

Topic to Tale: Combining Narrative-Focused Design Methods to Increase Understanding of Narrative Importance in Serious Games

Jan Willem de Ruig

May 2023

Abstract

Narrative plays a key role in serious games when it comes to player experience and learning, and likewise studies have found a compelling narrative to improve player investment. Yet despite the benefits, narrative in commercial games tends to lag in quality, and in serious games tends to only serve contextual purposes. Creating a truly interactive narrative is challenging due to various mechanical and practical constraints, and the pedagogical dimension added in educational serious games only increases complexity.

Methodologies proposed in the past, such as Narrative Serious Game Mechanics (NSGMs) and iterative learning cycles, show potential but are limited by the scope they respectively address. I have combined the best aspects of various such methods to alleviate these limitations. I created a prototype game, which was refined through focus groups, and expanded into two versions: One containing the proposed Narrative-to-Game and Game-to-Narrative mechanics, and one baseline containing conventional narrative techniques. To evaluate the relative influence of the proposed method on player experience and learning behaviour, the two versions were compared through a survey study.

Ultimately, the study failed to produce sufficient results to conclude an improvement in learning efficiency or player experience, due to various limitations. However, both the proposed methodology and baseline methodology are shown to facilitate learning behaviour. While it thus cannot be concluded that the proposed methodology performs better, there similarly are no sufficient results to reject it. Further research is required.

Contents

1	Introduction	3
2	Literature Review	4
2.1	Narrative Mechanics	4
2.2	Narrative in Serious Games	5
2.3	Narrative Serious Game Mechanics	5
2.4	Hermeneutic Learning Cycles	6
2.5	Learning Mechanic Integration	6
2.6	Summary	7
3	Method Plan	8
3.1	Combined Methodology	8
3.2	Prototype Creation	9
3.3	Comparison and Evaluation	10
4	Prototype Foundation	11
4.1	Initial Prototype	11
4.2	Evaluation Method and Goals	13
4.3	Results and Discussion	13
4.4	Iteration	14
5	Final Prototype	15
5.1	Narrative and Serious Content	15
5.2	Narrative Format Implementation	16
5.3	Variant I	18
5.4	Variant II	19
6	User study	21
6.1	Questionnaire	21
6.2	Data Collection	22
6.3	Participants	23
6.4	Analysis	23
6.5	Results	25
6.5.1	T-Test	25
6.5.2	RM-ANOVA	26
6.6	Discussion	27
7	Conclusion	28
8	Limitations and Future Work	28
9	Appendix A : QuickScan 1	33
10	Appendix B : QuickScan 2	42
11	Appendix C : Questionnaire	49

1 Introduction

Over the past decades, our understanding of interactive digital narratives has been steadily evolving. We have advanced from narratives as simple as 'Save the Princess', the bare minimum required to set the gameplay in context, to complex and living worlds such as that of Red Dead Redemption 2 [28][36]. In 2002, Henry Jenkins described well the ways in which locations and environments within games could be used to convey meaning and evoke feeling [17]: Environmental storytelling can be used to evoke narrative associations, to embed narrative information within the scene, or provide resources for emergent narratives. Bethesda's 'Radiant AI' fundamentally changed the way in which players could perceive the game world in 2006 by allowing in-world NPCs to act in more realistic and believable manners, and the technology has been improved for games released since [10]. And in a 2014 conference paper, Koenitz et al. compiled various theories to put forward a unified understanding of interactivity in digital narratives (IDNs), which cements IDNs' existence in the narrative universe and highlights its emphases on elements such as environment, character, and user interface. The understanding was expanded upon in a 2015 publication by mostly the same authors [21][20], engaging more in the history, theory, and practice that surrounds the IDN.

Indeed, research has repeatedly shown the significant role a narrative can play in digital games, and the benefits it brings to the gaming experience [13]: Interesting mechanics may make games appealing and enjoyable, but narrative can far more often elicit meaningful experiences. Yet, in a narrative analysis of Fallout 3 McDaniel et al. perform in a 2010 publication, the authors describe the ways in which narrative still seems to lag behind other core features of modern games, like graphics and gameplay mechanics [25]. While it is unheard of for successful modern games to feature major graphical or mechanical issues, game designers seem less concerned about producing consistent quality narratives.

On the topic of serious games, a similar trend becomes quickly apparent. Naul and Liu have documented the beneficial effects of a narrative on various important learning facets, such as motivation, immersion, and learning gains [27]. Several design methodologies to combine narrative with serious content have been proposed and discussed, such as methods of weaving together layers of mechanics through purpose and structure, or atomic iterative learning cycles using hermeneutic theory [23][5]. However, in a 2014 study by Lim et al., a vast minority of the serious games analysed incorporated dramatic structures or narrative elements properly for learning purposes, and instead use the narrative of a serious game mainly as a framing device [23]. This again goes to show that, while the theory exists, narrative in practice might not often enough be used to its fullest potential.

This begs the question, why? If narrative plays such major roles in both serious and non-serious games, why are games still found lacking in the implementation? The answer, according to McDaniel and Kampa et al., may have to do with technical challenges in making interactive digital narratives truly interactive [25][19]: The amount of required branching paths can easily become overwhelming, and narrative disconnect due to limitations in level design is hard to avoid. And indeed: For all the research that is out there, a complete and coherent methodology for creating such interactive digital narrative based on a serious topic seems nonexistent. Lim et al.'s Narrative Serious Game Mechanics (NSGMs) come close in the way they strive to unite different mechanics, but their proposed procedure focuses perhaps too much on the narrative and pedagogic mechanics and does not much consider the game itself [23]. Atmaja's approach with atomic iterative learning cycles takes the triumvirate of information, narrative, and interactivity much more closely into account, but the 'hermeneutic cycles' employed are vaguely defined and require, by the author's own words, more thorough investigation [5]. And yet other design methodologies, such as the mode of integrating learning mechanisms (LM-integration) proposed by Barbosa et al. in which the primary gameplay and learning mechanics are more independent of each other, are methodically sound yet

do not concern narratives at all [6].

It becomes clear that all these methodologies carry a piece of the puzzle, and while there is some overlap, complement each other's missing aspects. In this thesis, I will aim to combine these aforementioned methods; NSGMs, atomic iterative learning cycles, and LM-integration. Through doing so I will answer the question of how this combination affects the effectiveness of a narrative-based learning game in terms of complex issue understanding. I will create a simple prototype exemplifying this combined methodology, and evaluate its impact on motivation, immersion, and learning gains through a comparative user study, where it will be compared against a baseline prototype using more conventional narrative structures.

2 Literature Review

Important pieces of literature in this thesis are Lim et al.'s 2014 paper on Narrative Serious Game Mechanics, Atmaja's 2022 paper on hermeneutic cycles, and Barbosa et al.'s 2014 paper describing learning mechanic integration. Beyond these three, there exist several auxiliary studies, papers, and books that pertain to the design and development of a serious game, which I plan to consult in the creation of the prototype.

2.1 Narrative Mechanics

But first, let us take a closer look at how narrative mechanics typically work in digital games. Narrative mechanics have been discussed at length; Jenkins and Juul among several others engaged in the narratologist-ludologist debate on the way in which interactive games can tell stories in the first place, without diverging from gameplay [17][18]. And likewise, books and theses have been written to gain an understanding or establish a unified theory of how narrative mechanics work. [21][20][37]. According to Bauer et al., narrative mechanics can come in many shapes and forms. Like non-narrative game mechanics, they incorporate rules, rewards, and punishments – but, rather than a reward being an amount of currency gained or a punishment being some score lost, a narrative mechanic engages the narrative structure of the game in these rules, rewards, and punishments [37]. The decision a player makes does not just have an impact on the state of their inventory, or the numbers associated with their character, but directly affects the story this player perceives. Dialogue options in Mass Effect are a prime example of this; by choosing certain options with morally different connotations, the player directly affects the way NPCs will interact with them in the future [7].

In a 2019 literature review, Naul and Liu analyse various academic sources showing the positive impacts of a good narrative on the overall user experience and compile their broadly positive findings: Strong relationships were found between narratives, and immersion and engagement [27]. And, according to Dubbelman [12]:

“Narrative game mechanics invite agents, including the player, to perform actions that support the construction of engaging stories and fictional worlds in the embodied mind of the player.”

Thus, if narrative provides the player with motivation, engagement, and immersion, and if narrative mechanics allow the player to directly affect the course of the narrative, its effectiveness should increase. Arrambide reasoned as much [4]: The agency perceived by the player through these narrative mechanics can increase enjoyment and create a positive experience.

2.2 Narrative in Serious Games

As previously mentioned, narrative, and narrative mechanics by extension, are widely considered to be beneficial in serious games, again best exemplified in the meta-analysis performed by Naul and Liu, where the positive impact of narrative on motivation and learning gains were also discussed [27]. For instance, Parker and Lepper found both immediate learning and knowledge retention improved in third-grade students with narrative conditions in play, regardless of the narrative theme [29].

Naul goes on to explain how aspects of the narrative, such as an endogenous fantasy or compelling and empathetic characters, play an important role in fostering an effective learning environment:

“Characters to whom learners relate hold the potential to influence their attitudes and decisions.”

This is exemplified in a study performed by Kotler et al., finding that children tend to prefer foods their favourite characters like, and were more willing to try new foods these characters promoted [22]. And according to Routledge, narrative learning is all about empathy and personalization; making the player truly care about the outcome [33].

Reeve, meanwhile, argues the importance of the structure of narrative, and the way in which a narrative is provided to the player [32]. For instance, a narrative presented through cutscenes that take away player control can have detrimental effects to the player experience. Reeve describes this structure as a ‘comb narrative’, in which the over-arching narrative can be considered like a comb and user activity fits disconnected in the interstices between the ‘teeth’, and he contrasts this against the ‘mesh narrative’ which integrates uninterrupted activity and narrative. Using the serious game ‘Axon’, Reeve demonstrates that the relationship between user activity and narrative drama are not contradictory.

Narrative, therefore, plays important roles in serious games when it comes to immersion, learning, and motivation - but it is important to consider the way in which narrative is integrated, to make sure it does not incidentally take away from player agency and enjoyment.

2.3 Narrative Serious Game Mechanics

Now, let us consider the NSGMs proposed by Lim et al. [23]. A narrative serious game mechanic, plainly put, is a game mechanic that draws primarily on narratives to promote knowledge acquisition. As such, an NSGM often comprises several game mechanics in order to achieve its purpose and establish a process and structure. As they describe it, these mechanics are like individual strands, woven together on different layers - learning mechanics, game mechanics, and narrative mechanics - into one NSGM.

Additionally, there is an emphasis on narrative as being the first contact between author and audience. According to Lim et al., stories and narratives are vessels for meaning, a sentiment shared by Naul and Liu, Reeve, and Bauer et al [27][32][37]. Lim et al. go on to discuss how a proper narrative provides exposition, facilitates guidance to support and enhance the learning experience, and gives room for reflection and feedback. Within the pedagogical context of a serious game, these three narrative tasks are important aspects, so the paper concludes [23].

An example of such NSGM is provided by Lim et al., in the form of a breakdown of the mechanics at play in the GoVenture Any Business serious game [16]. Mechanics are broken down at the levels of Game Mechanics (GM), Learning Mechanics (LM), and Narrative Mechanics (NM), dispersed over the various periods of the game. This results in an overview of what mechanics work together during what gameplay section, such as narrative guidance working with experimentation learning during the decision-making period of the game, and showcases how LMs and NMs come together

to form an NSGM. Traditional narrative mechanics on the other hand, as discussed in the previous segment, do not always have this direct link to learning mechanics integrated.

The emphasis on narrative guidance, feedback, and reflection is in line with other papers discussed in the following subsections. Additionally, the NSGM being comprised of various mechanics drawn together by purpose, process, and structure means it lends itself well to combination with other theories and methodologies. This thesis proposes a structure in which the NSGM encapsulates mechanics on a macro-scale, while other theories address the individual mechanics themselves.

2.4 Hermeneutic Learning Cycles

Atmaja introduces the concept of hermeneutic learning cycles and attempts to adapt these to complex issues [5]. Hermeneutics describe a complex interpretation methodology that is most often used in the understanding of spiritual or philosophical works, such as the Bible [38]. It describes a circular process, where the context of a work informs understanding of the text, and new understanding gives new insight towards the context. However, the term has been used in the field of digital games as well, to describe both the understanding of games as cultural objects and the real-time understanding of the game world itself [3][24]. Atmaja describes the real-time hermeneutic cycle as a process in which the player acts in a context, receives feedback, and gains the opportunity to reflect on that feedback to gain a new understanding of the context [5]. In other words, through interaction with the game world a player gains a new understanding of that world, which, in turn, allows the player to find new ways of interaction.

This process might best be exemplified in metroidvania games, such as puzzle-platformer *Toki Tori 2* [8]. In this game, the player controls a chick in a world full of foreign flora and fauna. Aside from basic movement, the player has two core actions available throughout the entire game: The player can whistle, and the player can stomp. Other than a limited amount of utility songs the player unlocks over time, this is the extent of character ability progression in the game, and in order to win the player must learn about the world they inhabit and the interactions between themselves and the various other critters. Through trial and error, the player might witness that certain creatures are attracted by whistling and are scared off with a stomp. This new understanding of the environment leads to the player being able to discover and try out more interactions and learn about the environment from those in turn.

Atmaja poses that these cycles of hermeneutic understanding are not limited to in-game worlds or mechanics but may also apply to understanding complex issues. Furthermore, while usually this hermeneutic understanding develops throughout the extent of the entire game or game level, according to Atmaja they may be divided up into low-level 'atoms', each of which concerns a piece of knowledge. By chaining these cycles together, understanding of the larger issue becomes possible [5].

Such cycles of understanding may tie in well with the concept of narrative serious game mechanics proposed by Lim et al. [23]. However, where Atmaja places a focus on the the co-operation of information experts, narrative experts, and interactivity experts, Lim et al. focus more on the relation between narrative and pedagogy, and game design itself is considered to a much lesser extent. Still, the atomic nature of the hermeneutic cycle provides a unique opportunity for several such mechanics to be incorporated under the umbrella of one NSGM.

2.5 Learning Mechanic Integration

Lastly is the consideration of integration of serious mechanics. In a 2013 proceeding, Molnar and Kostkova already found that for some learning objectives, integrating these through game mechanics

as opposed to text has a positive effect on learning behaviour [26]. And in the same year, Procci et al. showed the positive effect of content-relevant minigames on learning [31]. The next year, Barbosa et al. proposed a new methodology for designing serious games and integrating serious game mechanics, partly based on these findings [6].

The theory suggested by Barbosa et al. is not dissimilar to the narrative comb structure Reeve discussed [32]. However, where Reeve described an over-arching narrative interleaved with user activity, Barbosa et al. consider an over-arching game with interluded learning mechanics. The idea is that by making the learning mechanics independent and parallel to the main game, it is easier to include and engage with learning content. Additionally, this proposal would allow the main game to focus primarily on fun [6].

Importantly, the original proposition of Barbosa et al. considered this form of LM-integration outside of the context of narrative learning. The serious minigames in Barbosa's example prototype, for instance, stand broadly outside of the scope of the rest of the gameplay; they are related to the main game and share the same thematics, but introduce different mechanics, and have no bearing on any over-arching plot threads [6]. This idea clashes with the concept of hermeneutical learning, which itself relies on this in-world feedback. However, the manner of interspersing important learning points throughout the game in a way that allows the gameplay itself to focus on fun is still an intriguing line of reasoning. The key difference between the approach taken by Barbosa et al. and this thesis is that instead of embedding learning mechanics in minigames, they will be embedded in the narrative instead.

2.6 Summary

Reviewing these facets of narrative in serious games, and various methods of integrating learning mechanics and facilitating narrative-based learning, it is clear that the subject is an incredibly complicated one. Interactive digital narratives must marry both ludological and narratological principles, integrate narrative mechanics with gameplay, and carefully consider the way in which narrative is structured and presented [37][32]. Additionally, serious games contain an extra dimension of pedagogical elements which must be included in such a way to facilitate learning. However, the importance of this process becomes similarly apparent [27][12]. There is a heavy emphasis on the pedagogical aspects, such as guidance, feedback, and room for reflection, which narratives are uniquely posed to make use of.

Narrative Serious Game Mechanics are one of several attempts to consider all these aspects in designing narrative and mechanics for a serious game. They address purpose, process, and structure, and provide a framework to consider how different layers of mechanics should interact [23]. However, the theory presented is quite dense, and does not elaborate much on the nature of the mechanics it strives to unite. Atomic hermeneutic cycles are another method, which places emphasis on the combination of information, narrative, and interactivity [5]. However, study on hermeneutics in games is limited, and the emphasis on an atomic nature risks losing sight of the grander system. The potential of these methods to supplement each other becomes apparent from their focus and magnitude: NSGMs consider the link between different layers of mechanics on a macro scale, where hermeneutic cycles take a more atomic approach to individualised concepts.

The manner in which these combined mechanics are integrated is the final point of consideration, where both Reeve's comb and mesh structures, and the LM-integration method proposed by Barbosa et al., become relevant as guides for how to weave the narrative and gameplay together [32][6]. The player should be able to explore the various learning mechanics at their own pace, after all, and without the game removing agency or reducing engagement.

However, since these various methods have been developed on their own, further study is required

to identify whether this combined approach has any merit compared to serious games without as much narrative mechanics. This is the research gap I address in my thesis.

3 Method Plan

My main goal was to combine several existing methodologies; the NSGMs proposed by Lim et al., the hermeneutic cycles proposed by Atmaja, and the LM-integration proposed by Barbosa et al. In order to determine whether this combination results in a viable method, a baseline comparison was required.

In order to develop and evaluate this combined methodology, there are several steps to follow.

- The methodology itself is worked out and set to paper, and a workflow be developed. Throughout this workflow, it is important to keep in mind the most vital parts of serious game design, such as identifying the goal of the game, and considering the interactions between developers and domain experts.
- In order to exemplify the proposed methodology, a prototype game following this design has been created. The focus is placed on the narrative, which will contain much of the serious content, to allow the gameplay to focus on fun.
- A second prototype variant was created alongside the first, using different design principles, for comparative purposes. This game mirrors the first in gameplay, but narrative and serious content is presented using contemporary narrative methods. This prototype will be used as a baseline to gauge the effectiveness of the proposed methodology.
- Lastly, the comparison and evaluation itself. This takes the form of an anonymous user study taken over a group of peers. Each individual has been assigned to one of two groups, which each received the different prototype variants. After having played through the prototype, each individual filled out a form containing various questions regarding their experience and learning behaviours. The responses are used to measure the effectiveness of each approach against the other.

Now let us consider these points in more detail.

3.1 Combined Methodology

In order to create a methodology for creating serious games, it is important to recognise the game's design goals in the first instance. Unlike entertainment games, for which inspiration can strike in the form of a mechanic element, a narrative theme, a broad subject or a small idea, serious games typically always start with their serious content. Stakeholders such as schools or institutions have a certain topic they would like to teach through gamified means, and as such the serious topic is almost always the initial point of design in a serious game. Atmaja's hermeneutic approach, to name an example, begins with considering the complex issue and informational requirements, and places the design considerations of the information expert prior to the considerations of the narrative and interactivity experts [5].

In fact, the order of the design process Atmaja proposes is methodically sound: Basic requirements are considered first, such as information requirements, audience specifications, the game genre, and lastly technical specifications [5]. After these have been established, an over-arching proto-narrative can be formed and iterated over. The identification and establishment of NSGMs,

being a backbone of narrative learning, would according to this flow-chart be considered during the formation of these initial specifications: it pertains to both information requirements and technical specifications of the game, and informs the shape of the proto-narrative. Within these identified NSGMs can then, during the iterative detailed design, several atomic learning cycles be formed.

Three key aspects of narrative learning that both Lim et al. and Atmaja touch upon, are guidance, feedback, and reflection [23][5]. The narrative should offer the user assistance when desired, though not to a point where it becomes overbearing. The narrative should offer feedback to the user's choices, which contain or reinforce the learning content. And the narrative should offer time for the user to reflect on their decisions, and how those might have led to this outcome. Given the pedagogical implications of these three aspects, they play a significant role in the proposed methodology as well.

Lastly the integration between narrative, narrative mechanics, and game mechanics. Indeed, the theory proposed by Barbosa et al. is at odds with some of these methods, as it considers the separation of learning mechanics and game mechanics, whereas both NSGMs and hermeneutic learning cycles embrace the togetherness of these facets. However, it is not this separation we are interested in, but rather the framework of enjoyable gameplay alternated by serious narrative.

Consider for instance an 'infinite' game, such as Tetris [30], wherein the core game mechanics do not stipulate a concrete end-state other than a 'Game Over'. By integrating hermeneutic narrative cycles within such gameplay, the player will be able to alternate between gameplay and learning without the gameplay ending before the player properly understands the serious content. These cycles can act as narrative gates, allowing players to continue to the next 'level' of understanding when they grasp the initial content, and allowing them to repeat the cycle of learning when they do not. This would allow players to learn at their own pace, without the game deciding a level is over before the content is fully assimilated yet.

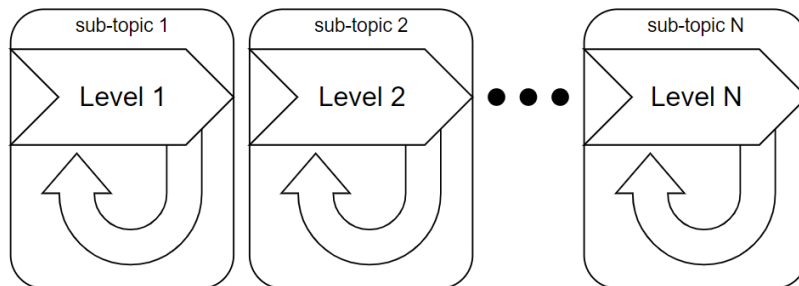


Figure 1: Diagram showcasing hermeneutic theory applied to an infinite-type game

3.2 Prototype Creation

As mentioned in the previous segment, when creating a prototype following this methodology the first step is to select a serious topic. Absolute generality of the methodology has not yet been established, but since it is intended to work for a broad range of topics, the need to emphasise my reasoning for choosing a certain complex issue is important only to contextualise the game. When choosing a complex issue, both the feasibility of the prototype design process and the manner in which the complex issue can interact with narrative and mechanics are more important to gauge motivation and learning gains than the content of the issue itself. As long as the content will be the same across the prototypes tested, the level by which it is internalised by the player can be

accurately measured across the prototypes.

However, this should not be taken as an invitation to simply make up a complex issue: While it is true that the content is less relevant, it is still important even in a prototype to not engage in falsehoods. For these reasons, I opted to select a topic I myself am passionate about, and for which I am confident in my ability to find sources and explanations grounded in scientific reality. In order to keep the prototype simple, the topic itself will be limited as well. The concept of 'induced demand' and its implications on urban design seems fitting as serious content, given its narrow scope.

The concept of 'induced demand' is a phenomenon that has been debated in urban planning circles for decades. In 2001, Cervero released a study to review and assess literary work on the topic, drawing a conclusion that both highlights the importance of recognising the effects of induced demand, while simultaneously admitting the nuance of the situation [9]. It appears the discussion has since bore little fruit, as much later, Speck (2012) described induced demand as "the great intellectual black hole in city planning, the one professional certainty that every thoughtful person seems to acknowledge, yet almost no one is willing to act upon." [35] In recent years, the discussion still crops up from time to time, in particular in video essays from urbanist YouTube channels, such as Alan Fisher, the armchair urbanist [14] and Not Just Bikes [34].

'Induced demand' describes the idea that increasing the ease of use of a particular mode of transport will be met with an increase in the use of that mode. This is mostly relevant in the automotive world, where proposed solutions for congestion can often boil down to construction of additional lanes - something which, according to the principle of induced demand, will only cause more cars to use those lanes in the end. Given that the car-centric design principles carried by many governments internationally seem to hold little regard for the concept, a serious game to educate people about the very real effects of induced demand and how to harness this seemed culturally relevant while remaining limited in scope.

Both versions of the SG prototype were created using the Unity engine, as I am experienced in both the tool and the language it uses. The primary loop of the narrative game contains a simple infinite-type game, during which the player will be allowed to make narrative choices on top of gameplay choices. The gameplay itself will be kept simple and slow, to allow the user to keep up with narrative dialogue throughout. Through narrative choice, the player can explore different facets of the serious topic, and once they show sufficient understanding they can break out of the hermeneutic loop and continue to the next level.

The narrative itself is crafted according to the methodology: The first consideration is the serious sub-topic that each level will address. Character archetypes are added to the game by considering stakeholders and their positions within the serious topic, and a narrative emerges from the way in which these characters might interact with each other or the choices the player makes towards these characters.

The baseline prototype will make use of contemporary narrative structures; primary gameplay will be the same as for the narrative game, but with narrative choices removed. Narrative instead is presented in segments prior to and after each level.

3.3 Comparison and Evaluation

A user study is conducted across peers in order to compare and evaluate the different prototypes. Participants were provided an informed consent form prior to the study and grouped in two different groups. Every group can only play one of the two prototypes, as otherwise learning effectiveness would cross-contaminate. In order to produce statistically significant results, however, it is important to have a sample size large enough to adequately cover both prototypes. More data is naturally always better, but I strived for a sample size of at least fifteen participants per group - though, if

possible twenty would be more satisfactory. I balanced the groups in terms of participant numbers and demographics as best I could, by first assessing the interest of individuals, and forming the groups accordingly. Aside from demographic information, data collected during these studies will focus primarily on the aspects of player experience and learning gains.

In order to measure the effectiveness of a serious game, learning gains are naturally an important variable. They can be measured by gauging the extent of participant knowledge on the subject both prior to the gameplay and after. The participants will be asked to rate their understanding of the sub-topics discussed on a Likert scale, and additionally will be asked to fill in a small quiz about the topic. Their understanding and level of growth can be extrapolated from these and compared accordingly.

Both engagement and enjoyment are important both for the serious aspect, and especially for the game aspect. Essential information here is to know whether the alternating of gameplay and narrative was disruptive of user engagement, and whether the cyclic nature of the game was repetitive to an annoying extent. Both of these are potential pitfalls in this design methodology, and thus important in the evaluation thereof to ensure the aspects of the serious game that are supposed to be fun, are actually fun.

Equally important to the enjoyment of the game as a whole, and the effectiveness thereof, is the enjoyment of the narrative. If the narrative contains serious topics, and is both enjoyable and memorable, the player may be able to remember the serious issues better by proxy. However, if participants indicate the narrative is bland or uninspiring, this will likely have similar, opposite effects on learning effectiveness.

The expectation is that by allowing the user to decide the pace by themselves through the cyclical nature of the learning experience, this methodology will perform better in terms of learning gains compared to the non-narrative approach. However, this does depend on both the core gameplay being fun, as well as the narrative being compelling and informative. Thus, all of these are to be measured.

4 Prototype Foundation

The implementation of the proposed methodology must begin with a suitable base prototype, a stable foundation on which the method can be tested rigorously. Since the methodology revolves around embedding serious content in narrative to allow the gameplay to focus on fun, said gameplay would need to be designed to properly facilitate this. The risk is that if a prototype with the methodology is implemented right away, but the gameplay itself is lacking, evaluation results caused by poor gameplay implementation can reflect negatively on the methodology. In order to avoid this, a base prototype was created and evaluated through a focus group.

4.1 Initial Prototype

The first prototype was created following design patterns identified in the method plan. Important to this creation is the understanding that, while this iteration of the game is lacking in serious content, the necessary building blocks to add this content eventually still need to be there. This extends to the mechanics of the game, but also the general aesthetic and design philosophy. NSGMs should still be considered for interaction with a non-serious narrative, as this narrative needs to be embedded in the gameplay, and not just layered on. And since the serious content to be embedded in the game will regard the concept of 'induced demand' in urban planning, a traffic-like aesthetic seems fitting for the game.



Figure 2: Initial UI layout

As mentioned, this iteration was developed in the Unity engine, and draws heavy inspiration from match-three games such as the Bejeweled series [15]. Given the game's thematics, the baubles take the form of common participants in traffic: yellow bikes, red cars, purple trucks, cyan pedestrians, and blue buses all inhabit a 5-by-9 grid. These dimensions were chosen arbitrarily, with a bias towards odd numbers and a wide rectangular shape.

The area beneath the grid hosts a textbox for narrative, and the side-margins contain various buttons, indicators, and image assets to accompany the text. There is an indicator to show when the player may make a move or when the game is working - i.e. dropping of baubles after a match, spawning new baubles, or looking for compound matches - which is stylised as a traffic light. There is a refresh-button present, to remove and regenerate all baubles on the field in case there are no more matches. Typically with a match-three game, this is internally calculated and results in a 'Game Over' state. However, given the relative complexity of these calculations, and the undesirable nature of a 'Game Over' state in this prototype, such a system was not implemented and delegated to the player instead. The button is normally invisible but appears when the player does not make a move in some time. Lastly there is a score indicator, included because the style of gameplay seemed to call for it, and under the presumption that it would add to intrinsic motivation. Score is calculated by multiplying the number of baubles in a match by nine, to ensure large numbers without the artificial inflation of trailing zeroes. A screenshot of the initial prototype UI can be found in Figure 2

Three important narrative game mechanics were included, as well. First, a mechanic linking gameplay to narrative: Whenever the player makes a valid move, the narrative advances to the next available line. Second, a mechanic linking gameplay decisions to narrative choice: The narrative can take branching paths depending on the type of match the player makes. For instance, when met with a narrative choice to go to either a bike rack or a parking lot, matching bikes or cars respectively will impact the flow of the narrative. Last, a mechanic which expands on this idea, using a pool of

prompts instead of a binary choice: The gameplay decision informs which prompt is taken from the pool when the player makes a move, which is repeated until the pool is exhausted. This particular mechanic will serve as the backbone of the hermeneutic cycles to be included later.

These three are examples of Game-to-Narrative (GtN) mechanics, exemplified by actions in the game influencing the narrative. The proposed methodology will also make use of Narrative-to-Game (NtG) mechanics, where events in the narrative may impact the game. However, no NtG mechanics have been implemented in this initial iteration. In order to house the GtN mechanics, a simple story was composed.

4.2 Evaluation Method and Goals

The primary goal of this base prototype's evaluation is to gauge the general enjoyability of the gameplay, the reception of the game's narrative and gameplay connections, and the clarity of game mechanics. To gather qualitative data on these subjects, focus groups were conducted, to which fellow Game and Media Technology students were invited. The technical and practical insight of those familiar with principles of game design and user experience was much appreciated and proved to be invaluable.

A total of six participants were interviewed, across three groups: One containing three participants, one containing two, and a one-on-one interview. The participants were sourced from university peers, consisting of young, educated individuals with experience in game and UI design. The participants were provided information about my thesis, and the purpose of this study within it. Informed consent was obtained prior to the start of the focus group, and the participants were made aware that the group discussions would be recorded.

The group started with each participant playing through the prototype demo. The game was built using WebGL technology and hosted on Itch.IO, meaning participants would be able to play the prototype in tandem on their own devices. Gameplay was surveyed, but as little external input as possible was given throughout, to avoid introducing bias into the gameplay experience. Afterwards, the groups discussed various talking points pertaining to the overall quality of the gameplay and narrative, the integration of game and narrative mechanics, and the intuitivity of gameplay. Various individual assets of the game, such as the traffic light indicator and the scoring system, were also discussed. Lastly, the participants were asked to reason about the importance of various hypothetical features, and aspects they felt were lacking in the prototype. After the focus group session, participants were offered candy and chocolate as a token of gratitude.

The recordings were later transcribed using the Office 356 Transcribe tool, and anonymised. After transcription, the original recordings were permanently deleted, with no known existing copies, as established in the informed consent documents. This is done both in compliance with the EU GDPR and the university Ethics and Privacy Scan. Those transcripts were then reviewed to group feedback on the various discussed elements throughout and common patterns and trends were identified in participant opinions.

4.3 Results and Discussion

The results from these focus groups are relevant to the study mainly from a position of perceived clarity of the proto-GtN mechanics, as mentioned in Section 4.2. For this reason, the emphasis of the analysis lies mainly on a player's interaction with the narrative.

General reception to the gameplay itself was broadly positive among the groups, though several important concerns arose pertaining to the various narrative mechanics. As I was the author of this game and these mechanics, their workings had seemed obvious to me, but for many of the

participants it took several turns before they realised their actions moved the narrative forwards. The initial implementation did not clarify this was the case, and because the participants had not extensively immersed themselves in the prototype as I had, the mechanic caused some confusion. This was most blatantly obvious with the first of the three mechanics, but a majority of participants also indicated being unaware that their action caused a specifically related narrative prompt to appear even after they realised their moves generated the prompts in the first place.

A second point of concern was the importance of the narrative. The layout of the initial prototype, with a central grid and textbox underneath, indicated to the participants that gameplay should be seen as more important than narrative. One participant indicated they perceived the text as complementary to the gameplay - a 'well-done!' sticker for making good matches - and paid little attention to it. The inclusion of the score, while indeed driving motivation, had the unintended effect of making players play for speed and score, which caused a further dismissal of the story. This genre of game was specifically chosen to allow players to take their time and choose their own pace, so they can easily consume the narrative even during gameplay. The inclusion of a scoring system undermined this thought entirely.

There were a few minor concerns regarding the combination of narrative and gameplay within the same space, and a point was made that it might be hard for players to have to switch back and forth between focusing on the narrative and focusing on making moves. However, after further discussion this concern was deemed not as problematic, if the player was aware they are the ones deciding the pace. A suggestion to have the narrative pop up whenever a new segment is triggered was considered, but given the frequency of pop-ups this would imply, it was deemed to be too distracting. Important non-narrative gameplay explanations would likely still benefit from this approach, as these would be used more sparingly and are more important to the player.

Lastly, the focus groups considered two hypothetical features, which I anticipated beforehand but had not yet implemented due to resource constraints. The first is a feature to access dialogue backlog throughout the course of the game. This was a feature brought up by various participants independently, even before my prompting, which indicates the importance thereof. In terms of accessibility, and especially if the narrative is central to the proposed methodology, it should be easy for the players to check back sections they may have missed, or incidentally skipped. The second is a feature to indicate when the player's moves set up a narrative decision. Even with narrative cues, participants indicated it was not obvious when a move correlated to a choice, or they did not notice narrative dialogue corresponding to their prior actions. Additional clarity towards the narrative mechanics would help in this regard, but an explicit UI element to signal when choices can be made and what baubles are available for this was generally agreed to be an important addition as well.

4.4 Iteration

Using the feedback from this study, I iterated over the base prototype and improved on the given concerns. The form of the UI was altered to make the textbox bigger and place it at the top margin. This places more importance on the narrative and makes it harder to ignore. The score feature was entirely removed, as it proved too distracting, and some minor graphic changes were implemented to make the various UI elements more intuitive.

A record of preceding dialogue was added to the game; a button in the corner of the textbox now opens up an overlay, where the user can scroll through previously seen dialogue. This dialogue is added in real-time, to take the branching paths and player decisions into account. Additionally, if a certain dialogue branch was accessed through an explicit bauble match, the colour of the bauble in question is included in the records.

An additional UI element was included in the form of a post-it note to indicate when choices are



Figure 3: Improved UI layout

available. This note is pinned onto the textbox to indicate relatedness to the narrative. By default, it instructs the player to simply match any bauble. In situations where GtN binary choices or pool choices can be made, however, the post-it shows the various options the player has. In the case of binary choice, it displays the two baubles which will lead to different paths. In the case of a choice pool, it will list all possible matches for which there is yet unexhausted dialogue. Once the pool is exhausted, the note returns to its default display.

The improved UI can be found in Figure 3.

5 Final Prototype

5.1 Narrative and Serious Content

As previously stated, the serious content used in this research will be kept small and simple in scope, centered around the concept of 'induced demand' as described in Section 3.2.

The story written around the concept is set up in two acts. The first simply explores what induced demand means, and how it works. During this first act, the story directly references the 'triple convergence' principle coined by Anthony Downs [11]. This principle outlines the three main means of convergence, from which demand is induced:

- **Spatial convergence** refers to travellers who used to take a different route, now converging to the renovated one
- **Temporal convergence** refers to travellers who used to avoid peak hours, now converging to those times after the renovation

- **Modal convergence** refers to travellers who used to take a different mode of transport, encouraged to switch modes by the renovation

The second act elaborates on some of the nuances surrounding induced demand in concrete, real-world situations, while additionally responding to some criticisms of the concept. This act touches on how 'induced demand' is not limited to car traffic, but also extends to other modes such as cycling and public transit. Some examples are given on how traffic congestion can be reduced by inducing demand for other modes, as well as reversing induced demand for car traffic.

The narrative is written in a CSV format, where information such as the name of the speaker and the image to render are recorded. It is written from the perspective of a high-school excursion to a traffic museum, with much of the content integrated as a dialogue between three students, Lucas, Susan, and the player, and their teacher, Mister Lane. During the second act, an antagonist in the form of Oliver Garch and his crony, Bob, show up, to offer counter-arguments for Mister Lane to rebut.

Due to the different ways in which the narrative is conveyed to the player between the two prototypes, the narrative will not be identical across them. However, the narrative structure and content outlined here is consistent across both.

5.2 Narrative Format Implementation

In order to create an interactive narrative, which could facilitate both GtN and NtG mechanics, the CSV format proved insufficient and had to be expanded. In simplest terms, a control flow of sorts had to be established, in order to offer branching paths dependent on in-game actions. While various solutions for this exist already, such as the Yarn Spinner tool for Unity, the extended nature of the mechanics I strove to implement made it difficult to properly adapt these packages. As such, I opted to outfit the prototype narrative manager with a rudimentary parsing program.

Due to the CSV formatting, certain meta-commands can be embedded in the first column of each line, which define how the line should be parsed. For instance, a line of narrative dialogue is preceded by the 'TEXT' flag, while a conditional branch would be preceded with 'OPTION'. The parser algorithm resolves these meta-commands and ensures the corresponding actions are taken, depending on both the extended arguments of the CSV line and the player-matched bauble. A full overview of these meta-commands, their arguments, and their effects, can be found in Table 1

In terms of code, the CSV file is read and instantiated as a `CSVAsset` class, which is a structure keeping track of its content, line pointer, recursive files, and other information pertaining to the various meta-commands. When a player-made move prompts a new line, the parser requests the next line of content from the `CSVAsset`, and depending on the nature of the meta-command either displays the text or performs a set of actions before requesting the next line. The system as described in this section contains the exhaustive set of commands required for the GtN and NtG mechanics the prototype makes use of; important to note is that prior versions of the prototype contained fewer commands, and similarly the baseline version described in Section 5.4 has many commands omitted as the GtN and NtG mechanics they facilitate should not be present in this variant.

While the system does have foundations in contemporary parsing strategies, it ultimately is somewhat of a patchwork solution: It certainly works, but with obvious limitations. Any expansion of the GtN or NtG mechanics would require some form of restructuring or adding more meta-commands, and in particular the 'GOTO' command - the backbone of this parser's control flow - makes the dialogue files more difficult to maintain. For any commercial product, this would hardly be an acceptable solution, however for the purposes of my prototype the main importance lies in providing a functional bite-sized experience without excessive production overhead. Under these prerequisites, the system works as intended.

Command	Argument(s)	Description
START		Denotes start of file.
END		Denotes end of file.
TEXT	name, content, portrait, modus	Describes text content to be displayed.
BLANK		The BLANK command is skipped by the interpreter.
GOTO	line	Indicates the line pointer to go to the specified line.
LOAD	source	Recursively loads a new file, identified by the source path.
OPTION	bauble, follow-up command, follow-up details	Facilitates binary choice. Different choice options (arguments) are placed on consecutive newlines. Follow-up commands are limited to GOTO and LOAD.
RANGE	size, range name, bauble, line	Facilitates a range of content to be visited in any order. Range option baubles and goto lines are placed on consecutive newlines, up to the range size.
ENDRANGE	line	Denotes the end of the range segment. The line argument indicates where the line pointer will continue.
SETMATCH	bauble, amount	Sets the number of baubles needed for a match.
RESETMATCH	bauble	Resets the number of baubles needed for a match to the default value of 3.
INCREASEODDS	bauble	Increases the odds of the given bauble spawning.
DECREASEODDS	bauble	Decreases the odds of the given bauble spawning.

Table 1: meta-commands and their arguments

5.3 Variant I

After the initial prototype was refined, and the narrative structure and serious content within was outlined, I developed the two variants of the prototype that will be used to evaluate the proposed methodology.

Variant I will expand on the base prototype by implementing a serious narrative that the player will be able to explore through both NtG and GtN mechanics. Together, these encapsulate the hermeneutic cycles of learning, and by integrating a serious narrative, form Narrative Serious Game Mechanics.

The NtG mechanics relevant to this variant were already present in the base prototype and iteration. Given the novel nature of these mechanics, which allow the player to control the narrative through in-game moves, it was imperative to gauge the effectiveness of these mechanics in a vacuum - that is, without serious content included. And indeed, the iteration improved many aspects of these mechanics to make them more intuitive to the player. As previously stated in Section 4, NtG mechanics allow actions within the game to influence the course of the narrative. This includes setting the player on different paths, depending on a choice of baubles matched, and offering the player one of various options depending on this choice.

The GtN mechanics added allow the narrative to change certain rules of the game in order to emphasise important narrative content. In a match-three bauble game such as this one, two important rules that affect gameplay are the rate at which baubles spawn, and the number of baubles needed to form a match. By default, every bauble has an equal chance of spawning, and every bauble needs a row or column of at least three. By integrating these GtN mechanics, both this spawn rate and match amount can be tweaked by narrative paths. This allows the narrative to simulate, for instance, traffic congestion by increasing both the spawn rate and match amount for cars. See Figure 4 for an example of how the board would look.

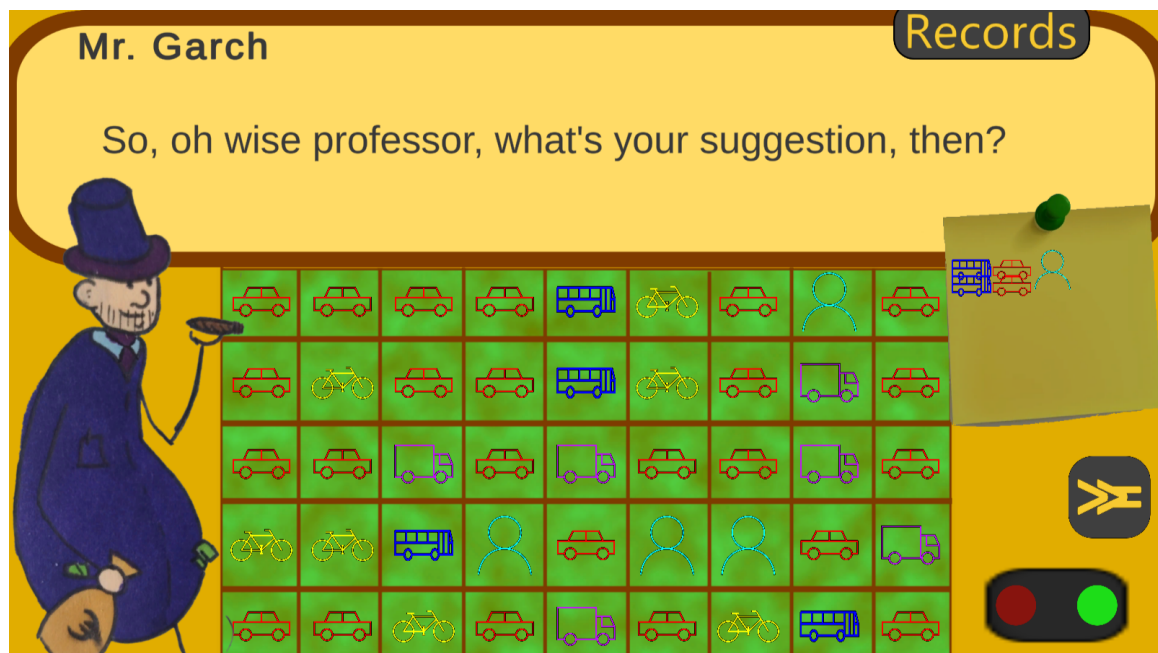


Figure 4: A board congested with cars

Both NtG and GtN mechanics are embedded in the narrative itself through the implementation described in Section 5.2.

5.4 Variant II

The second variant takes a more grounded approach to the formula. Many contemporary games in this genre alternate gameplay and narrative interactions, by introducing segments of the story both before and after each level. Variant II will modulate this method.

Deriving variant II from the base framework required some minor pruning and restructuring. As stated, the base did contain a handful of GtN mechanics, which would be obsolete in this variant. Additionally, the UI of the base framework contained various elements redundant to this variant. The textbox, the post-it note, and the extra room for character portraits are all relevant to the narrative, which in this variant does not coincide with gameplay. By removing these, the scale of the playing field could be increased.

Since in the base framework, a level is delineated by the narrative, and narrative is decoupled from gameplay here, a new method of deciding when a level is finished was needed. For this, the score mechanic originally present in the base framework was brought back. As there is no narrative to speak of during the gameplay sections, the risk of score distracting from story is no longer present, thus it is safe to reintroduce. Aside from the player's current score, a 'score goal' was added to the UI. When the player score reaches the score goal, the player has completed the level. An overview of the modified UI - both during gameplay and during narrative - can be found in Figure 5.

However, the serious content of this version is still embedded in the narrative, thus the narrative must still be a part of the game. A narrative overlay was implemented, containing the text box, the records access, and room for two character portraits.

To manage levels and dialogue, a simple level class was created. The class houses a reference to the pre-level dialogue, the score goal, and the post-level dialogue. Using this class, the level handler can easily delegate these parameters to the scoring system and dialogue system respectively, activating each when relevant. The narrative itself is still contained in CSV file format, though a much-simplified variant. Since both GtN and NtG mechanics are irrelevant in this version, all meta-commands pertaining to these have been removed. Thus, only the START, END, TEXT and BLANK commands remain. However, since the narrative overlay now contains two character portraits instead of one, the parameters of the TEXT command were extended to contain both the left and right portrait.



(a) Gameplay section



(b) Narrative section

Figure 5: Prototype II UI

6 User study

In order to draw conclusions on the efficacy of the combined methodology, a comparative user study is required. The study takes the form of a simply A-B test, in which participants were exposed to either one of the prototype variants and asked to report on their experiences playing the game through a questionnaire. After both variants were in a satisfactory state, both were uploaded to a respective Itch.IO page, and made available to play in the browser through WebGL. This will allow both variants to be easily accessible to study participants, without the need to download any data or receive a local copy, and thus lower the barrier of entry. To manage participants playing the games and record their reported experiences, a questionnaire was created in Qualtrics.

6.1 Questionnaire

The questionnaire was split up into three broad sections. First it introduced the basics of the study and provides contact information for relevant questions and concerns, before offering details regarding informed consent. Only after informed consent is obtained does the questionnaire continue.

Demographic information, such as age, gender, and level of education was collected, and the participant was asked to rate their pre-existing knowledge of the serious topic, 'Induced Demand', in three distinct statements. These are deliberately similar, but not identical, to ensure a reliable measure:

- I already know what 'Induced Demand' is
- I am already aware of how 'Induced Demand' impacts my surroundings
- 'Induced Demand' makes intuitive sense to me

Each of these statements are rated on a seven-point Likert scale.

After establishing prior knowledge, the participant would receive a hyperlink for the prototype game they will be playing. Every participant played either version I or version II, never both. The Qualtrics randomiser ensures either element is presented evenly. After completing the prototype, the player would receive a verification code, which when entered in the survey, allows them to continue.

The final part of the questionnaire concerns the collection of various quantitative data required for this study, pertaining to both the player's experience at a dynamic level, their experience at an aesthetic level, and the observed learning gains after playing the prototype. For the two former, the Player Experience Inventory (PXI) was consulted [1], and three measurement constructs from both the functional consequences and the psychosocial consequences respectively were selected to be most relevant to this study, listed below.

- **Functional consequences**

- Ease of control: Given version I introduces a new type of game-mechanic, it is important to evaluate how the player manages to control this.
- Goals and rules: Whether or not the inclusion of the game-mechanic has any drastic influence on the players' perception of what is expected from them in the game is similarly important
- Challenge: Neither version is made to be 'difficult' per se, yet the subtle differences between them, and certain mechanics such as congestion, may have implications for the game's challenge rating which might be interesting to observe.

- **Psychosocial consequences**

- Meaning: Both versions seek to teach the player serious content; how meaningful a game is to a participant can be indicative of the extent to which this content will be retained, and thus is relevant for the comparison.
- Curiosity: While version II is entirely linear, version I contains various methods for the player to discover content on their own. This metric will be important to gauge to what extent that discovery process boosts a player’s curiosity.
- Autonomy: Version I contains various mechanics to substitute linear story progression with a certain player influence over actions and directions. However, whether this autonomy is perceived as such, and in fact an improvement over the perceived autonomy of version II, will be important to measure effectiveness.

The PXI model contains three statements per construct, again to ensure a reliable measure. And again, are rated on a seven-point Likert scale, as expressed in the PXI bench user guide [2].

The observed learning gains will be measured using the same three statements used to gauge participant familiarity prior to playing, with two additional ones being added. These two query how the participant feels about their improvement in terms of understanding and interest, directly:

- I know what 'Induced Demand' is
- I am aware of how 'Induced Demand' impacts my surroundings
- 'Induced Demand' makes intuitive sense to me
- My understanding of 'Induced Demand' has increased
- My interest in 'Induced Demand' has increased

Finally, the participant is asked to briefly explain their understanding of the concept in an open textbox. This data differs from the rest of the study in being qualitative as opposed to quantitative, but this is an intentional distinction. Rather than using the qualitative data directly during the analysis section, it will instead be used to validate the self-reported quantitative data pertaining to learning gains. If for a certain data-point the participant expresses to be fairly knowledgeable in the subject, but their explanation indicates otherwise, this shows the data-point to be faulty, and it can be appropriately handled or discarded. The same is true the other way around.

6.2 Data Collection

The questionnaire was initially distributed among peers at the university. However, after some time saw an insufficient number of participants, the targeted demographic was increased in scope to include advertisements on LinkedIn. The use of survey-sharing sites such as SurveyCircle and SurveySwap was considered, but efforts made in this direction ultimately proved fruitless. Since the sites rewards participants for every survey made, it encourages more concise surveys – and while my survey in and of itself does not contain particularly many questions, the additional gameplay it requires does add more time which would dissuade many. In addition, the point-system these sites use would require completing a large number of other surveys in order to have my own listed even among the top-100. As a result, no participants were sourced through these websites.

Another issue which arose partway through was the randomisation provided by Qualtrics. Over time, it introduced a bias in the intermediate results: Despite the number of participants who received the version I and version II prototypes being roughly the same throughout the survey’s runtime, participants who played version I reported a significantly higher set of knowledge prior to the study, and participants who played version II reported significantly lower prior knowledge.

Since one of the metrics used to evaluate results is learning gains, or the difference between prior and posterior knowledge, having a similar (prior) baseline is important to accurately gauge a difference of differences.

To alleviate this bias, an additional clause was introduced. An embedded data item was used to average results of the three prior knowledge questions, and a conditional branch was established using this. Initially, this branch directed participants with higher prior knowledge to the version II prototype and participants with lower prior knowledge to version I, to average out the accumulated bias. After the record straightened, each branch was equipped with its own randomiser module, to ensure an even split between both higher and lower average prior knowledge.

6.3 Participants

Out of the twenty participants in this study, an average age of 27.95 was calculated. Ages broadly ranged from 22 to 30, with two outliers over 50. 70% of participants identified as male, 20% as female, and the remaining 10% non-binary. The vast majority of participants reported the highest level of education completed to be the bachelor level or higher, as shown in Figure 6, with only three participants reporting having just obtained a high-school degree. None reported no formal education obtained.

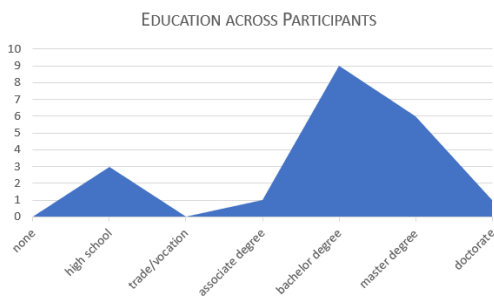


Figure 6: Reported education levels across participants

6.4 Analysis

After sufficient time had passed, the survey closed with a round twenty completed entries, split evenly among the two groups. All data was exported from Qualtrics and introduced in Excel, where it was categorised by group and refined. For each participant, raw data contained three responses for each PXI construct, rated on the same scale, which were averaged together into a singular value as expressed by the PXI bench user guide [2]. Similarly, the three statements querying participant knowledge were averaged together as well for both the prior and the posterior case. Lastly, the scale values were shifted over from the [0,7] range, as Qualtrics exported it, to [-3, 3]. This is both per the PXI bench user guide [2], and conveys more clearly the neutrality of the middle value. In total, the raw data was refined into ten data series, as described in Table 2.

Several graphs were generated from this refined data using averages over the two groups, and data was converted to long format to facilitate further analysis in JASP. This includes an independent-samples T-test over LE-2.2, LE-2.3, and the PXI measures, as well as a repeated measures analysis of variance (RM-ANOVA) over the LE-1 and LE-2.1 series as before and after factors. The results of these analyses are shown in Figures 7, 8, and 9, and Tables 3, 4, and 5.

Figures 7 and 8 are both derived directly from the refined Excel data, showing the data series averages for each version, with error bars indicating standard deviation. The data series described are the ones listed in Table 2, with the exception of LE-D in Figure 7, which instead describes the difference between LE-1 and LE-2.1.

Data series label	Elaboration
LE-1	Average of prior knowledge statements
LE-2_1	Average of posterior knowledge statements
LE-2_2	Self-reported increase of subject understanding
LE-2_3	Self-reported increase of subject interest
PXI-1.1	Functional PXI consequence; Average of ease of control statements
PXI-1.2	Functional PXI consequence; Average of goals and rules statements
PXI-1.3	Functional PXI consequence; Average of challenge statements
PXI-2.1	Psychosocial PXI consequence; Average of meaning statements
PXI-2.2	Psychosocial PXI consequence; Average of curiosity statements
PXI-2.3	Psychosocial PXI consequence; Average of autonomy statements

Table 2: Overview of data series labels

Table 3 describes results of the independent samples T-Test, performed in JASP. This test was performed over the six player experience data series, as well as the increase of understanding and increase of interest series.

Figure 9 Table 4 describe the RM-ANOVA results, which was performed to compare the effect of prototype variant on learning gains, where learning gains are described by LE-1 and LE-2.1 as knowledge prior to gameplay and after gameplay respectively. Additionally, Table 5 contains the post-hoc test results.

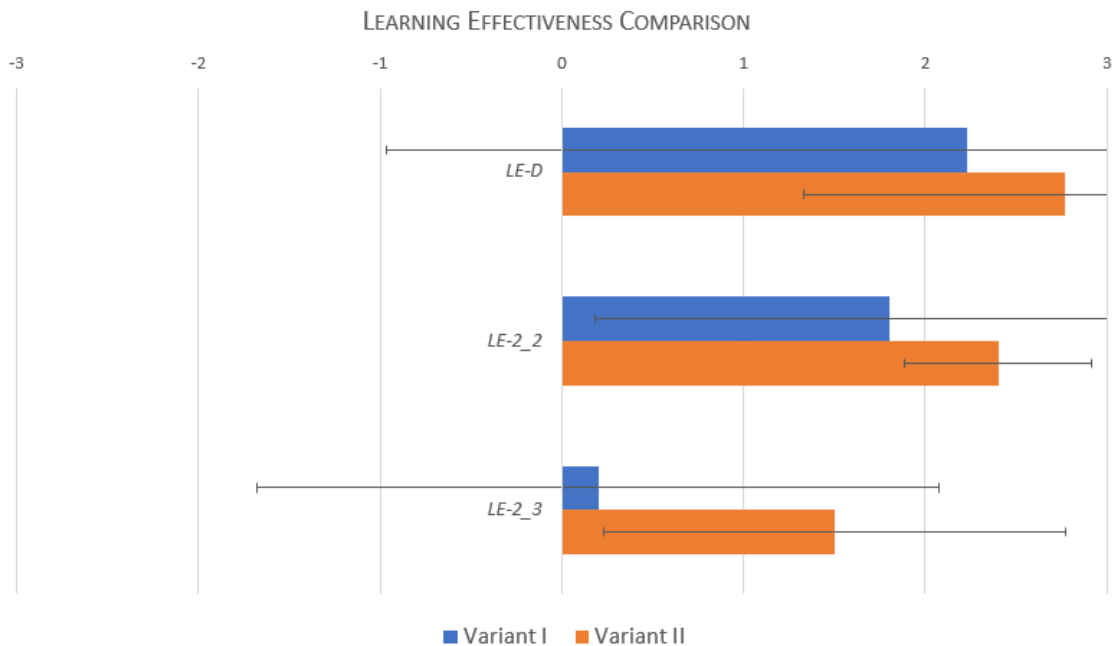


Figure 7: Learning effectiveness metrics between groups

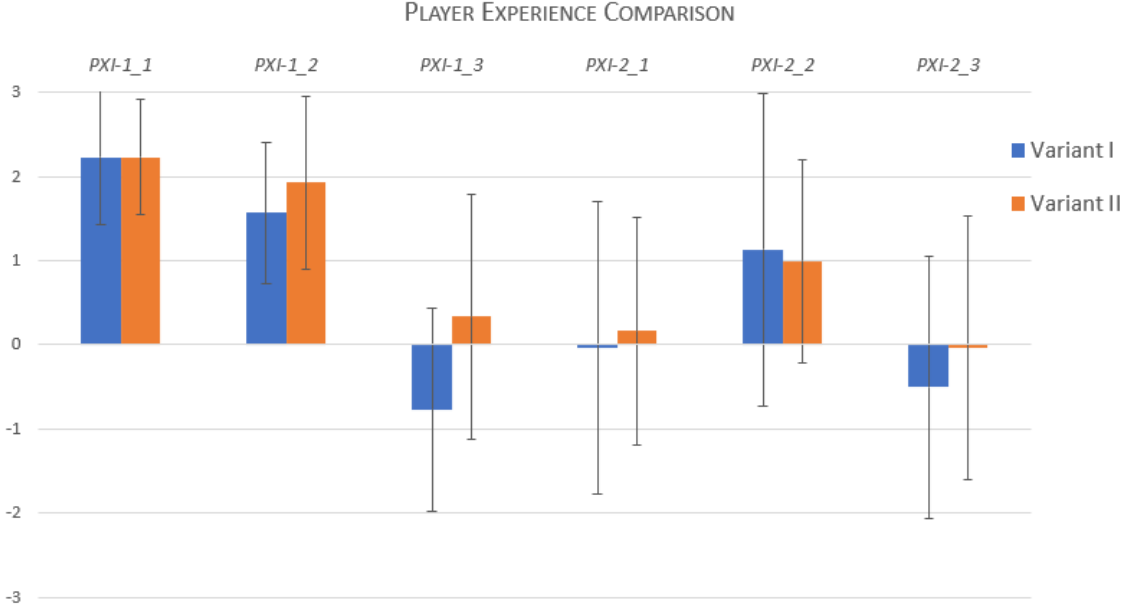


Figure 8: PXI metrics between groups

6.5 Results

6.5.1 T-Test

As described in Section 6.4, an independent-samples T-Test was performed over eight data series, the results found in Table 3.

In terms of learning behaviour, there was no significant result for either increased subject understanding, $t(18) = -1.1, p = .289$, or increased subject interest, $t(18) = -1.8, p = .086$, following the conventional $p = .05$ threshold. However, the comparison of increased subject interest does imply version II ($M = 1.5, SD = 1.3$) to perform better than version I ($M = 0.2, SD = 1.9$), as visible in Figure 7, LE-2.3.

In terms of functional player experience consequences, there was no significant result for either ease of control, $t(18) = -3.0 \times 10^{-10}, p = 1.0$, or clarity of goals and rules, $t(18) = -0.88, p = .392$. For perceived challenge, $t(18) = -1.84, p = .083$, the numbers imply version II ($M = 0.3, SD = 1.5$) to be more challenging than version I ($M = -0.8, SD = 1.2$), but again significance is limited. Figure 8 illustrates the similarity of mean and standard deviation across these groups and series.

In terms of psychosocial player experience consequences, there were no significant results for either meaning, $t(18) = -0.29, p = .777$, curiosity, $t(18) = 0.19, p = .851$, or autonomy, $t(18) =$

	t	df	p
LE-2.2	-1.116	18	0.279
LE-2.3	-1.816	18	0.086
PXI-1.1	-2.997×10^{-10}	18	1.000
PXI-1.2	-0.877	18	0.392
PXI-1.3	-1.838	18	0.083
PXI-2.1	-0.287	18	0.777
PXI-2.2	0.190	18	0.851
PXI-2.3	-0.668	18	0.513

Table 3: Independent Samples T-Test

$-0.67, p = .513$. Again, the compared mean and standard deviation can be found in Figure 8

Considering the T-Test results as a whole, it becomes apparent that results show high p-values across the board. None of the dependent variables show a value below the conventional $p = .05$ threshold, and additionally only two of the eight variables exhibit $p < .1$. The low significance is corroborated by the graphs of averages, where again only series LE-2.3 and PXI-1.3 show noteworthy difference between the version I group and version II group. Figure 8 shows that in terms of average challenge, version I was rated to be not so much challenging, whereas version II was rated to be somewhat challenging. And in Figure 7, it shows that version II was better at increasing subject interest than version I, though version I had not received a negative average score here. While these differences are interesting to consider, it is still important to realise that neither of these series' results are by conventional standards considered to be statistically significant. Thus, the data broadly shows no evidence to conclude these two groups are not the same.

6.5.2 RM-ANOVA

Cases	Sum of Squares	df	Mean Square	F	p
Time	62.500	1	62.500	20.327	< .001
Time * Group	0.711	1	0.711	0.231	0.636
Residuals	55.344	18	3.075		

(a) Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p
Group	0.178	1	0.178	0.106	0.748
Residuals	30.100	18	1.672		

(b) Between Subjects Effects

Table 4: ANOVA comparing difference in content knowledge before and after gameplay, across the two variant groups

The RM-ANOVA test, as mentioned prior, was performed to compare the effect of prototype variant on learning gains, through the LE-1 (prior) and LE-2.1 (posterior) reported knowledge series of data. Similar to the T-Test result, the RM-ANOVA results are broadly inconclusive. There was no statistically significant difference in learning gains between the two groups, $F(1, 1) = 0.23, p = .636$.

Despite the low significance, post-hoc tests using Bonferroni correction were carried out, shown in Table 5. Between both the before-before (MD = 0.40, SE = 0.69) and after-after (MD = -0.13, SE = 0.69) comparisons, showing respectfully $t = 0.58, p = 1.0$ and $t = -0.19, p = 1.0$, there is no statistically significant difference. Only when between-time effects are taken into consideration do we report higher significances. This is in line with the within subjects effects shown in Table 4a, where for Time we see $F(1, 18) = 20.33, p < .001$. Thus, while Group or Time * Group have no significantly obvious bearing on results, Time itself does.

This again is illustrated well in Figure 9, displaying the reported knowledge both before and after play, across the two groups of prototype variants. Both groups prior to gameplay report an average of around -0.3 to -0.8 on the Likert scale, whereas after gameplay both groups report an average around 2.0. While this result has no immediate implications to the study itself, it is

		Mean Difference	SE	t	<i>p</i> _{bonf}
Group1, Before	Group2, Before	0.400	0.689	0.581	1.000
	Group1, After	-2.233	0.784	-2.848	0.064
	Group2, After	-2.367	0.689	-3.435	0.010
Group2, Before	Group1, After	-2.633	0.689	-3.822	0.003
	Group2, After	-2.767	0.784	-3.528	0.014
Group1, After	Group2, After	-0.133	0.689	-0.194	1.000

Table 5: Post Hoc Comparisons - Group * Time

an interesting result regardless as it shows that both prototypes were effective at improving the participant’s knowledge.

6.6 Discussion

Based on the results outlined in the previous section, no conclusive answer to the research question can be provided. Importantly, the lack of significance does not imply that one particular variant performed substantially better than the other. However, this absence of evidence should not be considered as evidence of absence. The user study was unable to verify the increased merit of the combined methodology proposed, yet this does not immediately imply the methodology’s deficiency. Rather, let us consider the implications of the most striking of these results.

Regarding the implications of learning effectiveness results, the within-subjects RM-ANOVA analysis does indicate that both prototype variants served to improve the participants’ knowledge. From the gathered evidence it is impossible to state whether one improved this better or worse than the other, but the fact remains that, at the very least, both versions were able to facilitate learning.

Additionally, the implications of player experience results on the combined methodology of prototype variant I are particularly interesting, considering the baseline variant compared against. While it is true the results show no significant improvement of variant I over variant II, importantly the results show no significant deterioration either. Similarities between reported ease of control or clarity of goals and rules do seem to imply that, in terms of gameplay, the two prototype variants facilitate similar experience.

As such it can be reasoned that the approach including these GtN and NtG mechanics offers both similar learning behaviours and similar gameplay experiences. The main concern then becomes whether being similar is, in fact, enough. As became apparent during both the literature review and the prototype implementation and iteration, fully interactive digital narratives are time-consuming and complicated to develop. The question of why conventional methods appear to under-emphasize narrative might be attributed to the perception that, despite its well-documented benefits, the trade-off with the inherent complexities of narrative may not be deemed worthwhile.

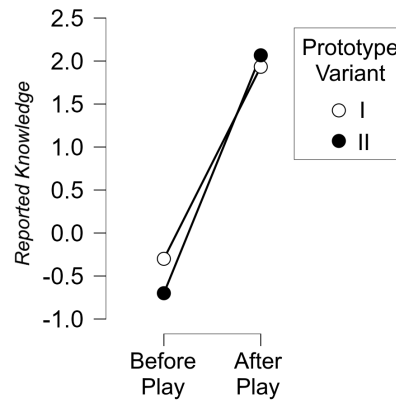


Figure 9: RM-ANOVA Descriptives plot

7 Conclusion

Throughout this thesis, I have attempted to develop a new methodology for narrative serious game design, which sought to combine various existing techniques to improve learning behaviour through stronger, more compelling, narrative engagement, and the idea that hermeneutic theory would allow narrative to dynamically match the different learning rates of different individuals. A prototype of an infinite-type game conducive to the cyclical hermeneutic theory was created, and with help of focus groups refined. However, in evaluating whether this combined methodology has any merit when compared to conventional narrative approaches, the survey study failed to produce relevant significant results.

While the study did imply the prototype variant making use of the combined narrative techniques facilitated learning to a similar degree as the baseline variant, and while in terms of gameplay experience both prototype variants exhibited similar properties – with the notable exception of perceived difficulty – this on its own is not enough to conclude the merit of the combined method.

In game design and development, trade-off and compromise are extremely important: It is easy to overburden the creative process with extraneous features that would or could be fun or engaging, but ultimately are not worth the time and resources required to make them be. Interactive digital narrative containing pedagogical elements, so it seems, is one such feature. This study failed to prove such narratives using the combined method offer a significantly appropriate improvement on player experience or learning gains, to warrant implementation. Indeed, in direct comparison they do not seem to compare worse to conventional methods, but if the additional work required is included in the consideration, 'just as good' is not good enough.

8 Limitations and Future Work

The limitations of this study become quickly apparent when taking another critical look at meta-information surrounding the survey results. In total, a mere twenty participants submitted to the survey, with each version being played by ten participants. This is less than half of the initially desired turnout, and results in a particularly small trial group. Additionally, due to an initial low turnout rate, the study's target audience was broadened during the procedure, which serves to further muddle the accuracy of results.

This lower number of participants is likely to be a combination of factors. For starters, as a master thesis project, this study does not have access to extensive funding in order to compensate participants or make use of services such as Amazon MTurk. As such, all participants had to be sourced from those willing to spend their free time working on a survey – which introduces another factor, duration.

The survey itself is not particularly in-depth or intricate, but it does require the participant to spend a certain amount of time playing the prototype game. During development, a goal was set to keep this mandatory playtime below twenty minutes, yet even with this goal met, it still results in an average survey duration of half an hour. This time-requirement is similarly likely to dissuade many from participating, and indeed I observed many incomplete survey submissions that were abandoned halfway through. Based on the portions completed and left incomplete, it seems likely that these participants reached the gameplay section, and here decided the study was taking too long.

However, regrettably, with a study like this it is near impossible to reduce the time required. In order to gauge the learning effectiveness of a digital game, it is merely a fact that this game's content needs to be experienced. Shorter games will likely not be able to contain as much in terms of mechanical depth or information, and so we are at an impasse where either the duration of study or depth of the game must take priority.

That said, various psychological methods may be exploitable in order to speed up the perceived time playing. Players entering a state of flow, for instance, may be more likely to forget their playtime and surroundings. A sufficiently alluring game could, following this line of reasoning, entice players to spend more time playing without realising it. In terms of duration of study, this may be one such method to increase mechanical depth and increase playtime without the study at large feeling like it costs too much time or effort.

Another limitation to discuss is the state of the game prototype itself. Despite the simple bauble-matching mechanic of the base prototype, the game does contain various narrative mechanics which, throughout the creation and testing processes, proved to be less than intuitive. Facilitating narrative player choice through matching baubles is certainly an interesting take – but whether it is mechanically optimal is a different question. Ultimately, prototype variant I introduced several unconventional mechanics, the implementation of which may very well have distracted from gameplay and learning. Perhaps, modifying tried and tested gameplay mechanics to include hermeneutic engagement would see more positive effects, as opposed to wrapping unconventional hermeneutical theory in unconventional gameplay loops.

As mentioned in the conclusion, further study would be required in order to validate the effectiveness of the proposed methodology. Such further study would benefit both from an increased sample size, as well as an improved prototype game. Despite the focus group surrounding the initial prototype's efficacy, gameplay and mechanics still left much to be desired should this game compete with market experiences.

Additionally, this study focused primarily on comparing the proposed methodology version to a version containing more conventional methods of conveying narrative. Further study might also be interested in considering the implications of the proposed methodology itself in further depth. Relevant information, such as how the target audience recognises and utilizes the various hermeneutic cycles available to them, ultimately was not considered in this study, yet might provide additional insight towards how best to make use of the methodology.

References

- [1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies*, 135:102370, 2020.
- [2] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, Kathrin Gerling, and Aqeel Haider. Pxi bench user guide.
- [3] Jonne Arjoranta. How are games interpreted? hermeneutics for game studies. *Game Studies*, 22, 08 2022.
- [4] Karina Arrambide. Interactive narratives in games: Understanding player agency and experience. In *CHI PLAY '19 Extended Abstracts*, CHI PLAY '19 Extended Abstracts, page 1–5, New York, NY, USA, 2019. Association for Computing Machinery.
- [5] Pratama Wirya Atmaja and Sugiarto. When information, narrative, and interactivity join forces: Designing and co-designing interactive digital narratives for complex issues. In *Interactive Storytelling: 15th International Conference on Interactive Digital Storytelling, ICIDS 2022, Santa Cruz, CA, USA, December 4–7, 2022, Proceedings*, page 329–351, Berlin, Heidelberg, 2022. Springer-Verlag.
- [6] André Barbosa, Pedro Pereira, João Dias, and Frutuoso Silva. A new methodology of design and development of serious games. *International Journal of Computer Games Technology*, 2014:1–8, 06 2014.
- [7] BioWare. Mass effect, 2007.
- [8] Two Tribes BV and Soedesco. Toki tori 2, 2013.
- [9] Robert Cervero. Induced demand: An urban metropolitan perspective. 2001.
- [10] Zoe Delahunty-Light. Remember skyrim’s radiant ai? it’s got the potential to revolutionise rpgs, 2018.
- [11] Anthony Downs. *Stuck in Traffic: Coping with Peak-hour Traffic Congestion*. Brookings Institution Press, 1992.
- [12] Teun Dubbelman. Narrative game mechanics. In *Interactive Storytelling*, pages 39–50, 11 2016.
- [13] Malte Elson, Johannes Breuer, James Ivory, and Thorsten Quandt. More than stories with buttons: Narrative, mechanics, and context as determinants of player experience in digital games. *Journal of Communication*, 64:521–542, 06 2014.
- [14] Alan Fisher. Alan fisher.
- [15] PopCap Games. Bejeweled, 2001.
- [16] MediaSpark Inc. Goventure anybusiness.
- [17] Henry Jenkins. Game design as narrative architecture. *Computer*, 44, 01 2002.
- [18] Jesper Juul. Games telling stories? *Game Studies*, 1, 07 2001.

- [19] Antonia Kampa, Susanne Haake, and Paolo Burelli. *Storytelling in Serious Games*, volume 9970, pages 521–539. Springer, 10 2016.
- [20] Hartmut Koenitz, Gabriele Ferri, Mads Haahr, Digdem Sezen, and Tonguc Sezen. *Interactive Digital Narrative: History, Theory and Practice*. routledge, 04 2015.
- [21] Hartmut Koenitz, Mads Haahr, Gabriele Ferri, and Tonguc Sezen. *First Steps Towards a Unified Theory for Interactive Digital Narrative*, pages 20–35. Springer, 04 2014.
- [22] Jennifer A. Kotler, Jennifer M. Schiffman, and Katherine G. Hanson. The influence of media characters on children’s food choices. *Journal of Health Communication*, 17(8):886–898, 2012.
- [23] Theodore Lim, Sandy Louchart, Neil Suttie, Jannicke Hauge, Ioana Stefan, Ivan Martinez-Ortiz, Pablo Moreno Ger, Francesco Bellotti, Maira Carvalho, Jeffrey Earp, Michela Ott, Sylvester Arnab, and Riccardo Berta. Narrative serious game mechanics (nsgm) – insights into the narrative-pedagogical mechanism. In *Games for Training, Education, Health and Sports*, pages 23–34. Springer, 04 2014.
- [24] Tomasz Z. Majkowski. Feeling good about myself : real-time hermeneutics and its consequences. 2017.
- [25] Rudy McDaniel, Stephen M. Fiore, and Denise M. Nicholson. Serious storytelling: Narrative considerations for serious games researchers and developers. In *Serious Game Design and Development: Technologies for Training and Learning*. Information Science Reference, 2010.
- [26] Andreea Molnar and Patty Kostkova. On effective integration of educational content in serious games: Text vs. game mechanics. In *2013 IEEE 13th International Conference on Advanced Learning Technologies*, pages 299–303, 2013.
- [27] Emily Naul and Min Liu. Why story matters: A review of narrative in serious games. *Journal of Educational Computing Research*, 58:073563311985990, 07 2019.
- [28] Nintendo. Super mario bros., 1985.
- [29] Louise Parker and Mark Lepper. Effects of fantasy contexts on children’s learning and motivation: Making learning more fun. *Journal of personality and social psychology*, 62:625–33, 05 1992.
- [30] Aleksej Pazjtnov. Tetris, 1984.
- [31] Katelyn Procci, Clint Bowers, Christopher Wong, and Anya Andrews. Minigames for mental health: Improving warfighters’ coping skills and awareness of mental health resources. *Games for Health Journal*, 2(4):240–246, 2013.
- [32] Carlton Reeve. *Narrative-Based Serious Games*, pages 73–89. Springer, 01 2009.
- [33] Helen Routledge. Narrative driven instructional design: 3 ways serious games produce compelling learning, 2017.
- [34] Jason Slaughter. Not just bikes.
- [35] Jeff Speck. *Walkable city; How Downtown Can Save America, One Step at a Time*. North Point Press, 2013.

- [36] Rockstar Studios. Red dead redemption 2, 2018.
- [37] Beat Suter, René Bauer, and Mela Kocher. *Narrative Mechanics: Strategies and Meanings in Games and Real Life*, pages 19–50. transcript Verlag, Bielefeld, 2021.
- [38] Henry A. Virkler. *Hermeneutics: Principles and Processes of Biblical Interpretation*. Baker Book House, 1981.

9 Appendix A : QuickScan 1

Please find the Ethics and Privacy QuickScan results for the prototype evaluation focus group on the pages below. Despite the survey having triggered a warning in survey Section 3, point H9, pertaining to the possibility of serious games research being used for misinformation or propaganda, this issue was not deemed sufficiently severe to shut down the study. As such, the study was allowed to proceed.

Response Summary:

Section 1. Research projects involving human participants

P1. Does your project involve human participants? This includes for example use of observation, (online) surveys, interviews, tests, focus groups, and workshops where human participants provide information or data to inform the research. If you are only using existing data sets or publicly available data (e.g. from Twitter, Reddit) without directly recruiting participants, please answer no.

- Yes

Recruitment

P2. Does your project involve participants younger than 18 years of age?

- No

P3. Does your project involve participants with learning or communication difficulties of a severity that may impact their ability to provide informed consent?

- No

P4. Is your project likely to involve participants engaging in illegal activities?

- No

P5. Does your project involve patients?

- No

P6. Does your project involve participants belonging to a vulnerable group, other than those listed above?

- No

P8. Does your project involve participants with whom you have, or are likely to have, a working or professional relationship: for instance, staff or students of the university, professional colleagues, or clients?

- Yes

P9. Is it made clear to potential participants that not participating will in no way impact them (e.g. it will not directly impact their grade in a class)?

- Yes

Informed consent

PC1. Do you have set procedures that you will use for obtaining informed consent from all participants, including (where appropriate) parental consent for children or consent from legally authorized representatives? (See suggestions for information sheets and consent forms on [the website.](#))

- Yes

PC2. Will you tell participants that their participation is voluntary?

- Yes

PC3. Will you obtain explicit consent for participation?

- Yes

PC4. Will you obtain explicit consent for any sensor readings, eye tracking, photos, audio, and/or video recordings?

- Yes

PC5. Will you tell participants that they may withdraw from the research at any time and for any reason?

- Yes

PC6. Will you give potential participants time to consider participation?

- Yes

PC7. Will you provide participants with an opportunity to ask questions about the research before consenting to take part (e.g. by providing your contact details)?

- Yes

PC8. Does your project involve concealment or deliberate misleading of participants?

- No

Section 2. Data protection, handling, and storage

The General Data Protection Regulation imposes several obligations for the use of **personal data** (defined as any information relating to an ~~identified or identifiable~~ living person) or including the use of personal data in research.

D1. Are you gathering or using personal data (defined as any information relating to an ~~identified or identifiable~~ living person)?

- Yes

High-risk data

DR1. Will you process personal data that would jeopardize the physical health or safety of individuals in the event of a personal data breach?

- No

DR2. Will you combine, compare, or match personal data obtained from multiple sources, in a way that exceeds the reasonable expectations of the people whose data it is?

- No

DR3. Will you use any personal data of children or vulnerable individuals for marketing, profiling, automated decision-making, or to offer online services to them?

- No

DR4. Will you profile individuals on a large scale?

- No

DR5. Will you systematically monitor individuals in a publicly accessible area on a large scale (or use the data of such monitoring)?

- No

DR6. Will you use special category personal data, criminal offense personal data, or other sensitive personal data on a large scale?

- No

DR7. Will you determine an individual's access to a product, service, opportunity, or benefit based on an automated decision or special category personal data?

- No

DR8. Will you systematically and extensively monitor or profile individuals, with significant effects on them?

- No

DR9. Will you use innovative technology to process sensitive personal data?

- No

Data minimization

DM1. Will you collect only personal data that is strictly necessary for the research?

- Yes

DM4. Will you anonymize the data wherever possible?

- Yes

DM5. Will you pseudonymize the data if you are not able to anonymize it, replacing personal details with an identifier, and keeping the key separate from the data set?

- Yes

Using collaborators or contractors that process personal data securely

DC1. Will any organization external to Utrecht University be involved in processing personal data (e.g. for transcription, data analysis, data storage)?

- No

International personal data transfers

DI1. Will any personal data be transferred to another country (including to research collaborators in a joint project)?

- No

Fair use of personal data to recruit participants

DF1. Is personal data used to recruit participants?

- No

Participants' data rights and privacy information

DP1. Will participants be provided with privacy information?
(Recommended is to use as part of the information sheet: For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at www.uu.nl/en/organisation/privacy.)

- Yes

DP2. Will participants be aware of what their data is used for?

- Yes

DP3. Can participants request that their personal data be deleted?

- Yes

DP4. Can participants request that their personal data be rectified (in case it is incorrect)?

- Yes

DP5. Can participants request access to their personal data?

- Yes

DP6. Can participants request that personal data processing is restricted?

- Yes

DP7. Will participants be subjected to automated decision-making based on their personal data with an impact on them beyond the research study to which they consented?

- No

DP8. Will participants be aware of how long their data is being kept for, who it is being shared with, and any safeguards that apply in case of international sharing?

- Yes

DP9. If data is provided by a third party, are people whose data is in the data set provided with (1) the privacy information and (2) what categories of data you will use?

- Not applicable

Using data that you have not gathered directly from participants

DE1. Will you use any ~~personal data~~ that you have not gathered directly from participants (such as data from an existing data set, data gathered for you by a third party, data scraped from the internet)?

- No

Secure data storage

DS1. Will any data be stored (temporarily or permanently) anywhere other than on ~~password-protected University authorized computers or servers~~?

- No

DS4. Excluding (1) any international data transfers mentioned above and (2) any sharing of data with collaborators and contractors, ~~will any personal data be stored, collected, or accessed from outside the EU?~~

- No

Section 3. Research that may cause harm

Research may cause harm to participants, researchers, the university, or society. This includes when technology has dual-use, and you investigate an innocent use, but your results could be used by others in a harmful way. If you are unsure regarding possible harm to the university or society, please discuss your concerns with the Research Support Office.

H1. Does your project give rise to a ~~realistic risk to the national security of any country?~~

- No

H2. Does your project give rise to a ~~realistic risk of aiding human rights abuses in any country?~~

- No

H3. Does your project (and its data) give rise to a realistic risk of damaging the University's reputation? (E.g., bad press coverage, public protest.)

- No

H4. Does your project (and in particular its data) give rise to an increased risk of attack (cyber- or otherwise) against the University? (E.g., from pressure groups.)

- No

H5. Is the data likely to contain material that is indecent, offensive, defamatory, threatening, discriminatory, or extremist?

- No

H6. Does your project give rise to a realistic risk of harm to the researchers?

- No

H7. Is there a realistic risk of any participant experiencing physical or psychological harm or discomfort?


- No

H8. Is there a realistic risk of any participant experiencing a detriment to their interests as a result of participation?

- No

H9. Is there a realistic risk of other types of negative externalities?

- Yes

Ethics Warning.  **As you replied yes to one (or more) of H1-H9, a fuller ethical review is required. Please provide more detail here on the potential harm, and how you will minimize risk and impact:**

The research aims to further understanding of serious games, on the topic of embedding serious content within the narrative. The goal is to develop a method that would allow game mechanics to focus primarily on fun, while the narrative is the primary vessel of learning content. Various studies regarding the role of narrative in games have concluded that a good story can increase engagement with both the game and the story.

The risk of this research is the chance that this methodology will be used to perpetuate falsehood; a 'serious game' might be released following this method, which contains 'learning content' akin to propaganda, or factually incorrect information.

The scope of the current study pertains just a focus group to discuss the initial game prototype (without any serious content included). As such, this study will likely not suffer from this risk, rather it is the research at large which may potentially lead there. I plan to minimize impact by including various disclaimers in the research document. However, it will be difficult to assess the nature of individuals accessing the research.

Section 4. Conflicts of interest

SECTION 4. CONFLICTS OF INTEREST

C1. Is there any potential conflict of interest (e.g. between research funder and researchers or participants and researchers) that may potentially affect the research outcome or the dissemination of research findings?

- No

C2. Is there a direct hierarchical relationship between researchers and participants?

- No

Section 5. Your information.

This last section collects data about you and your project so that we can register that you completed the Ethics and Privacy Quick Scan, sent you (and your supervisor/course coordinator) a summary of what you filled out, and follow up where a fuller ethics review and/or privacy assessment is needed. For details of our legal basis for using personal data and the rights you have over your data please see the [University's privacy information](#). Please see the guidance on the [ICS Ethics and Privacy website](#) on what happens on submission.

Z0. Which is your main department?

- Information and Computing Science

Z1. Your full name:

Jan Willem de Ruig

Z2. Your email address:

j.w.deruig@students.uu.nl

Z3. In what context will you conduct this research?

- As a student for my master thesis, supervised by:
Dr. J. Frommel

Z5. Master programme for which you are doing the thesis

- Game and Media Technology

Z6. Email of the course coordinator or supervisor (so that we can inform them that you filled this out and provide them with a summary):

j.frommel@uu.nl

Z7. Email of the moderator (as provided by the coordinator of your thesis project):

coordinator.cosc@uu.nl

Z8. Title of the research project/study for which you filled out this Quick Scan:

Topic to Tale: Combining Narrative-Focused Design Methods to Increase Understanding of Narrative Importance in Serious Games

Z9. Summary of what you intend to investigate and how you will investigate this (200 words max):

This study intends to investigate how several narrative-focused design methods for serious games can come together to improve learning rates by embedding serious content within the narrative of the game, rather than the gameplay itself. This would allow the gameplay to focus on fun, while the narrative can be tailored to the desired topic.

Z10. In case you encountered warnings in the survey, does supervisor already have ethical approval for a research line that fully covers your project?

- No
-

Scoring

- Privacy: 0
 - Ethics: 1
-

10 Appendix B : QuickScan 2

Please find a replica of the Ethics and Privacy QuickScan results for the survey user study on the pages below. This QuickScan had been appropriately completed prior to carrying out the study, however the file has been misplaced since. The replica below is a similarly completed version included for archival purposes.

Response Summary:

Section 1. Research projects involving human participants

P1. Does your project involve human participants? This includes for example use of observation, (online) surveys, interviews, tests, focus groups, and workshops where human participants provide information or data to inform the research. If you are only using existing data sets or publicly available data (e.g. from Twitter, Reddit) without directly recruiting participants, please answer no.

- Yes

Recruitment

P2. Does your project involve participants younger than 18 years of age?

- No

P3. Does your project involve participants with learning or communication difficulties of a severity that may impact their ability to provide informed consent?

- No

P4. Is your project likely to involve participants engaging in illegal activities?

- No

P5. Does your project involve patients?

- No

P6. Does your project involve participants belonging to a vulnerable group, other than those listed above?

- No

P8. Does your project involve participants with whom you have, or are likely to have, a working or professional relationship: for instance, staff or students of the university, professional colleagues, or clients?

- Yes

P9. Is it made clear to potential participants that not participating will in no way impact them (e.g. it will not directly impact their grade in a class)?

- Yes

Informed consent

PC1. Do you have set procedures that you will use for obtaining informed consent from all participants, including (where appropriate) parental consent for children or consent from legally authorized representatives? (See suggestions for information sheets and consent forms on [the website](#).)

- Yes

PC2. Will you tell participants that their participation is voluntary?

- Yes

PC3. Will you obtain explicit consent for participation?

- Yes

PC4. Will you obtain explicit consent for any sensor readings, eye tracking, photos, audio, and/or video recordings?

- Not applicable

PC5. Will you tell participants that they may withdraw from the research at any time and for any reason?

- Yes

PC6. Will you give potential participants time to consider participation?

- Yes

PC7. Will you provide participants with an opportunity to ask questions about the research before consenting to take part (e.g. by providing your contact details)?

- Yes

PC8. Does your project involve concealment or deliberate misleading of participants?

- No

Section 2. Data protection, handling, and storage

The General Data Protection Regulation imposes several obligations for the use of **personal data** (defined as any information relating to an identified or identifiable living person) or including the use of personal data in research.

D1. Are you gathering or using personal data (defined as any information relating to an identified or identifiable living person)?

- Yes

High-risk data

DR1. Will you process personal data that would jeopardize the physical health or safety of individuals in the event of a personal data breach?

- No

DR2. Will you combine, compare, or match personal data obtained from multiple sources, in a way that exceeds the reasonable expectations of the people whose data it is?

- No

DR3. Will you use any personal data of children or vulnerable individuals for marketing, profiling, automated decision-making, or to offer online services to them?

- No

DR4. Will you profile individuals on a large scale?

- No

DR5. Will you systematically monitor individuals in a publicly accessible area on a large scale (or use the data of such monitoring)?

- No

DR6. Will you use special category personal data, criminal offense personal data, or other sensitive personal data on a large scale?

- No

DR7. Will you determine an individual's access to a product, service, opportunity, or benefit based on an automated decision or special category personal data?

- No

DR8. Will you systematically and extensively monitor or profile individuals, with significant effects on them?

- No

DR9. Will you use innovative technology to process sensitive personal data?

- No

Data minimization

DM1. Will you collect only personal data that is strictly necessary for the research?

- Yes

DM4. Will you anonymize the data wherever possible?

- Yes

DM5. Will you pseudonymize the data if you are not able to anonymize it, replacing personal details with an identifier, and keeping the key separate from the data set?

- Not applicable

Using collaborators or contractors that process personal data securely

DC1. Will any organization external to Utrecht University be involved in processing personal data (e.g. for transcription, data analysis, data storage)?

- No

International personal data transfers

DI1. Will any personal data be transferred to another country (including to research collaborators in a joint project)?

- No

Fair use of personal data to recruit participants

DF1. Is personal data used to recruit participants?

- No

Participants' data rights and privacy information

DP1. Will participants be provided with privacy information? (Recommended is to use as part of the information sheet: For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at www.uu.nl/en/organisation/privacy.)

- Yes

DP2. Will participants be aware of what their data is used for?

- Yes

DP3. Can participants request that their personal data be deleted?

- Yes

DP4. Can participants request that their personal data be rectified (in case it is incorrect)?

- Yes

DP5. Can participants request access to their personal data?

- Yes

DP6. Can participants request that personal data processing is restricted?

- Yes

DP7. Will participants be subjected to automated decision-making based on their personal data with an impact on them beyond the research study to which they consented?

- No

DP8. Will participants be aware of how long their data is being kept for, who it is being shared with, and any safeguards that apply in case of international sharing?

- Yes

DP9. If data is provided by a third party, are people whose data is in the data set provided with (1) the privacy information and (2) what categories of data you will use?

- Not applicable

Using data that you have not gathered directly from participants

DE1. Will you use any personal data that you have not gathered directly from participants (such as data from an existing data set, data gathered for you by a third party, data scraped from the internet)?

- No

Secure data storage

DS1. Will any data be stored (temporarily or permanently) anywhere other than on password-protected University authorized computers or servers?

- No

DS4. Excluding (1) any international data transfers mentioned above and (2) any sharing of data with collaborators and contractors, will any personal data be stored, collected, or accessed from outside the EU?

- No

Section 3. Research that may cause harm

Research may cause harm to participants, researchers, the university, or society. This includes when technology has dual-use, and you investigate an innocent use, but your results could be used by others in a harmful way. If you are unsure regarding possible harm to the university or society, please discuss your concerns with the Research Support Office.

H1. Does your project give rise to a realistic risk to the national security of any country?

- No

H2. Does your project give rise to a realistic risk of aiding human rights abuses in any country?

- No

H3. Does your project (and its data) give rise to a realistic risk of damaging the University's reputation? (E.g., bad press coverage, public protest.)

- No

H4. Does your project (and in particular its data) give rise to an increased risk of attack (cyber- or otherwise) against the University? (E.g., from pressure groups.)

- No

H5. Is the data likely to contain material that is indecent, offensive, defamatory, threatening, discriminatory, or extremist?

- No

H6. Does your project give rise to a realistic risk of harm to the researchers?

- No

H7. Is there a realistic risk of any participant experiencing physical or psychological harm or discomfort?

- No

H8. Is there a realistic risk of any participant experiencing a detriment to their interests as a result of participation?

- No

H9. Is there a realistic risk of other types of negative externalities?

- No

Section 4. Conflicts of interest

C1. Is there any potential conflict of interest (e.g. between research funder and researchers or participants and researchers) that may potentially affect the research outcome or the dissemination of research findings?

- No

C2. Is there a direct hierarchical relationship between researchers and participants?

- No

Section 5. Your information.

This last section collects data about you and your project so that we can register that you completed the Ethics and Privacy Quick Scan, sent you (and your supervisor/course coordinator) a summary of what you filled out, and follow up where a fuller ethics review and/or privacy assessment is needed. For details of our legal basis for using personal data and the rights you have over your data please see the [University's privacy information](#). Please see the guidance on the [ICS Ethics and Privacy website](#) on what happens on submission.

Z0. Which is your main department?

- Information and Computing Science

Z1. Your full name:

Jan Willem de Ruig

Z2. Your email address:

j.w.deruig@students.uu.nl

Z3. In what context will you conduct this research?

- As a student for my master thesis, supervised by:
Dr. J. Frommel

Z5. Master programme for which you are doing the thesis

- Game and Media Technology

Z6. Email of the course coordinator or supervisor (so that we can inform them that you filled this out and provide them with a summary):

j.frommel@uu.nl

Z7. Email of the moderator (as provided by the coordinator of your thesis project):

gmt-ethics@uu.nl

Z8. Title of the research project/study for which you filled out this Quick Scan:

Topic to Tale: Combining Narrative-Focused Design Methods to Increase Understanding of Narrative Importance in Serious Games

Z9. Summary of what you intend to investigate and how you will investigate this (200 words max):

This study intends to investigate how several narrative-focused design methods for serious games can come together to improve learning rates by embedding serious content within the narrative of the game, rather than the gameplay itself. This would allow the gameplay to focus on fun, while the narrative can be tailored to the desired topic.

Z10. In case you encountered warnings in the survey, does supervisor already have ethical approval for a research line that fully covers your project?

- Not applicable

Scoring

- Privacy: 0
 - Ethics: 0
-

11 Appendix C : Questionnaire

Please find the questionnaire used throughout the user study in the pages below.

Preamble

Welcome!

In this survey, you will be asked to play a short serious game prototype, and answer a few questions about your playing and learning experience. Please make sure you are filling in this survey on a desktop or laptop, as the prototype does not support mobile browsers

This survey will take around half an hour to complete; twenty minutes of that will be playing the prototype, and the survey itself will about take ten minutes to complete.

The survey is split in three main sections. First, there will be room for you to give informed consent. Second are a small amount of demographic questions, to contextualise the survey result. And third, after the prototype playtime, you will be questioned on your experience playing the game.

If you would like to receive further details about the survey or the study goals, feel free to contact me at j.w.deruig@students.uu.nl

Informed Consent

Please read carefully through this consent form. By ticking the box at the bottom of the page, you agree to the following:

- I confirm that I am 18 years of age or older
- I confirm that the research project has been explained to me, or I have had the opportunity to ask questions about the project and have had these answered satisfactory. I had enough time to consider whether to participate
- I consent to the material I contribute being used to generate insight for the research project
- I consent to have limited personal information measured in this survey, including information about my age, gender, and education. I understand this information will only be used to contextualise aggregate survey findings
- I understand that if I give permission, my personal data will be held confidentially and anonymously on the Qualtrics servers, for as long as is necessary to process these, and for no longer than six weeks
- I understand that my participation in this research is voluntary and that I may withdraw from the study at any time without providing a reason, and that if I withdraw any personal data already collected from me will be erased
- I consent to allow the fully anonymised data to be used in future publications and other scholarly means of disseminating the findings of this research project

- I understand that the data acquired will be securely stored by researchers, but that appropriately anonymised data may in future be made available to others for research purposes. I understand that the University may publish appropriately anonymised data in appropriate data repositories for verification purposes and to make it accessible to researchers and other research users.

I understand the above listed, and agree to take part in the research project on "Combining Narrative-Focused Design Methods to Increase Understanding of Narrative Importance in Serious Games"

- Yes, I understand
- No, take me back

Demographics

Please enter your age in years

What is your gender? *Select all that apply*

- Male
- Female
- Non-binary

Prototype II

You have been selected to play Prototype II.

The game is hosted on Itch.IO, and can be played online in your desktop browser; you can find a link to it [here](#). Please open the link in a new tab, to prevent loss of data. No information will be collected during gameplay. After the game has finished, please enter the verification code obtained from it below to continue the survey.

Verification code:

Prototype I

You have been selected to play Prototype I.

The game is hosted on Itch.IO, and can be played online in your desktop browser; you can find a link to it [here](#). Please open the link in a new tab, to prevent loss of data. No information will be collected during gameplay. After the game has finished, please enter the verification code obtained from it below to continue the survey.

Learning Effectiveness

Rate the following statements on your familiarity with the targeted learning content after playing the game

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I know what 'Induced Demand' is	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am aware of how 'Induced Demand' impacts my surroundings	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
'Induced Demand' makes intuitive sense to me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My understanding of 'Induced Demand' has increased	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My interest in 'Induced Demand' has increased	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

To validate the learning effectiveness, please provide a short summary of your understanding of 'Induced Demand' (max. 200 words)