Narrative categorization in digital game-based learning: Engagement, motivation & learning

Fredrik S. Breien and Barbara Wasson

Fredrik S. Breien is an industrial PhD candidate at Center for the Science of Learning and Technology (SLATE), Department of Information Science & Media Studies (Infomedia), University of Bergen and the Bergen Science Center VilVite. His background is from the serious games industry, his area of interest concerns narratives in game design. Barbara Wasson is Director of SLATE and Professor at Infomedia, University of Bergen. She has over 150 publications in the field of Technology Enhanced Learning (TEL). Wasson was one of the founders of Kaleidoscope, a European Network of Excellence on TEL. The European Commission often uses her as an expert evaluator. Her current research interests include interaction design, CSCL, learning games, AI in education, e-assessment, teacher inquiry and learning analytics. Address for correspondence: Fredrik S. Breien, Department of Information Science & Media Studies, Centre for the Science of Learning & Technology, University of Bergen, Bergen 5020, Norway. Email: fredrik.breien@gmail.com

Abstract
Previous research shows that digital game-based learning (DGBL) can have positive effects on engagement, motivation and learning, and that using narratives may reinforce these effects. A systematic review identified 15 DGBL systems that report effects from their use of narratives. A gap in the field, however, is the lack of a common model to categorize and isolate narratives in DGBL to enable an analysis and comparison of how, and under what conditions, narratives have effects on learning in DGBL systems. The ludo narrative variable model (LNVM) that has been used to isolate and categorize narratives in research on commercial video games is a candidate to fill this gap. This research has investigated the potential of this model for DGBL and resulted in an extended LNVM (eLNVM) that can be used to isolate and categorize narratives in DGBL. The 15 DGBL systems were categorized on the eLNVM and the results show that there are characteristics of DGBL systems with positive self-reported effects that separate them from other DGBL systems. Furthermore, it was possible to identify characteristics of the narrative modeling that are associated with positive effects on engagement, motivation and learning. The paper concludes with a description of how the eLNVM will be used in future research.

Introduction
Digital Game-Based Learning (DGBL), first termed by Prensky (2001), refers to any form of use or integration of digital games into learning environments. DGBL spans from serious games developed with learning or instructional objectives, to student game design activities to explore a subject matter, and the use of commercial off-the-shelf (COTS) digital games to address both content-based and high-order learning outcomes (Van Eck, 2006). In their review of gamified learning, Subhash and Cudney (2018) define the digital domain of Game-Based Learning (GBL) to separate between non-game systems such as Learning-Management Systems (LMS) that use some game elements and mechanics, called gamification (Deterding, Dixon, Khaled, & Nacke, 2011), and games designed for, or used in, educational settings. The latter range from role-play
games (RPGs), serious games, gamified applications, mobile-based learning games and 3D simulation games for learning. DGBL can also include COTS games used in educational settings, and probably also a wider set of video game genres added to RPGs and 3D simulation games, which are listed as main game genres alongside strategy, and action in Apperly’s (2006) critical approach to game genres. A simple and summative definition of DGBL is that it “incorporates educational content or learning principles into video games” (Coffey, 2009, p. 1). This study adopts the definition provided by Coffey and refers henceforth to any digital game used for education, instruction or training as a DGBL system.

One key aspect of DGBL is the narrative or story, which Ermi and Mäyrä (2005) show to be one of three core game structures, alongside rules and space. Serious games (DGBL systems) “harness the power of digital games for training or education” (Fu-Hsing, Kuang-Chao, & Hsien-Sheng, 2012, p. 240). Serious games are growing in popularity (Hersh & Leporini, 2018), and support learners in learning effectively and to enjoy learning (e.g., Coffey, 2009; Ferguson et al., 2019; Fu-Hsing et al., 2012; Rushby, 2012). Rushby (2012), however, concludes that while “there is evidence serious games do help people learn, there is very little evidence as to how they do that” (Rushby, 2012, p. 179) and calls for further study related to (1) which characteristics make good serious games, and (2) how good serious games can be made even more effective.

Several literature reviews (e.g., Clark et al., 2016; Connolly et al., 2012; Novak, 2015) have explored narrative effects on engagement, motivation, and learning in DGBL. There is a consensus among these reviews that DGBL can have positive effects on, and offer novel venues for, learning in some
Narrative Categorization in DGBL circumstances, by motivating and engaging learners using narratives. Since narratives are a key aspect of DGBL, we explored how narratives are used to enhance engagement, motivation and learning in DGBL systems.

It is important to understand how each game design feature affects the specific learning outcomes and/or benefits when examining their relationship to learning (Novak, 2015). Lester et al. (2014) call for further study to isolate specific game design features to understand which features are relevant to student engagement. While investigating narrative effects on learning, Armstrong and Landers (2017) found a scarcity of current studies on narrative DGBL systems that isolate narratives from other game elements. According to them this makes unclear the conditions under which narratives themselves impact learning, and whether effects attributed to narratives come from them at all. Consequently, they call for further study of how, and under what conditions, narratives have effects on learning in order to conduct meaningful future research. This leads to the current study of how narratives are categorized in DGBL, and what characteristics, if any, define DGBL systems with positive effects on engagement, motivation and learning. We concur with the findings of Armstrong and Landers; not only is there a shortage of studies that isolate narratives from game play effects on learning, the papers reviewed in this study also reveal that there are many ways that researchers categorize narratives in DGBL. While this leads to various means and opportunities to analyze individual DGBL systems, it makes comparison challenging.

The study reported in this paper contributes to current research by presenting a model for the isolation and categorization of narrative effects from other game mechanics by drawing on, and extending, Aarseth’s (1997, 2005, 2012) work on narrative game theory. Through a systematic literature review (Catalá-López, Stevens, Garrity & Hutton, 2017; Kitchenham, 2007; Kitchenham & Brereton, 2013) of narrative DGBL systems from the last 10 years, and by categorizing the evaluated systems on an extension of Aarseth’s LNVM, our research explores whether the proposed extended LNVM (eLNVM) is a suitable tool to isolate and categorize narratives in DGBL. Furthermore, to investigate whether effects can be characterized using eLNVM, the DGBL systems’ categorizations are clustered by their reported effects on engagement, motivation and learning.

Our results show there are distinct and isolated characteristics of narratives that are integrated in DGBL systems that are also evaluated to successfully reinforce engagement, motivation, and learning. eLNVM is useful to perform categorization of narrative characteristics, which can further be used to provide development guidelines that can be used by educators and game developers, separately or in multidisciplinary teams. The use of these design guidelines has the potential to optimize the use of resources that go into designing narratives in effective DGBL. Furthermore, the results show that eLNVM is an effective tool to isolate narratives and narrative effects on learning in single DGBL systems, as well as when comparing them to one another.

Background for the study

DGBL, learning theories and reach

According to Coffey (2009), DGBL draws upon a constructivist theory of learning. In their review of GBL and 21st century skills, Qian and Clark (2016) analyzed 22 studies and found that of the 28 explicit references to 10 established learning theories, 13 of the references were to constructivism, with the second most referenced learning theory being constructionism, with only four references. Qian and Clark (2016) show that DGBL often fail to use learning theoretical foundations, however, they go on to show that DGBL tends to yield positive outcomes when learning theories are incorporated, and that particularly a sociocultural theory of learning (Vygotsky, 1978).
and flow theory (Csikszentmihalyi & Csikzentmihaly, 1990) are well-suited to game design and learning outcomes (Qian & Clark, 2016).

As described in “Innovating pedagogy 2019” (Ferguson et al., 2019), research shows that digital games have been used by all age groups, and that they can help learners develop skills such as collaboration, problem solving and creativity. Confirming the broad reach of DGBL across age and subject matters, Clark, Tanner-Smith and Killingsworth’s (2016) review of digital games, design and learning generate a dataset of 69 studies of DGBL used in K-16 education, published between 2000 and 2012. Through a meta-review they further identify 65 DGBL used in adult workflow training between 1976 and 2009, and 71 DGBL used by all age groups between 1986 and 2012.

In their granular review of 29 studies of implemented DGBL, Qian and Clark (2016) identified 28 game design elements. The most implemented game design elements were collaboration (eight times) and role playing (seven times), with narrative (six times) being the third most utilized. The strength of narratives as a game element is evidenced in commercial video gaming through successes of narrative based titles such as Grand Theft Auto and Red Dead Redemption, which have sold 290 and 24 million units to date respectively (Take Two Interactive annual report, 2019) and have excellent review scores in past and present instalments (https://www.metacritic.com/game/playstation-4/grand-theft-auto-v; https://www.metacritic.com/game/playstation-4/red-dead-redemption-2).

This study is motivated by the above-mentioned research which shows that DGBL can employ effective learning theories, and moreover, is able to provide a meaningful platform for learning for wide demographics. The research seeks to provide a means for isolating categories in DGBL, and to explore which characteristics enable a better understanding of the conditions under which positive learning effects occur.

**Effect of narratives on engagement, motivation and learning**

Engagement, motivation and learning in DGBL systems has been explored for nearly 40 years. Thomas W. Malone’s seminal paper “Toward a Theory of Intrinsically Motivating Instruction” (Malone, 1981) is considered by many (eg, Dempsey, Lucassen, Gilley, & Rasmussen, 1993; Fulya Eyupoglu & Nietfeld, 2019; Habgood, Ainsworth & Benford, 2005) to be the initial research in this area. Specific research on narratives in DGBL was thoroughly examined and defined by Marc Prensky through his pioneer work in the field (eg, Prensky & Thiagarajan, 2007).

Narratives used to facilitate learning in other forms of instructional contexts (eg, through books, movies, audio, etc.) have been shown to be superior to expository texts with regards to fostering comprehension, retention and recall (Armstrong & Landers, 2017). While advantages of pairing narratives with DGBL have been suggested (eg, Dettori & Paiva, 2009), evaluations of narrative DGBL have yielded mixed results. Clark et al. (2016) find insignificant results that show DGBL may be more effective without, than with narratives. Novak (2015) does not find positive motivational effects from narratives but argues that narratives may still provide contextual anchoring and facilitate better knowledge construction and information organization (Novak, 2015). Other studies, such as empirical tests on the Crystal Island narrative DGBL (Lester et al., 2014; Rowe, Lucy, Bradford, & James, 2010; Rowe, Shores, Mott, & Lester, 2011; Lee, Mott, & Lester, 2010; Xu & Woodruff, 2017), have shown positive effects from narratives on engagement and motivation, as well as on learning objectives and processes such as cognitive modeling and memory.

With regards to learning outcomes, warnings are raised as to: (1) if narrative conditions in DGBL impose an excessive cognitive load on learners, which negatively affects their learning outcomes
(Novak, 2015); (2) that cognitive capacity is consumed by following a game narrative, rendering the learner with insufficient capacity to think deeply about the academic material in DGBL (Pilegard & Mayer, 2016); (3) that switching between story text and processing of other information may cause cognitive overload with detrimental effects on learning (Ross, Pye, & Randell, 2016); and (4) that young children may be particularly prone to such overload, due to their immature cognitive and attention skills (Courage et al., 2015).

The above research shows that while narratives in DGBL are not well understood, they enforce positive effects on learning in other domains. Narratives are important game elements and have the potential to positively affect learning in DGBL. Thus, this study focuses on narratives in DGBL and their effect on engagement, motivation and learning.

The Ludo Narrative variable model

In his book “Cybertext: Perspectives on Ergodic Literature” Aarseth defines Ergodic Literature as any piece of literature where nontrivial effort is required to allow the reader to traverse a text (or story), beyond trivialities such as eye movement and the periodic, or arbitrary, turning of pages (Aarseth, 1997). The separation between triviality and nontriviality intrinsically defines all narrative games as Ergodic literary works, in that a story that is traversed only using trivial motoric abilities may not be considered a game of any form, while the interactivity of any game can be defined as nontrivial and needed to drive a game narrative forward. In a further study, Aarseth introduces a narrative theory of games and the LNVM (Aarseth, 2012), which is based on his related work on game quest structure and storytelling (Aarseth, 2005).

Recognizing computer games as a dominant cultural form that influence others such as cinema, TV, literature, theater, painting and music, Aarseth (2012) poses the question of whether narrative theory should be modified or expanded to incorporate games. In seeing critical self-reflection remaining a hallmark of scholarship, Aarseth (2012) reminds us that when we examine a phenomenon with critical tools developed for another type of phenomenon, the tools must be critically examined for theoretical concepts to remain meaningful when transported to new fields. To this end Aarseth (2012) proposes the LNVM as a usable model for categorizing and studying narratives in games inside four ontological dimensions that are present in all games as well as narratives.

Table 1 shows the LNVM, which constitutes a categorized graph from the narrative pole to ludic (or game) pole inside each of the ontologies that have been shown to be shared by all games and narratives. The narrative pole offers high author agency, at the sacrifice of game agency, and the game pole offers high game agency, at the sacrifice of author agency. The object and agents ontologies are categorized on linear scales. For objects this implies a scale of objects from something that may only be observed to exist under the non-interactable category, to something that can be invented in the game, by for example using programing code available inside the game interface under the inventible category. For agents this implies a scale of character depth, from characters described in detail with regards to background, motivation, psychology, and so on, under the deep, rich, round category, to characters that exist with no purpose other than gameplay, such as hordes of aliens in a shooting game under the bots, no individual identity category. Contrastingly, the world and events ontologies are categorized as taxonomies. These are different from each other and imply no scale, but rather implementation methodologies in video games that to a larger and lesser extent offer polar game to author agency. While the world and event ontologies have emerged in commercial video game development since the 1970s, and were both scientifically described by Aarseth (2005, 2012), these ontologies merit further description. Table 2 describes the taxonomies for the world ontology, concerning how a game environment evolves based on
user action. Table 3 describes the taxonomies of the events ontology. These taxonomies concerns kernels, cardinal functions of a narrative that are both necessary and sufficient for a particular story to exist, also known as nuclei (Barthes & Dulsit, 1975), opposed to satellites, who in narratology are divided into three optional expansions to a story; catalyzes, who are consecutive but not consequential to the story; and, implicit and explicit indices, adding information to narrative events (Kwiat, 2008).

Using LNVM to categorize a game and analyzing the game’s effect on “something” before comparing it to other games with the same kinds of effects, characteristics may emerge as to which degree a game uses game- and author agency inside the separate ontologies to successfully reach its goals.

This study recognizes the need to further categorize and characterize narratives in DGBL and to promote efficient strategies for effective design and comparable evaluation of future systems. This research applies the LNVM to explore its potential as a tool to meaningfully isolate and categorize narratives from game elements in DGBL and explores whether there are characteristics that are associated with positive effects on engagement, motivation and learning.

**Research questions**

Previous research shows that DGBL can have positive effect on learning for broad demographics when properly applying learning theories, and that narratives in DGBL may enforce learning. A
gap in the current research is the lack of a common model to isolate and categorize narratives and other game elements in DGBL systems. This makes it difficult to identify how, and under what conditions narratives have effect on engagement, motivation, and learning. We address this gap by investigating whether the LNVM can be such a common model. The study is guided by two research questions:

1. How can the LNVM be used to isolate and categorize narrative DGBL systems?
2. What characterizes narrative DGBL systems that positively affect engagement, motivation and learning?

Method
A systematic literature review with a search protocol and inclusion criteria process (Kitchenham, 2007; Kitchenham & Brereton, 2013) limited to the simplifications set forth by Catalá-López et al. (2017) in describing rapid reviews for evidence synthesis was carried out. The development and implementation of the search protocol comprised three stages:

1. The unstructured initial informal search across 10-fold databases to test different combinations of search words and database results,
2. The test protocol to develop search string and criteria, and to select the final databases to use, and
3. The main protocol, in which the actual search protocol was conducted using our search query and inclusion criteria.

A search for research papers about narrative DGBL systems published in scientific journals or conference proceedings between January 2009 and August 2018 used the following search string:

Title: (Narrative OR Story) AND Title/Abstract/Keywords: (Game) AND (Learning OR Training OR Education) AND (Motivation OR Engagement)

The results were evaluated through four exclusion rounds to produce a dataset of DGBL systems’ that could be coded by their self-reported effects from narratives on engagement, motivation, and learning. Based on their qualitative descriptions, the DGBL systems in the dataset were then categorized on the LNVM. The clusters were compared and analyzed to explore characteristics that isolated and separated clusters of DGBL systems showing different effects from one another.

| Taxonomy                        | Description                                                   |
|---------------------------------|---------------------------------------------------------------|
| Fully plotted                   | Neither kernels nor satellites content or sequence is affected by interaction |
| Dynamic satellites, playable story | Kernels content and sequence remains fixed, while satellites can be altered or rearranged |
| Dynamic kernels                 | Content of kernels may be changed by interaction, leading to potentially different stories |
| No kernels                      | A game with no narrative                                      |
| Year  | ACM Digital Library (IT) | IEEE Xplore (IT) | PsycINFO (Psychology) | Science Direct (General) | Web of Science (General/learning) | Total per year |
|-------|-------------------------|------------------|-----------------------|--------------------------|---------------------------------|----------------|
| 2009  | 0                       | 0                | 1                     | 1                        | 1                               | 3              |
| 2010  | 1                       | 0                | 1                     | 0                        | 1                               | 3              |
| 2011  | 2                       | 1                | 3                     | 2                        | 2                               | 10             |
| 2012  | 1                       | 1                | 1                     | 2                        | 1                               | 6              |
| 2013  | 0                       | 2                | 1                     | 1                        | 1                               | 5              |
| 2014  | 1                       | 0                | 0                     | 3                        | 3                               | 7              |
| 2015  | 1                       | 1                | 0                     | 0                        | 1                               | 3              |
| 2016  | 1                       | 1                | 0                     | 1                        | 3                               | 6              |
| 2017  | 2                       | 1                | 1                     | 3                        | 4                               | 11             |
| 2018(7)| 3                       | 0                | 1                     | 2                        | 1                               | 7              |
| Total per database | 12                      | 7                | 9                     | 15                       | 18                              | 61             |

Table 4: The distribution of papers as found per database per year
Results
The search query, see Table 5, produced 61 papers, the distribution of which is presented as obtained across five databases, by year, in Table 4. The five databases were selected after showing good coverage during initial unstructured search and test protocols, and for covering a broad range of disciplines, two with IT related content also including games studies (ACM, IEEE), one with focus on psychology (PsycINFO), and another two of general purpose (Science Direct, Web of Science), the later with a strong relation to learning sciences.

Completing the main search protocol and applying the inclusion criteria resulted in a dataset comprising of 14 papers and 15 DGBL system evaluations, as one paper (Barab et al., 2010) evaluated two different DGBL systems. In Table 5, the search protocol and inclusion criteria rounds are described as they were performed: Main search, Duplicate removal, False positive removal, and four Inclusion rounds (abstract, full text, full text, system description).

The evaluations of the 15 DGBL systems were explored for self-reported effects on engagement, motivation, and learning. These were coded positive, negative, neutral, or not applicable (N/A). In cases where an evaluation reflected on several demographics, a positive report was liberally noted if positive effects were found for one or more of the demographics. Table 6 shows the distribution of evaluated effects reports in each paper.

As the categorization of DGBL systems was undertaken, two shortcomings were found in the LNVM, resulting in an extension of the model itself. The first extension is related to the category naming in the agents ontology, the second is related to the characteristics of categories in the events ontology.

While Aarseth provides a linear scale of characteristics in agents depth from his Deep, rich, round to Flat character categories in the agents ontology that is visible in his categorization of different works across an open scale between them in his original presentation of a narrative model for games (Aarseth, 2012), he does not, however, label the interim categories by name. To label these and thus provide a clearer terminology for future discourse when analyzing narrative games, including narrative DGBL systems, the agents ontology is extended with three categories as shown in Table 7 (extended categories in bold):

When plotting the DGBL systems on the events ontology, it became apparent that the Dynamic satellites/ playable story category encompasses two types of characteristics that constitute fundamentally different event structures in games. To capture this granularity the events ontology is extended as shown in Table 8 (extended categories names in bold):

Based on the extensions, Table 9 shows the extended LNVM (eLNVM), which is further used to categorize the DGBL systems.

The 15 DGBL systems were categorized on the eLNVM as shown in Table 10. When described by ontology, most were linear corridor worlds (7) offering static, usable objects (8) to grounded, consistent agents (8) that navigated interchangeable fixed kernels events (5).

Analysis
Using the eLNVM ontological categorization for each DGBL system (Table 10) and the DGBL systems’ evaluated effects (Table 6) as basis, an analysis was made to explore how clustering is distributed for each of the ontologies. The coded systems were sorted as follows:

1. Fully positive (nine DGBL system evaluations): Learning outcomes are positive, both engagement and motivation are either positive or N/A.
Table 5: Search protocol and paper inclusions criteria

| Main search protocol (in databases listed in Table 4) |
|-----------------------------------------------------|
| Search query:                                        |
| “Title: (Narrative OR Story) AND Title/Abstract/Keywords: (Game) AND (Learning OR Training OR Education) AND (Motivation OR Engagement)” |
| 61 papers                                            |
| ↓                                                   |
| **Removal of duplicates**                           |
| 50 papers                                           |
| ↓                                                   |
| **Removal of false positives**                      |
| E.g. positive return of paper by search query by inclusion of search words inside other words with different meaning (like the word “history” giving a positive result on the embedded word “story”), or reference to “narrative evaluation methods”—Verified by two researchers |
| 40 papers                                           |
| ↓                                                   |
| **Inclusion round one: Abstract**                   |
| “The study must be about the design—or evaluation of a fully or partially interactive digital tool, system or process that teaches or trains someone in a subject matter, skill, trait or process”—Verified by two researchers |
| 27 papers                                           |
| ↓                                                   |
| **Inclusion round two: Full text**                  |
| Three conditions (1) The study does not target users with special assistive needs or medical conditions, (2) The study is targeted at school children, students or adults in professional or academic settings, and (3) The study provide results, observations, comments, recommendations, practices or other relevant information about narratively obtained motivation and/or engagement and learning outcomes from the learner or practitioner—Verified by two researchers |
| 19 papers                                           |
| ↓                                                   |
| **Inclusion round three: Full text**                |
| Classify by (1) Study type, (2) Sample size, (3) Demography, (4) Method, (5) Subject matter, (6) Game type/genre. Papers removed if problematic—Verified by two researchers |
| 18 papers                                           |
| ↓                                                   |
| **Inclusion round four: System description**        |
| Confirm that the evaluated DGBL systems were qualitatively sufficiently described to meaningfully categorize them on the Ludo-Narrative variable model (by reading the papers, but also by following links to on-line resources, playing games, watching academic videos of game play sessions, and/or by following reference links to previous papers in which the DGBLs were further described)—Performed in detail by one researcher, reported for all eliminations to another researcher |
2. Neutral learning (two DGBL system evaluations): Learning outcomes are neutral, both engagement and motivation is either positive or N/A, 
3. No positive effects (three DGBL system evaluations): Neither engagement nor motivation is positive, and 
4. Negative learning (one DGBL system evaluation): Learning outcomes are negative.

Only six of the 15 DGBL studies report on both engagement and motivation, but all studies report on either one or the other. Therefore, the distinction between clusters 1 and 3 was intentionally made to separate evaluations that showed positive effects on at least one of engagement or motivation, as well as no neutral or negative effect on the other, from those evaluations that reported no positive effects on learning outcomes, engagement or motivation. Among the nine fully positive evaluations, two failed to report on engagement, while four failed to report on motivation.

The DGBL systems clustered by coded effects and categories, show the characteristics that both separate and isolate DGBL systems with fully positive effects from all others. This is shown and described for each ontology in Figures 1 through 4.

Figure 1 shows that there is a clear characteristic for fully positive DGBL systems to gravitate towards the taxonomical hubshaped quest landscape and open landscape categories, and moreover, only positive coded evaluations of DGBL systems fall into these two categories. Only two positive coded DGBL systems are categorized under the linear corridor category, which is also a gravitational category for neutral learning and no positive effects DGBL systems. The negative learning coded DGBL system falls under the inaccessible category. None of the selected systems fall into the single room and multicursal labyrinth categories.

Figure 2 shows that while eight DGBL systems fall under the static, usable category, only two of these are fully positive. The other fully positive DGBL systems gravitate alone to the modifiable and creatable categories. The events ontology is a linear scale, where a higher degree of game agency also implies all lower degree categories for the same objects, for example, a creatable object is necessarily also modifiable, since creating any new object may also be viewed as modifying another with larger or smaller changes to different parameters. Furthermore, an object of any type can also be treated as one of higher author agency than itself, eg, any object of higher game agency than modifiable may be considered static or even non-interactable. It cannot be known from this data whether the reason that the two DGBL systems under the creatable category have positive effects is because objects are creatable, or because they are modifiable. It is thus considered safest to characterize fully positive coded DGBL systems as modifiable or higher game agency categories in this research. No DGBL systems fall under the noninteractable, destructible or inventible categories, however, the former two are both implied as part of the DGBL systems that fall under the creatable category.
Table 6: Evaluations effect reports per paper in dataset

| Study                  | Effect on engagement | Effect on motivation | Effect on learning outcomes |
|------------------------|----------------------|----------------------|-----------------------------|
| Barab et al. (2010)    | Positive             | N/A                  | Positive                    |
| Dickey (2011)          | Positive             | Positive             | Positive                    |
| Garneli et al. (2017)  | Neutral              | Negative             | Neutral                     |
| Jemmali et al. (2018)  | Positive             | Positive             | Positive                    |
| Lester et al. (2014)   | Positive             | Positive             | Positive                    |
| Marsh et al. (2011)    | Positive             | N/A                  | Positive                    |
| Pilegard and Mayer (2016) | Neutral         | Neutral              | Positive                    |
| Ross et al. (2016)     | Positive             | Neutral              | Negative                    |
| Rowe et al. (2010)     | Positive             | N/A                  | Positive                    |
| Rowe et al. (2011)     | Positive             | N/A                  | Positive                    |
| Sangalang et al. (2013)| N/A                  | Positive             | Neutral                     |
| Wang et al. (2017)     | N/A                  | Neutral              | Neutral                     |
| Wouters et al. (2011)  | N/A                  | Positive             | Neutral                     |
| Zhang et al. (2018)    | Positive             | N/A                  | Positive                    |

*a*Three studies about the same narrative DGBL but are treated separately due to (1) modification and (2) different research questions and focuses.

*b*This study reports on different game versions. Criteria and effects in the table focus on the “Rich Narrative Version.”

*c*This study reports on two systems where only one is a DGBL. Criteria and effects in the table focus on “Fantastic Flying Books of Morris Lessmore.”

*d*This study reports on two narrative DGBL which for research criteria and effects are equal, but that are categorized differently in the LNVM ontologies.

Table 7: The extended agents ontology

| Category                                | Description                                                                                                                                 |
|-----------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Deep, rich, round characters            | Aarseth’s definition                                                                                                                                 |
| Grounded, consistent characters         | Characters that are described in terms of background, origin, morale, culture, etc., and that act in a seemingly consistent and sensible manner, but that are not fleshed out and described in the detail like the deep main characters found in literary works by e.g. Dostojevskij, Tolstoy, or Hemingway |
| Sensible characters                     | Characters with little background detail, but that have a personality that may seem sensible in the ongoing discourse, without providing grounding enough for the observer to evaluate whether the characters behaviour is consistent with the characters pre-events history or not |
| Flat characters                         | Aarseth’s definition                                                                                                                                 |
| Stereotypical characters                | Characters that have cultural roles and act as such e.g. The Butler, The Bartender, The Detective, The Noble, etc                                                                                   |
| Bots, no individual identity            | Aarseth’s definition                                                                                                                                 |

Figure 3 shows that all fully positive coded DGBL systems fall into categories from sensible through deep, rich, round, gravitating to the grounded, consistent category and as such are all on the high author agency end of the ontology, which is linear as is the objects ontology. The grounded,
The consistent category is also the most populated, comprising eight DGBL systems. Three of these are of either no positive effects or of negative learning clusters. Additionally, one of the three DGBL systems in the sensible category shows neutral learning. There are no DGBL systems that use the bots, no individual identity category.

Figure 4 shows that fully positive coded DGBL systems categorize as linear fixed kernels, dynamic satellites, or interchangeable fixed kernels, or dynamic kernels. However, while four of the five DGBL systems that categorize as interchangeable fixed kernels are fully positive, only two of four DGBL systems that categorize as linear fixed kernels are fully positive. All the three DGBL systems that categorize as dynamic kernels are fully positive. All three DGBL systems categorized as fully plotted show either no positive effects on engagement or motivation, or have negative effect on learning. There are no DGBL systems categorized with no kernels, which according to Aarseth (2012) would mean it cannot be a narrative.

### Table 8: The extended events ontology

| Category                                      | Description                                                                 |
|-----------------------------------------------|-----------------------------------------------------------------------------|
| Fully plotted                                 | Aarseth’s definition                                                        |
| N/A                                           | N/A                                                                         |
| Linear fixed kernels, dynamic satellites      | Predetermined kernels that arrive in a fixed sequence from story beginning to end, in which the satellites are dynamic |
| Interchangeable fixed kernels, dynamic satellites | Predetermined kernels that arrive at an interchangeable sequence, in which the satellites may be dynamic |
| Dynamic kernels                               | Aarseth’s definition                                                        |
| No kernels (pure game)                        | Aarseth’s definition                                                        |

### Table 9: The extended ludo narrative variable model (eLNVM)

| Ontology                  | World          | Objects      | Agents          | Events                                      |
|---------------------------|----------------|--------------|-----------------|---------------------------------------------|
| Polarity                  |                |              |                 |                                             |
| Narrative                  | Inaccessible   | Non-interactable | Deep, rich, round | Fully plotted                              |
| High                       | Single room    | Static, usable | Grounded, consistent sensible | N/A                                         |
| author                     | Linear corridor | Modifiable   | Sensible        |                                             |
| agency                     | Multicursal labyrinth | Destructible | Flat            | Linear fixed kernels, dynamic satellites   |
| High                       | Hub shaped quest landscape | Creatable | Stereotypical | Interchangeable fixed kernels, dynamic or fixed satellites |
| game                      | Open landscape | Inventible | Bots, no individual identity | Dynamic kernels |
| Ludic                      |                |              |                 | No kernels (pure game)                      |
| (game) pole                |                |              |                 |                                             |
One result of this study is eLNVM, a common model that allows separate and comparative investigation of the effects that DGBL systems have on learning and the conditions under which it occurs. Fifteen DGBL systems comprise a small dataset, a result of the very strict exclusion criteria of this study; however, none of the DGBL in the studies failed to be categorizable. The original LNVM was introduced using only five commercial games (Aarseth, 2012).

It is possible to observe categorization clustering of fully positive DGBL systems that separates them from all other DGBL systems on the eLNVM, revealing characteristics for fully positive DGBL systems in this research. While all four ontologies show clear categorization clustering of fully positive DGBL systems, only three of them—World, Objects and Events—are also isolated under their plot categories. The Agents ontology shows mixed isolation results.

The second result of this study shows that a narrative DGBL with positive effects on engagement, motivation and learning has the following four characteristics in the eLNVM:

1. A quest-based hub-landscape world, in which confined areas of the full world (hubs) are explored at will, and where reaching objectives in sequences (quests) open new hubs in the landscape and close others, depending on player choices.
2. **Modifiable objects**, that are objects in the world that may be altered, combined, or configured to overcome challenges and reach objectives,

3. **Grounded and consistent actors**, which implies characters in the game narrative with sensible and well described backgrounds, psychologies and goals, and

4. **Interchangeable fixed or dynamic kernels**, which are storyline events of high importance that are either pre-written and interchangeable among each other, allowing for traversing a
story in different ways or that are dynamically written by the system, allowing for development of new stories based on player choices.

### Limitations

Reflection on the study has revealed multiple areas that could be improved.

**Verification of self-assessment in dataset**

This study does not verify the subjective self-assessments of the DGBL systems’ effects on engagement, motivation and learning. This is due to the lack of detailed enough information about the 15 DGBL systems as described in the papers and in the online information about the systems. Furthermore, the evaluations are described with differing foci, different methods, in different depths and variation, and the DGBL systems themselves are often not available in playable form, reproducing and verifying the results would be extremely difficult. In future work we will design and evaluate systems of our own making, and while this study does not have the opportunity to verify self-assessments, the goals of arriving at a model that is usable for inspecting characteristics of DGBL with particular effects remains valid.

**Scarcity of studies that evaluate effects on both engagement and motivation**

While this study explores effects on learning, engagement and motivation, most of the evaluations (9 of 15) fail to evaluate on both engagement and motivation alongside learning. This study, however, still isolates and separates characteristics of DGBL systems with only positive reports. In order to increase the dataset, further study of narrative DGBL under less strict exclusion criteria is warranted.

**Broadening future searches to solidify findings**

While previous literature reviews referenced in this research had some emphasis on narratives, this research defined the search protocol to identify only papers that referenced narratives in the title. This was a conscious decision to narrow the DGBL systems to a dataset with the main emphasis on narratives. Perhaps this search criterion is the reason why these results show a higher percentage of fully positive effect evaluations than what has been shown in previous research. Further research would be required to verify such a claim. Such research may seek to categorize DGBL systems where narratives are named in the abstract, or among the keywords and may even

---

**Figure 4:** Distribution of DGBL system categorizations on the extended ludo narrative variable clustered by evaluated effects, events ontology.

| Categorization                          | Fully positive | Neutral learning | No positive effects | Negative learning |
|----------------------------------------|----------------|------------------|---------------------|-------------------|
| Fully plotted                          |               |                  |                     |                   |
| Linear fixed kernels, dynamic satellites|               |                  |                     |                   |
| Interchangeable fixed kernels, fixed satellites|           |                  |                     |                   |
| Dynamic kernels                         |               |                  |                     |                   |
| No kernels (pure game)                  |               |                  |                     |                   |
inspect DGBL system evaluations where narratives are not mentioned, but where narrative components are identifiable through analyzing the DGBL systems’ qualitative descriptions.

This research found that the LNVM needed refined category naming in one ontology, Agents and separation of category characteristics in another ontology, Events. Thus, the categorized DGBL systems are only partially comparable to the original systems that were categorized by Aarseth (2012). To further nuance results in future research, the eLNVM should be tested and possibly further extended through the categorization of both commercial games and DGBL systems. eLNVM contributes to the debate on isolation and categorization of narrative development in all games, whether they are created for commercial purposes or if they constitute DGBL systems.

LNVM is part of a larger narrative theory of games. For example, ontologies are either ludic or story drivers, stories are separated by degree of emergence and the games can be categorized on a single axis from a ludic to a narrative pole. These, however, are all aspects that may serve to define games as narratives, but only after they have been categorized ontologically. Since our study concerns itself with determining if the LNVM is usable to categorize narratives in DGBL, further inspection was not conducted. Based on the positive indications revealed here, such aspects warrant further study.

**Conclusion**

A review of previous research revealed that DGBL has the potential to positively enforce engagement, motivation and learning in broad populations across various subject matter and 21st century skills. Positive enforcement is particularly evident when applying learning theories, such as constructivism, socio-cultural learning and flow theory. Furthermore, narratives are considered a main game element, which can be introduced in DGBL to strengthen such positive effects. In other instructional contexts, narratives have been shown to increase learning, also by engaging and motivating learners. The field, however, is lacking a common model on which to compare and contrast the characteristics of narratives in DGBL. This research fills this gap by proposing such a model, eLNVM. We extended and showed how Aarseth’s LNVM is a usable tool to isolate and categorize the narrative from other game elements. Fifteen narrative DGBL that have self-reported evidence related to effect of narratives and a system description available to enable categorization on the eLNVM, have been analyzed and categorized. Thus, we provide a common model that allows separate and comparative investigation of the effects that DGBL systems have on learning and the conditions under which it occurs.

By making two extensions to Aarseth’s LNVM, which has implications for work in game studies as well as in studies of DGBL systems, this research has shown that the eLNVM can be used to isolate and categorize narrative DGBL. The study revealed that there are ontological categorization clusters that separate, and to a certain extent isolate, the characteristics of DGBL systems that have positive effects on engagement, motivation and learning from DGBL systems that show less than positive effects. This is an important contribution to current research, since it shows how previous, current and future studies of narratives in DGBL systems can be (re-)categorized and (re-)evaluated, alone and in comparison, using common ontological categorization spanning author to game agency.

In future studies, the eLNVM can be used directly, which has the potential of creating a corpus of studies over time that have been evaluated using the same model, enabling comparison across DGBL systems without further refinement of DGBL system categorization. This requires, however, sufficiently detailed descriptions of the DGBL systems for researchers to be able to plot them using
In future evaluations of DGBL systems, we encourage and recommend the inclusion or direct reference to system descriptions suitable for categorizing the DGBL system. While this study explored effects from narratives in DGBL systems on engagement, motivation, and learning, it is emphasized that the eLNVM is not only useful for evaluation of effects on those three concepts, but also any other concepts or features in DGBL systems under scrutiny.

Finally, this research contributes to the future design of DGBL systems by showing which ontological characteristics that are expected to optimize the potential for positive effects on engagement, motivation and learning. In ongoing work, a process using the eLNVM as basis for efficient design of effective narrative DGBL by multidisciplinary teams is being tested and evaluated.

Acknowledgements
The authors thank the reviewers for invaluable feedback and Nils Petter Hauan for comments on a previous version of this paper. This research is part of an industrial PhD at the University of Bergen and Bergen Science Center (VilVite), funded by the Research Council of Norway project number 277769.

Statements on open data, ethics and conflict of interest
A detailed diary, showing all excluded studies, inclusion/exclusion and a reference list of original searches is available from the first author.

The study adheres to the ethical guidelines monitored by the Norwegian Center for Research Data. VilVite has a mandate to advance science teaching in Norway in cooperation with schools. VilVite is not motivated to affect the research results in any way and the employment there does not constitute a conflict of interest.

References
Aarseth, E. (1997). *Cybertext: Perspectives on ergodic literature*. Baltimore: Johns Hopkins University Press.
Aarseth, E. (2005). From Hunt the Wumpus to EverQuest: Introduction to Quest Theory. Paper presented at the Entertainment Computing—ICEC 2005, Berlin, Heidelberg.
Aarseth, E. (2012). A narrative theory of games. Paper presented at the Proceedings of the International Conference on the Foundations of Digital Games, Raleigh, North Carolina. Retrieved from http://delivery.acm.org/10.1145/2290000/2282365/p129-aarseth.pdf?ip=129.177.94.75%26id=2282365%26acc=ACTIVE%20SERVICE%26key=CDADA77FFDD8BE08%2E5DBA179404D3BE6F%2E4D4702BOC3E3B35%2E4D4702BOC3E3B35%26__acm__=1545225006_ad15b4bf2296c827e438c8ea66b8b5b
Apperley, T. H. (2006). Genre and game studies: Toward a critical approach to video game genres. *Simulation & Gaming*, 37(1), 6–23. https://doi.org/10.1177/1046878105282278
Armstrong, M. B., & Landers, R. N. (2017). An evaluation of gamified training: Using narrative to improve reactions and learning. *Simulation & Gaming*, 48(4), 513–538. https://doi.org/10.1177/1046878117703749
BarthesR., & Duisit L. (1975). An introduction to the structural analysis of narrative. *New Literary History*, 6(2), 237–272. http://dx.doi.org/10.2307/468419
Barab, S. A., Dodge, T., Ingram-Goble, A., Pettyjohn, P., Peppler, K., Volk, C., & Solomou, M. (2010). Pedagogical dramas and transformational play: Narratively rich games for learning. *Mind Culture and Activity*, 17(3), 235–264. https://doi.org/10.1080/10749030903437228
Catalá-López, F., Stevens, A., Garrity, C., & Hutton, B. (2017). Rapid reviews for evidence synthesis. *Medicina Clinica*, 148, 424–428.
Narrative Categorization in DGBL

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. Review of Educational Research, 86(1), 79–122. https://doi.org/10.3102/0034654315582065

Coffey, H. (2009). Digital game-based learning. Learn NC. Retrieved from https://www.dsu.univr.it/documenti/Avviso/all/all036209.pdf

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. Computers & Education, 59(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Courage, M. L., Bakhtiar, A., Fitzpatrick, C., Kenny, S., & Brandaek, K. (2015). Growing up multitasking: The costs and benefits for cognitive development. Developmental Review, 35, 5–41. https://doi.org/10.1016/j.dr.2014.12.002

Csikszentmihalyi, M., & Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience (Vol. 1990). New York: Harper & Row.

Dempsey, J., Lucassen, B., Gilley, W., & Rasmussen, K. (1993). Since Malone’s theory of intrinsically motivating instruction: What’s the score in the gaming literature?. Journal of Educational Technology Systems, 22(2), 173–183.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining “gamification”. Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments, 9–15.

Dettori, G., & Paiva, A. (2009). Narrative learning in technology-enhanced environments. In Technology-enhanced learning (pp. 55–69). Dordrecht: Springer.

Dickey, M. D. (2011). Murder on Grimm Isle: The impact of game narrative design in an educational game-based learning environment. British Journal of Educational Technology, 42(3), 456–69. https://doi.org/10.1111/j.1467-8535.2009.01032.x

Ermi, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. Ferguson, R., Coughlan, T., Egeland, K., Gaved, M., Herodotou, C., Hillaire, G., ... Whitelock, D. (2019). Innovating Pedagogy 2019. The Open University. Retrieved from https://www.learntechlib.org/p/207292/

Fu-Hsing, T., Kuang-Chao, Y., & Hsien-Sheng, H. (2012). Exploring the factors influencing learning effectiveness in digital gamebased learning. Journal of Educational Technology & Society, 15(3), 240–250. www.jstor.org/stable/jedutechsoci.15.3.240

Fulya Eyupoglu, T., & Nietfeld, J. L. (2019). Intrinsic motivation in game-based learning environments. D. Ifenthaler & Y. Kim (Eds.), Game-based assessment revisited, advances in game-based learning (pp. 85–102). Cham: Springer.

Garneli, V., Giannakos, M., & Chorianopoulos, K. (2017). Serious games as a malleable learning medium: The effects of narrative, gameplay, and making on students’ performance and attitudes. British Journal of Educational Technology, 48(3), 842–859. https://doi.org/10.1111/bjet.12455

HabgoodM. P. J., Ainsworth S. E., & Benford S. (2005). Endogenous fantasy and learning in digital games. Simulation & Gaming, 36(4), 483–498. http://dx.doi.org/10.1177/1046878105282276

HershM., & Leporini B. (2018). Editorial: Serious games, education and inclusion for disabled people. British Journal of Educational Technology, 49(4), 587–595. http://dx.doi.org/10.1111/bjet.12650

Jemmali, C., Bunian, S., Mambretti, A., & El-Nasr, M. S. (2018). Educational game design: an empirical study of the effects of narrative. Paper presented at the Proceedings of the 13th International Conference on the Foundations of Digital Games, Malmö, Sweden. Retrieved from http://delivery.acm.org/10.1145/324000/3235783/a34-jemmali.pdf?ip=129.177.94.75%26id=3235783%26a=ACTIVE%20SERV%26key=CDADA77FFFD8BE08%2E5DBA179404D3BE6F%2E4D4702B0C3E8B35%2E4D4702B0C3E8B35%26_acm=_1537206268_7b51768161a95f71b51c4164c730a31

Kitchenham, B. (2007). Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report EBSE-2007-01.

Kitchenham, B., & Brereton, P. (2013). A systematic review of systematic review process research in software engineering. Information and Software Technology, 55(12), 2049–75. https://doi.org/10.1016/j.infsof.2013.07.010
Wang, J. J., Baranowski, T., Lau, P. W. C., Buday, R., & Gao, Y. (2017). Story immersion may be effective in promoting diet and physical activity in Chinese children. *Journal of Nutrition Education and Behavior, 49*(4), 321–329.e321. https://doi.org/10.1016/j.jneb.2017.01.001

Wouters, P., van Oostendorp, H., Boonekamp, R., & van der Spek, E. (2011). The role of game discourse analysis and curiosity in creating engaging and effective serious games by implementing a back story and foreshadowing. *Interacting with Computers, 23*(4), 329–336. https://doi.org/10.1016/j.intcom.2011.05.001

Xu, Z., & Woodruff, E. (2017). Person-centered approach to explore learner’s emotionality in learning within a 3D narrative game. Paper presented at the Proceedings of the Seventh International Learning Analytics &#38; Knowledge Conference, Vancouver, British Columbia, Canada. Retrieved from http://delivery.acm.org/10.1145/3030000/3027432/p439-xu.pdf?ip=129.177.94.75%26id=3027432%26acc=ACTIVE%20SERVICE%26key=CDADA77FFDD8BE08%2E5DBA179404D3BE6F%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%26_acm__=1537206291_a8d963b3f9f2ae96439fb0253c746ee3

Zhang, E., Culbertson, G., Shen, S., & Jung, M. (2018). Utilizing narrative grounding to design storytelling games for creative foreign language production. Paper presented at the Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal QC, Canada. Retrieved from http://delivery.acm.org/10.1145/3180000/3173771/paper197.pdf?ip=129.177.94.75%26id=3173771%26acc=ACTIVE%20SERVICE%26key=CDADA77FFDD8BE08%2E5DBA179404D3BE6F%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%26_acm__=1537206319_d8900227fd75a6ff57c046cbe70a02f4