved human participants – you are expected to write this work up in the format expected in Section 4.2.

## 4.1 Software Development Projects

Include this section if you are undertaking a software development project. You should discuss:

1. Requirements elicitation, gathering, collection and analysis
2. Design
3. Building and programming
4. Testing
5. Operation

### Requirements of the Game

It has already been stated that the game developed for use in this project must be simple: it needs to have a way to display the generated historical content, and said content must be spread across multiple areas such that the player discovers them bit by bit – allowing them to form a narrative themselves through exploration. However, beyond this, the ‘game’ element can take any number of forms.

During the design phase, several ideas were presented and eventually narrowed down. These included a traditional dungeon crawler, a text-based exploration game, a card game etc. There was even the idea for a simple linear display of the PCG worldbuilding elements where the player could just click through them one by one, however this was quickly put aside as it would not provide any insight into the effectiveness of the system within games – only an insight into the quality of the system itself. Ideas like the dungeon crawler held promise – the use of a map of rooms for a player to walk around in would allow for an even distribution of worldbuilding elements, much like how games like Skyrim (2011) distribute books and other reading materials across the game world. A map similar to Enter the Gungeon (2016) could be simple to implement, though creating a PCG landscape to traverse would be a complex task that would have no effect on the project objectives. The game needs to have an element to drive the player – there needs to be cause to explore and collect the texts outside of being told to play the game for research purposes – but these motivations must be simple so as to not alienate the player. Finally, the game must be able to be completed start to finish in a relatively short space of time, as the aim of the project is to have the game played multiple times over to compare differences between different worldbuilding methods.

### Requirements of the PCG

The system for procedurally generating a narrative is going to mimic that of the system used in Caves of Qud: using a generative grammar system to create an individual then applying that individual to a series of scenarios, modifying the individual as needed (Grinblat and Bucklew, 2017). When an individual is created in this system, they begin with a set of properties: their name, age, occupation etc – these properties will need to be stored in memory somewhere outside of the system itself, so that Unity can access them. This is so that we can manipulate the game world in accordance with these events.

The number of events generated can be arbitrary for the purposes of this project, though it would be sensible to keep them to a number reflective of the size of the map being used for the game. If the game uses a PCG map that alters in size with each playthrough, then the generative system used will need to work in accordance to generate an appropriate number of events before they are distributed. If the game uses a hand-built map, then the number of events generated can similarly be handpicked to whatever feels appropriate.

Finally, the number of possible events that can occur within the narrative needs to be enough such that there is little in the way of repetition – if the player is to replay the game, then the experience needs to be different to keep things interesting. However, since each of these events will warrant additional assets to reflect the effect of that event on the game world, too many events will become an overly time-consuming task that will negatively affect the quality of each event. Thus, it is important to find a sweet spot between these two extremes.

### Designing the PCG

Players are going to be replaying this instance multiple times in order to show the variety of results the system can produce. As such, it is important that the instance of gameplay overall is rather short. This will also mean that the player will be able to recall the game in its entirety during subsequent playthroughs – preventing them from forgetting previous instances of procedural generation.

In initial development, the first port of call was to create the figure by which events will occur – Since I decided early on to set this game in a library, I decided to name these figures ‘archivists’. At creation, each archivist would have generated for them a set of properties – their name, age, particular field and specialty, as well as their pronouns – this was a particularly tricky process since each of these properties was then saved to a class instance, and some compound properties such as the pronouns were rather difficult to separate from the generator using the Tracery library. Once this was done however,

## 4.2 Research Projects

If your project includes primary research components it is expected that you present this work in a manner appropriate to a scientific report:

1. Participant recruitment
2. Evidence that ethical procedures have been followed
3. Study design (short summary of research methods section) – including hypotheses/research question as appropriate
4. A detailed description of the procedure
5. Results of experiment
6. Analysis of results. Consider the results of your work with respect to both your own specific hypotheses/research question and wider context identified in your literature review.

Chapter 5

# Conclusions

The results from this project indicate that ...

Chapter 6

# Reflective Analysis

The project went well ...

# References

Compton, K. (2015) Tracery: *An Author-Focused Generative Text Tool.* [online] Available from: http://fdg2015.org/papers/fdg2015\_extended\_abstract\_18.pdf [Accessed 7 November 2021]

Compton, K. (2016) *So you want to build a generator?* [blog]. 22 February. Available from:

<https://galaxykate0.tumblr.com/post/139774965871/so-you-want-to-build-a-generator> [Accessed 2 November 2021]

Floridi, L. and Chiriatti, M. (2020). *GPT-3: Its Nature, Scope, Limits, and Consequences* [commentary] February 20. Available from:

<https://link.springer.com/content/pdf/10.1007/s11023-020-09548-1.pdf>

GDC (2016). *Dwarf Fortress, Moon Hunters, and Practices in Procedural Generation.* [vlog]. May 27. Available from:

https://www.youtube.com/watch?v=v8zwPdPvN10&list=PL0yMAu5TQBt8iL-Sdd5h1wNvFiFcYK9Hs&index=5 [Accessed 13 September 2021].

GDC (2021). *End-to-End Procedural Generation in Caves of Qud.* [vlog]. June 23. Available from:

https://www.youtube.com/watch?v=jV-DZqdKlnE&list=PL0yMAu5TQBt8iL-Sdd5h1wNvFiFcYK9Hs&index=4 [Accessed 13 September 2021].

GDC (2016*). Interior Design and Environment Art: Mastering Space, Mastering Place.* [vlog]. November 12. Available from:

https://www.youtube.com/watch?v=WWXsmnlmADc&list=PL0yMAu5TQBt8iL-Sdd5h1wNvFiFcYK9Hs&index=2 [Accessed 6 November 2021]

GDC (2018a). *Procedurally Generating History in Caves of Qud.* [vlog]. June 12. Available from:

https://www.youtube.com/watch?v=H0sLa1y3BW4&list=PL0yMAu5TQBt8iL-Sdd5h1wNvFiFcYK9Hs&index=3 [Accessed 5 September 2021]

GDC (2018b). *Ten Principles for Good Level Design.* [vlog]. February 9. Available from:

https://www.youtube.com/watch?v=iNEe3KhMvXM&list=PL0yMAu5TQBt8iL-Sdd5h1wNvFiFcYK9Hs&index=1 [Accessed 4 November 2021]

Grinblat, J. and Bucklew, C., (2017*). Subverting Historical Cause & Effect: Generation of Mythic Biographies in Caves of Qud.* [online] Available from:

https://www.freeholdgames.com/papers/Generation\_of\_Mythic\_Biographies\_in\_CavesofQud.pdf [Accessed 27 October 2021].

Hall, H., Williams, B. and Headleand, C., (2017) *Artificial Folklore for Simulated Religions.* Lincoln, UK [online]

Available from: https://ieeexplore.ieee.org/document/8120332

IJsselsteijn, W.A., de Kort, Y.A.W., Poels, K. (2013) *Game Experience Questionnaire.* [online] Available from:

https://eclass.uoa.gr/modules/document/file.php/DI411/PAPERS/Evaluation/Game%20Experience%20Questionnaire%20English.pdf

[Accessed 6 November 2021]

Johnson, D., Gardner J., Perry R. (2018) *Validation of two game experience scales: The Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ).* [online]

Available from: https://bit.ly/3qPJFBV [Accessed 6 November 2021]

Johnson, D., Watling, C., Gardner, J., Nacke, L. (2014) *The Edge of Glory: The Relationship between Metacritic Scores and Player Experience.* [online] Available from:

https://dl.acm.org/doi/pdf/10.1145/2658537.2658694?casa\_token=7quzFk9KTMEAAAAA:lJgwxXjEJkdS\_F89cHIVkzqlADmirpmeYgCd-RbZvrqNfJI1-UnsVDOx2lpaxaKE2wli23DBekUk [Accessed 6 November 2021]

Johnson, M. (2016) *Towards Qualitative Procedural Generation.* [online] Available from:

https://web.archive.org/web/20170610155631id\_/http://www.ccgworkshop.org:80/2016/Johnson.pdf [Accessed 10 November 2021].

James Vincent (2019) *OpenAI’s new multitalented AI writes, translates, and slanders.* The Verge[article] Available from:

<https://www.theverge.com/2019/2/14/18224704/ai-machine-learning-language-models-read-write-openai-gpt2> [Accessed 12 January 2022]