Validating Video Games as Complex Artifacts

MATTIA BELLINI, University of Tartu, Institute of Cultural Research, Estonia

The narrative sense-making of video games relies on different sources of information. Among them is the multimodal system through which games are made sensorially perceivable, the sensorimotor experiences they afford and require, and the mnemonic recollection through which they are made understandable. These sources show strong feedback loops and give rise to an overall meaning that is more than the sum of its parts. As such, narrative-driven video games can be considered complex in themselves. In order to test this theoretical model, a custom-made video game has been developed. By employing simple mechanics and simple graphics, the game will be the basis for conducting a think-aloud session. The session will give insights into the actual cognitive mechanisms of players, to investigate how their sense-making works.

CCS Concepts: • **Software and its engineering → Interactive games**; • **Applied computing → Computer games**.

Additional Key Words and Phrases: Video games, Interactive Digital Narratives, Representing complexity, Complex systems theory, Social good, Think-aloud

ACM Reference Format:

Mattia Bellini. 2023. Validating Video Games as Complex Artifacts. In *ACM International Conference on Information Technology for Social Good (GoodIT '23), September 6–8, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3582515.3609539

1 INTRODUCTION

Video games are not simple. Salen and Zimmerman, in the influential and holistic view on video games they provide in their Rules of Play, similarly argue that "games are certainly one of the best examples of how entertainment can be far from simple" [13]. The inherent complexity of video games comes from their idiosyncrasies as medium, that is, from their participatory nature and from their audio-visual(-haptic) modes of representation, at the very least.

1.1 Theoretical framework

Video games rely on different sources of information, which need to be synthesized to give rise to the overall more-than-combinatorial meaning. Three of these sources of information have been identified, namely the multimodal system, the sensorimotor experiences resulting from interactivity, and the recollection of background knowledge [2].

The multimodal system [1], [9] is the variably rich system of semiotic resources (audio, visual, and haptic) through which games are made sensorially perceivable. The modes comprising the multimodal system can be seen themselves as exhibiting features of complexity. For instance, in many introductions to film composition are reported several experiments in which spectators of the same scene but with changed background music reported very different overall meaning of the portrayed events. This means that, at least to some degree, the meaning (and in some ways also the story) of the soundtrack is not simply combinatorially added to the meaning of the visual representation, but something more emerges from the encounter of the two. Emergence is one of the aspects that define a system as complex.

The sensorimotor experiences [6] video games afford and require are the result of the participatory nature of digital media, of which video games are one of the manifestations [10]. In video games, acting is required by the computing

Manuscript submitted to ACM

device to generate a sensorially perceivable representation of the fictional environment. In other words, to instantiate a particular narrative experience from the often countless available in the protostory [8], players have to embody their characters and enact their roles within the game world (and story) [7]. In addition, in line with the enactivist approach to human cognition [5], the world of the game is properly understandable (or even perceivable) only in terms of interaction affordances. The sensorial part of sensorimotor experiences is based on the multimodal system: when players perform an action, the game adapts accordingly. In Super Mario, when players press the button to jump, they will see Mario jumping - which means, the game will generate a perceptually available representation of the character jumping, including visual and audio elements, and possibly also vibration (a haptic percept).

However, the very act of Mario jumping would make little sense should the players have no previous experience of jumping (themselves, or seeing others perform the action). This means that video games draw, to be understandable, on players background knowledge. This knowledge comes both in terms of "experiential traces" [16], i.e. experientially-grounded information (such as gravity, jumping), but also in terms of shared schemata (in super Mario, enemies look sentient) [12], scripts (one should jump on enemies) [14], and popular and personal culture (e.g., game conventions on jumping, like the possibility to "double-jump", or previous experiences of jumping in games). This sort of mnemonic recollection, through which games are contextualized [3], [4] and made understandable, configurates therefore as a third source of information in video games. These memories are built and updated dynamically by interacting with the fictional world, through its multimodal presentation. For instance, keeping Super Mario as an example, while one needs no explanation of jumping or gravity, as said, soon one will need also no explanation regarding questionmark boxes. Once players interact with them, their behavior (giving rewards) will become part of the background knowledge of such players. In further encounter of questionmark boxes (within the same game, or in different games in which they appear as "similar enough"), this knowledge could be retrieved and used to understand further interaction, which will build new memories, and so on.

As per the discussion provided so far, the three systems of information show strong feedback loops, they are mutually dependent and they give rise to an overall meaning that is more than the sum of its parts. Therefore, being conveyed through these three systems and relying on them for their meaning constitution, video games can be conceived as complex artifacts [2]. This complex-systemic understanding of videoludic artifacts translates in a holistic rhetoric of video games, forming a theoretical framework useful for analyzing the very medium as a cultural phenomenon, and the idiosyncrasies of its expressive language, in a productive way.

2 RESEARCH HYPOTHESIS AND RESEARCH QUESTIONS

The hypothesis guiding this research is that the overall meaning of a video game is constituted through at least three systems of information: the multimodal system, the sensorimotor experiences, and the recollection of background knowledge. To validate this theory, a think-aloud [15] qualitative study will be conducted, followed by an interview. The specific research question for this experiment therefore is:

- Are video games understood through the complex system of information coming from the multimodal presentation, players' interaction and players' knowledge?

These will be addressed through the following specific research questions:

(1) Are the multimodal system, the interaction and the knowledge mutually interacting?
(2) Are these three systems presenting feedback loops?
(3) Do the three systems give rise to a unique coherent whole (the video game story and the video game in general)?

3 THEORY VALIDATION

3.1 Method

Think-aloud [15] is a technique used in cognitive psychology and user research to understand a person's thought processes as they perform a task or solve a problem. In a think-aloud session, participants are asked to verbalize their thoughts and narrate their actions while they are engaged in a particular activity. Think-aloud is particularly suited for this kind of study as its main objective is to gain insights into the participant's mental processes, decision-making strategies, and problem-solving approaches. By vocalizing their thoughts, participants of a think-aloud session provide researchers with valuable information about their perceptions, interpretations, uncertainties, and reasoning behind their actions.

Think-aloud sessions are often used in usability testing and user experience research to evaluate the effectiveness and efficiency of a product or interface. It is one of the main qualitative research methods in user-centered design. Participants are encouraged to express their impressions, expectations, difficulties, and suggestions while interacting with a website, application, or prototype. This helps researchers identify usability issues, areas for improvement, and potential design flaws.

Approximately 12 think-aloud sessions will be conducted while one participant at a time engages with a custom-made video game, of the approximate length of 4 minutes. This game is currently in development, and has been designed to be relatively simple. The main idea behind this design decision is that by reducing to the very core the elements that comprise the testing game, it could be possible to obtain a "cleaner", i.e. less noisy set of data. With the same idea in mind, given the expectably varied nature of the sources of information participating in video games sensemaking, and the difficult task of looking into one's cognition, it was decided to rely on a group of participants as homogeneous as possible, in terms of ethnicity, age, education and socio-cultural group (while genders will be equally represented). This decision, which could be seen at first glance at a weakness of the study, is driven by the very practical need to reduce to the minimum possible the amount of noise in my data set which, in this specific test, had a high risk of rending opaque my results. It was also decided to look only for participants with significant experience in video games' play. This is due to the fact that it was needed to exclude as much as possible the noise coming from a possible difficulty of dealing both with interaction hardwires (the controller, in this case), and with interaction mechanisms.

The development of the video game, following the design decisions based on the abovementioned abstraction, has been conducted under my supervision by two students of a bachelor degree in computer science and engineering, experienced players themselves. The creation of the game has been therefore carried out by people belonging to the same group as those the game will be tested on, to further minimize the generation of possible noise.

After the think-aloud, participants will be interviewed. The interview will revolve around participants' behavior while interacting with the game, including the reasons for their decisions, and their sense-making mechanisms intended in a broad sense. Participants will be presented with a log of their actions during the play session. The investigator will hold an "expected behavior sheet", in which previsioned behavior is described for each salient moment of the game, according to the theoretical model presented above. Participants will be asked to comment on their choices on the basis of the "expected behavior sheet". Discrepancies between the "expected behavior" and the actual behavior of players as resulting from the gameplay log will be inquired and investigated further. The exact content of the methodology for the interview is still currently under construction.

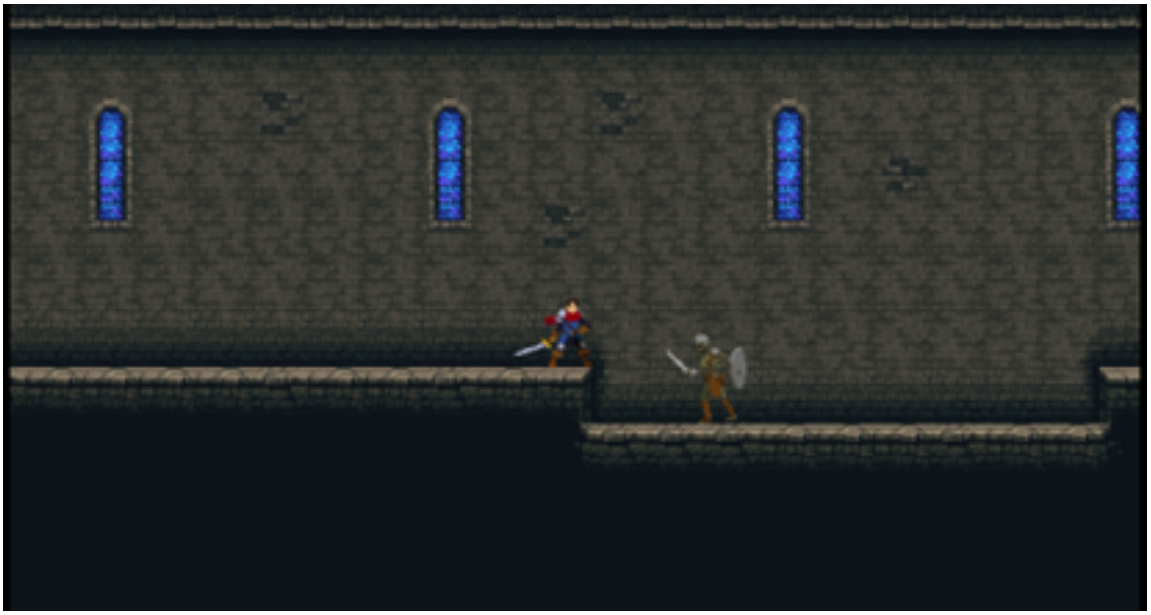

Fig. 1. Screenshot of the video game developed specifically for the current study [All rights reserved]

3.2 Expected findings

"Understanding the connections between the rules and processes of both human and computer systems is vital to the progression of this field, as we strive to create frameworks that effectively produce meaningful, dynamic interactions between players and game systems", write Nacke and colleagues [11]. Pursuing this understanding however is not only productive to advance the field of game studies (and possibly of game production), but also to better understand the affordances of video games for social good.

An implication of my theory is that, due to their reliance on different systems of information, authors can lower the required cognitive load of players and therefore make more easily understandable even complex topics that constitute contemporary social challenges, like climate change and the migration crises. Thanks to the multimodal systems on which they are built, video games can represent a wealth of details and can portray large pieces of representation, due to the multiplicity of perceptions they involve synchronically. Because they rely on sensorimotor experiences, they can adapt and respond in a coherent manner to players' inputs, which enables a deeper and more all-encompassing understanding of the subject matter. This also permits video games to afford replayability, and to present different outcomes from multiple interactions with the same artifact, which further enhance comprehension. Furthermore, thanks to mnemonic recollection, they can rely on previous knowledge of complex topics to aid their comprehension inside the artifacts themselves - even though this is not exclusive of video games.

Empirically proving that all these systems of information participate in an overall integrated understanding of the artifact could be very beneficial in informing designers on how to achieve better, more cognitively efficient representations, to explain in easier ways ever more complex topics (see also the work of the COST Action 18230 – Interactive Narrative Design for Complexity Representation (https://indcor.eu). This could help to effectively inform

the general public on the complexity of the world and on the multiplicity of perspectives and interests that often participates in this complexity. An informed citizen is the absolute prerequisite of effective democracy.

4 FUTURE WORK

After this initial validation, future investigations should expand the pool of participants, both in terms of numbers and in terms of variability of the demographics. The results of such changes could bear extremely interesting results. It could even highlight differences in the sense-making of different groups of people, e.g. in different cultures, different ages, different education or level of expertise with video games.

Additional insights could come also from analyzing in similar ways different video games genres, and even full-fledged games. This sort of study, which would expectably provide much noisier data, could be nonetheless very helpful in pinpointing the workings of these interacting systems of information to the particulars, functioning as an addition to the overall study conducted though my experiment, adding a different level of granularity to it.

ACKNOWLEDGMENTS

This research is supported by the University of Tartu. Sincere thanks to my supervisor, Prof. Dr. Marina Grishakova, and my co-supervisor, Dr. Frank Nack, for the help and guidance. I also thank the COST Action 18230 Interactive Narrative Design for Complexity Representation, and its members for the insightful discussions.

REFERENCES

[1] Elisabetta Adami. 2017. Multimodality. In *The Oxford Handbook of Language and Society*, Ofelia García, Nelson Flores (Linguist), and Massimiliano Spotti (Eds.). Oxford University Press. Google-Books-ID: ugU1DQAAQBAJ.

[2] Mattia Bellini. 2022. Interactive Digital Narratives as Complex Expressive Means. *Frontiers in Virtual Reality* 3 (2022). https://www.frontiersin.org/article/10.3389/frvir.2022.854960

[3] Jerome S. Bruner. 1961. The act of discovery. *Harvard Educational Review* 31 (1961), 21–32. Place: US Publisher: Harvard Education Publishing Group.

[4] Marco Caracciolo. 2014. *The Experientiality of Narrative.* De Gruyter. https://www.degruyter.com/document/doi/10.1515/9783110365658/html Publication Title: The Experientiality of Narrative.

[5] Ezequiel Di Paolo. 2018. The Enactive Conception of Life. In *The Oxford Handbook of 4E Cognition*, Albert Newen, Leon De Bruin, and Shaun Gallagher (Eds.). Oxford University Press, 70–94. https://doi.org/10.1093/oxfordhb/9780198735410.013.4

[6] Marina Grishakova and Maria Poulaki. 2019. Introduction: Narrative Complexity. In *Narrative complexity: cognition, embodiment, evolution*, Marina Grishakova and Maria Poulaki (Eds.). University of Nebraska Press, Lincoln.

[7] H. Jenkins. 2003. Game Design as Narrative Architecture. /paper/Game-Design-as-Narrative-Architecture-Jenkins/f82f061e7a44530d1dee281b96d9b1640485aa74

[8] Hartmut Koenitz. 2015. Towards a Specific Theory of Interactive Digital Narrative. In *Interactive digital narrative: history, theory and practice*, Hartmut Koenitz, Gabriele Ferri, Mads Haahr, Digdem Sezen, and Tonguc Ibrahim Sezen (Eds.). Number 7 in Routledge studies in European communication research and education. Routledge, Taylor & Francis Group, New York. https://doi.org/10.4324/9781315769189-8

[9] Gunther R. Kress. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication.* Taylor & Francis, Milton Park. Google-Books-ID: ihTm_cI58JQC.

[10] Janet H. Murray. 2017. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace.* MIT Press, Cambridge. original-date:1997.

[11] Lennart E. Nacke, Rina R. Wehbe, Samantha N. Stahlke, and Pedro A. Noguiera. 2016. Games of the Heart and Mind. Affective Ludology and the Development of Emotionally Aware Player Experiences. In *Video Games and the Mind: Essays on Cognition, Affect and Emotion*, Bernard Perron and Felix Schröter (Eds.). McFarland, Jefferson, North Carolina.

[12] David E. Rumelhart. 1980. Schemata: The Building Blocks of Cognition. In *Theoretical Issues in Reading Comprehension.* Routledge. Num Pages: 26.

[13] Katie Salen and Eric Zimmerman. 2004. *Rules of Play: Game Design Fundamentals.* MIT Press.

[14] Roger C. Schank and Robert P. Abelson. 1977. *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures* (1st edition ed.). Psychology Press.

[15] Maarten Van Someren, Yvonne F. Barnard, and J. Sandberg. 1994. The think aloud method: a practical approach to modelling cognitive. *London: AcademicPress* 11 (1994), 29–41. Publisher: Citeseer.

[16] Rolf A. Zwaan. 2008. Experiential Traces and Mental Simulations in Language Comprehension. In *Symbols and Embodiment: Debates on Meaning and Cognition*, Manuel de Vega, Arthur Glenberg, and Arthur Graesser (Eds.). Oxford University Press.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009