

Gamified Educational Software for Data Literacy - A Research Through Design Approach to GANDALF

Nikoletta-Zampeta Legaki
zampeta.legaki@tuni.fi
Gamification Group, Tampere
University
Tampere, Finland

Mattia Thibault
mattia.thibault@tuni.fi
Gamification Group, Tampere
University
Tampere, Finland

Juho Hamari
juho.hamari@tuni.fi
Gamification Group, Tampere
University
Tampere, Finland

ABSTRACT

Due to the datafication of our society the necessity to increase data literacy is at an all time high. In this brief paper we present the design of the gamified application GANDALF, aiming to promote data literacy and used in experiment-based research. Following a methodology of research through design, we account for the different phases of the process and frame them as explorative and insightful epistemic stages. Such recount allow us some critical reflection over the design of GANDALF and to formulate five key design considerations that we hope will be useful in the future for the creation of software with similar purposes.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; • **Software and its engineering** → **Designing software**; • **Applied computing** → **Interactive learning environments**.

KEYWORDS

game design, educational games, gamification

ACM Reference Format:

Nikoletta-Zampeta Legaki, Mattia Thibault, and Juho Hamari. 2022. Gamified Educational Software for Data Literacy - A Research Through Design Approach to GANDALF. In *FDG '22: International Conference on the Foundations of Digital Games, September 05–08, 2022, Athens, Greece*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3555858.3555912>

1 INTRODUCTION

The digitization and *datafication* of our society is shedding some light on the necessity to increase the data literacy of the general public. “Reading and writing with data” [1] is a key set of competencies that support individuals in being aware, and understanding the world around them through data. Data literacy is a “set of abilities around the use of data as part of everyday thinking and reasoning for solving real-world problems” [19]. Such a key competence deserves innovative and creative pedagogical efforts, nevertheless there is only a limited number of initiatives and empirical research exploring data literacy pedagogy [11, 15, 20].



This work is licensed under a Creative Commons Attribution International 4.0 License.

FDG '22, September 05–08, 2022, Athens, Greece
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9795-7/22/09.
<https://doi.org/10.1145/3555858.3555912>

In order to address this challenge, and based on studies suggesting the efficacy of gamification as a teaching strategy [12], we have designed GANDALF, a gamified data literacy learning application part of an EU-funded research process. GANDALF is a web-based application, that supports random assignment among four versions i) baseline interactive visual analytics; ii) interactive visual analytics with integrated challenge-based gamification; iii) immersion-based gamification; iv) social-based gamification, according to [21]. GANDALF was used as an experiment setting to test the effects of these different gamification strategies on the learning performances of participants. Preliminary results of the experiments, having students participants are already published [10] and suggest that more than specific strategies of gamification, it was the intrinsic interactivity and playfulness of the application that contributed to the learning success. The use of GANDALF, no matter the version, resulted to a statistically significant improvement in students’ performance of almost 44%. In this paper, instead we will focus on the design process and on what we learned from in in terms of how to design gameful interactive systems to improve data literacy.

2 METHODS

In this paper, we analyse the design process of the GANDALF application from a perspective of *research through design*. This methodology is based on acknowledging the epistemological nature of design activities considered as ways to explore both practical and theoretical dimensions of their fields of application. Research through design provides academic insight, but based on different principles that research through science. According to Gaver ([2]), while research through science is based on strong hypothesis and aims to produce results that can be replicated and generalised, research through design has a more exploratory nature, and proceeds from some nebulous expectations to make something new. The accountability of this epistemology, then, is not rooted in the rigour of the process followed to acquire certain results, but on the *aesthetic qualities* of the final product itself. Research through design has been used as a methodology in the field of games and gamification, where *game design* becomes a nodal point to investigate issues related to the many socio-technical and cultural layers related to games and play [8] or [3].

We believe that a research through design approach to GANDALF will offer additional insights - complementary to those emerged from the experiment - about the challenges and opportunities of designing gamified application for teaching data literacy. In this short paper, we will follow the main design phases outlined by [2] to highlight the main elements and design choices behind our

application. This will allow us to emerge a series of design considerations that can work as general takeaways for the design of gamified applications with similar purposes.

3 DESIGN PROCESS

3.1 Exploring context

GANDALF is a research software whose main objectives are to support civic understanding of data and to evaluate the impact of different individual gamification elements on learning outcomes, using real data in the educational contents. The software was created as part of a EU-funded research project and its core design principles were already formulated in the MSCA-IF proposal description, where they were connected to research questions based on gaps in the literature and research. The theoretical background for the project was articulated around the need for more empirical research in gamification [5, 7], in particular emphasizing the separate impact of different gamification elements [7], and for the exploration of the potential of gamified strategies to increase learners' motivation and support a data-based worldview [17].

The design process was initiated with a thorough exploration of the state of the art. A literature review focusing on existing approaches to gamified pedagogy was supported by a scoping analysis of existing gamified applications and online courses (e.g. on DataCamp, Udemy, Brilliant) related to Exploratory Data Analysis.

The second step was identifying a clear aesthetics and a specific tone for the application. Due to the playful nature of gamification, and in relation to the name of the research project (i.e. Gamification and Datafication towards Learning Forecasting (GANDALF)), we decided to focus on the creation of a character that would serve as representative logo of our research, as well as fictional interlocutor within the application. The design of the logo was the result of input from 10 gamification experts and two artists and it was finalised by a professional design studio.

Having created a basis for kick-starting the application design process, the last step of this stage was the organisation of a brainstorming and co-design session with 7 gamification experts. To facilitate the session we prepared a simple testing prototype of the application using R and Rshiny. The prototype was composed of 5 pages supporting slightly different interfaces around the central concept of presenting an interactive chart with real data, interactive calculators/sliders, and close-ended questions. The participants to the session were informed about the key question (i.e. what individual gamification element we could use to support challenge-, immersion-, and social- based gamification [21]) and were asked to provide their input through a brainstorming map on “mindmeister”. The brainstorming offered several foundational concepts for GANDALF's design, in particular in terms of engagement with the data (e.g., using layman terminology, offering extra guidance and explanations for statistical concepts related to *EDA*), of interaction design (e.g. juiciness, clear guidance), gamification (specific strategies related to the different kinds of gamification implemented) and experiment design (e.g., using follow-up tests, assigning participants randomly to different gamification element or to the control group).

3.2 Developing a design space

In the context exploration phase we accumulated knowledge about existing applications and literature, as well as precious advice and ideas from gamification experts and artists. The next phase was to define the design space for our application, in particular in terms of: i) technical characteristics (programming language based on requirements: web-based, data visualization), ii) educational content, iii) gamification design.

Technical characteristics. After some testing we decided to continue using R and Rshiny. The language worked adequately and, due to the fact that it is mainly used by statisticians for educational purposes, guidelines for reporting and interactive data visualization were easily available.

Educational content. We did not find any standardized test about data literacy or *EDA*, that could be used as a pre-post-test. For this reason, we reviewed the related literature [6, 16, 18] and selected 5 key topics: Central Tendency; Spread of Data; Charts interpretation; Re-expression of Data and Regression.

The topics were to be presented in the application on dedicated pages and in a progression in regard of difficulty. In order to increase the relevance of such topics, we decided to use real data related to societal challenges (e.g. CO2 emissions, life expectancy, death-causes, global population, COVID19- spread, gender-gap) with specific attention to common misconception that could be addressed by engaging the data, or global issues calling for action.

The learning materials would consist in 4-7 questions per topic, each with an interactive chart (and often interactive calculators) helping users to answer it correctly. Data visualisation, hence, was selected as a key component of the educational material.

Gamification design. While the brainstorming session offered several insights in regard of the game design elements to include, their implementation still needed deeper reflection. For this reason we realised a second prototype (more complete content-wise) showcasing several possibilities. A meeting with the project supervisor was then used to discuss how to best integrate gamification within the application. The key takeaways of the meeting regarded in particular the importance of progress and of its evident and clear visualisation through game mechanics and indicators (according to the gamification type: badges, a story board and avatars or competition). Progress, creating a parallel between the gamified experience, the learning path and the progress through the application was meant to offer, at the same time, feedback on the learning performance and motivation to continue in the process.

3.3 Refinement and making

After deciding about the technical characteristics, educational content and gamification design of GANDALF, we started a phase of technical experimentation and an iterative cycle of prototyping and testing different version of the software with different audiences.

The first author of this study, being an experienced developer, statistician and researcher, implemented the first complete version of the application. The software started with the user registration (considering informed consent and data privacy and protection) followed by a pre-test. Users would then be randomly assigned to 4



Figure 1: Motivational affordances, topics and a question from the GANDALF application.

different versions of the application (challenge-based, immersion-based and social-based gamification, plus the control group). The application would finally, present the post-test.

The game elements used in GANDALF (artefacts of stories, avatars, badges and motivational texts about users' ranking) were created in collaboration with the same design studio that created the character / logo of the project, so to maintain aesthetic coherence and to provide high-quality visuals for the gamification design. Figure 1 depicts the representative integrated features for, social-, challenge-, and immersion-based gamification respectively. It also shows some modal buttons that present important topics for the specific page and one of the questions that a participant needs to answer along with an interactive calculator, which helps the participants to understand the topics.

Once the platform was ready, we initiated a pilot phase according to human-computer interaction and gamification design guidelines [4, 9, 14]. The participants of the pilot were divided into two groups: a) gamification research experts ($n=12$, 8 completed the application) and b) economics experts (Ph.D. in Economics or more than 5 years of working experience in statistics) ($n=8$, 6 completed the application). The participants were instructed to give feedback on the content, interface design, and system failures.

Based on the feedback received, we made improvements in the UI, fixed technical issues and updated some of the content. After this phase, the application was ready to be deployed in the experiment.

3.4 Assessment and learning

The GANDALF application was used in experiment-based research. We conducted four-group random assignment pre- post-test online experiments with students ($N=181$) from various courses, schools, and educational levels and with users from MTurk Amazon services ($N=243$). Based on a preliminary assessment of the students results, the overall use of the application resulted in a 44% improvement (baseline and all the gamified versions). Different gamification

strategies did not alter significantly the learning process. As mentioned, the results of the experiment are published elsewhere [12].

3.4.1 Feedback. Due to the length of the application and of the pre- and post-tests, we were unable to ask for a structured feedback of the users' experience – something that would have helped us greatly in assessing the quality of our final design.

Several forms of unsolicited feedback seemed to support a generally positive reception of the experience. We received four unprompted e-mails by student participants proposing improvements or expressing interest in participating in research in the field. While expressing appreciation, the students also mentioned that they found the pre- and post-tests tiring.

Unexpectedly, we also received unsolicited feedback from several MTurk users, which also expressed appreciation for the application.

"I do appreciate the opportunity to work on this project. It was very engaging and somewhat enjoyable. I hope to work with you guys again in the future.";

"Ok!! That was SUPER cool!!! :)".

Interestingly, a MTurk user expressed frustration for a lack of a dedicated space for feedback on several aspects of the application.

"I would have really liked to write about how I felt about each page, because I really felt like I was killing it up until page 4. Then, on page 5, I aced it. I also had thoughts about the pop-up messages in the lower right that I had no way to express"

Finally, teachers involved in administering the software to their students also expressed that the experience was generally described as favorable by their pupils.

3.4.2 Assessment and self-reflection. The design process of GANDALF was long and laborious, but we believe that the final product – while not perfect – was a good example of effective educational software. We believe that, while the experimental gamified elements did not impact significantly the learning performances, the intrinsic playfulness of the interactive data visualisation tools and the lighthearted tone set by the GANDALF character and the

aesthetics of the application contributed to making the experience pleasant for the users and, possibly, to its educational success. An important potential issue that emerged in the design of GANDALF's database was that of data protection. Cautiously safekeeping of users' data (i.e., encrypted and accurate data) was important for being able to give feedback, checking technical/design problems and conduct research. Having the same person responsible for the design and the implementation was crucial to address this issue.

We are also committed to our choice of using real data, connected to important societal challenges and issues in our learning materials. While we have no indication on how they might or might not have affected the users' experiences or learning, we believe it of central importance from a pedagogical perspective aiming at facilitating *transformative learning* [13].

3.4.3 Design considerations. Designing GANDALF was an in-structive process. After the experiment, the designers had a final meeting to discuss their own findings, reflections and considerations around the process. These findings, related to the design of gamified application promoting data literacy, can be articulated in 5 main considerations.

- It is a multidisciplinary and complex process. It is of capital importance to integrate successfully the different elements of the application (learning materials, data visualisation, gamified elements and theme and aesthetics) to provide to users an experience that is both playful and educational.
- The design needs different kinds of expertise. Experts' contribution to the design and feedback was essential. Involving experts early in the process was a key point in the design. Even though the communication between different disciplines (gamification, computer science, economics, semiotics, journalism, engineering and arts) was challenging, the synergy between them allowed us to pool the necessary competences for the creation of a multifaceted application.
- The design needs unity in vision. Ideally, the main designer of the education software should be closely involved in the entire procedure: educational content design; design and integration of gamification; chosen examples or data; artistic design; software implementation and research design. It is paramount to keep the main designer strongly involved in all the process and manage the different aspects but also make the synthesis and the reflection of different inputs.
- The design needs unity of aesthetics. The aesthetic qualities and elements of the application should be clear and harmonious. Setting a tone and a visual regime for the application is important to give an impression of unity and to avoid the disarticulation that could be generated by the compresence of many different elements.
- Progress is a unifying factor. Progress is typical both of games and learning: it should be interwoven with the content and the gamification so to give purpose and direction to the users.

4 CONCLUSIONS

In this brief paper we provided a short account of the aim and of the design process of the gamified application GANDALF, aiming to promote data literacy and used in experiment-based research. Following a methodology of research through design, such process

has helped us formulate five design considerations that we believe useful for the creation of similar software. We hope that our considerations will be useful for future studies focusing on the design of interactive and creative visualization as effective ways to teach, entertain or provide experience and engagement to discover new insights from the rapidly increasing volume of data in a variety of sectors and disciplines.

ACKNOWLEDGMENTS

This work was supported by the Academy of Finland Flagship Programme (337653 Forest-Human-Machine Interplay (UNITE)), the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska Curie, grant agreement No 840809, and the Academy of Finland Centre of Excellence in Game Culture Studies (GameCult, grant 312396).

REFERENCES

- [1] Catherine D'Ignazio and Rahul Bhargava. 2016. DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics* 12, 3 (2016).
- [2] William Gaver. 2014. Science and design: The implications of different forms of accountability. In *Ways of Knowing in HCI*. Springer, 143–165.
- [3] Alan Hook and Paul Coulton. 2017. Games design research through game design practice. In *Game Design Research: An Introduction to Theory and Practice*. ETC Press, 97.
- [4] Karl M Kapp. 2012. *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.
- [5] Jonna Koivisto and Juho Hamari. 2019. The rise of motivational information systems: A review of gamification research. *International Journal of Information Management* 45 (2019), 191–210.
- [6] Matthieu Komorowski, Dominic C Marshall, Justin D Saliccioli, and Yves Crutain. 2016. Exploratory data analysis. *Secondary analysis of electronic health records* (2016), 185–203.
- [7] Richard N Landers, Elena M Auer, Andrew B Collmus, and Michael B Armstrong. 2018. Gamification science, its history and future: Definitions and a research agenda. *Simulation & Gaming* 49, 3 (2018), 315–337.
- [8] Petri Lankoski and Jussi Holopainen. 2017. *Game design research*. ETC Press Pittsburgh, PA.
- [9] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research methods in human-computer interaction*. Morgan Kaufmann.
- [10] Nikolett-Zampeta Legaki, Daniel Fernandez Galeote, and Juho Hamari. 2022. The Impact of different gamification types in the context of data literacy: An online experiment. (2022).
- [11] Nikolett-Zampeta Legaki and Juho Hamari. 2020. Gamification in statistics education: A literature review. (2020).
- [12] Nikolett-Zampeta Legaki, Nannan Xi, Juho Hamari, Kostas Karpouzis, and Vassilios Assimakopoulos. 2020. The effect of challenge-based gamification on learning: An experiment in the context of statistics education. *International journal of human-computer studies* 144 (2020), 102496.
- [13] Jack Mezirow. 2018. Transformative learning theory. In *Contemporary theories of learning*. Routledge, 114–128.
- [14] Benedikt Morschheuser, Lobna Hassan, Karl Werder, and Juho Hamari. 2018. How to design gamification? A method for engineering gamified software. *Information and Software Technology* 95 (2018), 219–237.
- [15] Luci Pangrazio and Julian Sefton-Green. 2020. The social utility of 'data literacy'. *Learning, Media and Technology* 45, 2 (2020), 208–220.
- [16] Ronald K Pearson. 2018. *Exploratory data analysis using R*. CRC Press.
- [17] Hans Rosling. 2019. *Factfulness*. Flammarion.
- [18] John W Tukey et al. 1977. *Exploratory data analysis*. Vol. 2. Reading, Mass.
- [19] Annika Wolff, Daniel Gooch, Jose J Caverio Montaner, Umar Rashid, and Gerd Kortuem. 2016. Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics* 12, 3 (2016).
- [20] Annika Wolff, Michel Wermelinger, and Marian Petre. 2019. Exploring design principles for data literacy activities to support children's inquiries from complex data. *International Journal of Human-Computer Studies* 129 (2019), 41–54.
- [21] Nannan Xi and Juho Hamari. 2019. Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management* 46 (2019), 210–221.