

Triadic Game Evaluation: A Framework for Assessing Games with a Serious Purpose

Casper Harteveld

Delft University of Technology

Jaffalaan 5, 2600 GA

Delft, the Netherlands

+31 1527 81874

c.harteveld@tudelft.nl

ABSTRACT

Evaluatory research for games with a serious purpose is still at its infancy. It especially faces a lack of methodology. To contribute to this gap, this paper puts forward an evaluation framework called “Triadic Game Evaluation (TGE).” This framework stresses that three different perspectives have to be taken into account in the beginning, during, and at the end of the evaluation process.

Keywords

Triadic Game Design, evaluation, methodology, serious games, framework, assessment

INTRODUCTION

The use of games for serious purposes has gained much attention since the beginning of this 21st century [2]. In contrast to the century before, games, and especially digital games, are also applied for a wider variety of domains and purposes. Most institutions nowadays believe games are a *potentially* powerful tool. The question remains, will it reach its potential?

For this to be answered, *evaluatory* research is necessary. By assessing whether a game reaches its serious purpose and in what manner, it can be concluded to what extent games really are such powerful tools. Additionally, we can use the results to improve existing designs and practices. For pursuing this, we hit, however, another question: How do we perform evaluatory game research?

This question cannot be easily answered, because we are facing a *methodological gap* when it comes to evaluating games. This gap can be noted from the huge inconsistency in evaluation studies that have been done so far [3]. Scholars evaluate their games without a structured way of thinking of how to evaluate them.

To contribute to this methodological gap, this paper proposes a framework for assessing games with a serious purpose. This evaluation framework is called “Triadic Game Evaluation (TGE)” and is based on a game design philosophy called “Triadic Game Design (TGD)” [2]. In a nutshell, this philosophy stresses that designing a game

involves three equally important “worlds”: the worlds of Reality, Meaning, and Play. These worlds have to be taken into account while evaluating a game as well and this is what TGE is about.

In discussing TGE further, I will first go more in-depth about what it takes to assess games. Subsequently, TGD is discussed. Following this, a generic description is given of TGE. Before the paper is concluded, I will highlight how this framework is employed for a specific game called *Levee Patroller* [2].

ASSESSING GAMES

Despite the quite long tradition of using games for serious purposes, which can even be traced back to the beginning of human civilization [2], one of the most fundamental questions on this area – does it work? – remains largely unanswered. This is due to a number of reasons.

First, while a long tradition can be observed, academically game research is still at its *infancy*. We are still exploring the medium; we are trying to understand what it is and to what extent it differs from everything else.

Second, games are *complex* objects to study. Games are rich and dynamic environments that consist of many aspects, such as sound and graphics, and are affected by many environmental variables, such as the context in which the game is played. To study this complexity, a better understanding is needed of what variables need to be considered and how they relate to one another.

Third, it is difficult to *proof* the value beyond the game. For some games this is directly visible, but in many cases, if not most, this is quite difficult. For testing the effects in the real world, many intervening variables can play a role. If it is possible to demonstrate any effects, the question remains to what extent all the effort – in terms of design, deployment, and time spent on playing the game – was actually worth it.

Fourth and finally, performing game research is quite *intensive*. It requires a lot of effort to first design a good game. To setup a good evaluation process in the second place, is often too much to be asked for due to budget and time constraints. Most scholars, therefore, stick to the so-called “smile-sheets” to assess their games [6].

A framework will not solve all these problems. It will, on the other hand, help scholars in determining what to do and

LEAVE BLANK THE LAST 2.5 cm (1”) OF THE LEFT
COLUMN ON THE FIRST PAGE FOR THE
COPYRIGHT NOTICE.

it will help in structuring evaluatory game research. For TGE and possibly other game evaluation approaches, a number of first stepping stones have been laid. I will discuss these before I get to the idea of TGE.

The First Stepping Stones

The work of Kriz and Hense [5] is one of the few attempts to build a methodology for evaluating games. Their theory-oriented evaluation has been important for highlighting that when evaluating a game, the game itself should not be seen as a “black box.” Due to the complexity of games, a lot could happen in between. For example, players could get stuck and, as a consequence, get frustrated and stop playing. Therefore, the process, or the “throughput” as they call it, needs to be considered next to the input and output.

Another reason why the work of Kriz and Hense [5] is important is that they made a first attempt to determine what variables play a role in what part of the game cycle and how they relate to each other. For example, they argue that player’s expectations, an input, need to be measured as this influences the throughput and eventually the output. In other words, they mapped some causal relationships between important variables.

Both aspects, the importance of the process and the mapping of variables, are illustrated in the “Design, Play and Experience (DPE)” framework as well [7]. This framework explains that in the game design process three components are relevant that each relate to an aspect of the game cycle: the development of a game by the designer (“Design”), playing the game by the player (“Play”), and the experience the player derives from playing the game (“Experience”). Each of the components is subdivided into layers, like user interface and story, which scholars need to take into account when evaluating a game.

Another stepping stone concerns the four dimensional framework by De Freitas and Oliver [1]. This framework stresses that when evaluating a game, four dimensions are relevant: the context, the pedagogy, the representation, and the learner (or player). For each of these dimensions they developed a number of questions, like “Which pedagogic models and approaches are being used?” and “What level of realism is needed to achieve the learning objectives?”

Although each framework focuses at educational games, they are the first attempts to tackle the methodological gap and are, therefore, important stepping stones for evaluatory game research. With this in mind, I will now turn to TGD, a design philosophy for creating games.

TRIADIC GAME DESIGN

Triadic Game Design is a design philosophy that stresses that three different but equally important “worlds” have to be taken into account when designing a game with a serious purpose [2]. Each world is inhabited by different people, practices, theories, criteria, and so forth.

The first world is “Reality.” Games have always – how abstract they may be – a relationship with the real physical world. For games with a serious purpose this connection is even more important, as in the end, the real world needs to be affected by the game. This world is grounded in the

disciplines related to the subject matter is represented by the different stakeholders, with their expertise and opinions.

The second world is “Meaning.” Although no game can be considered “meaningless,” a more elaborate consideration needs to be made to intentionally achieve a meaningful effect beyond the game experience. The world of Meaning is concerned with this creation of value. In considering this, other disciplines, such as the learning sciences, should be involved and, depending on the purpose, people, like teachers, that know how to achieve a certain value.

The third world is “Play.” Games are first and foremost a tool (or medium). When we think of games, we think of highly interactive and engaging tools that immerse people into a fictive situation. For creating such an environment, aspects, disciplines, and people, like game designers and modelers, need to be involved.

The three worlds are not disconnected from each other. In fact, they are tightly coupled and even overlap in certain areas. The difficult task for designers is to find a synthesis between the three worlds. This task is difficult, because as I noticed, often the three worlds clash with each other on design issues. The synthesis, or a balance, is necessary, since in the end a game with a serious purpose needs to have an effect on the real world, achieve its purpose, and be fun and engaging to play.

While this philosophy is specific to game design, the ideas can easily form the basis of an evaluation framework. As the DPE framework points out [7], the design of a game influences how it is played and, eventually, what results are achieved. This means that if the principles of TGD are used to design a game, it can (and should) also be tested whether the criteria of all three worlds are achieved. With the development of TGE such an evaluation becomes possible.

TRIADIC GAME EVALUATION

Based on the first stepping stones and TGD an evaluation framework has been developed that involves a way of thinking and a way of working.

The Way of Thinking

Similar to TGD, TGE considers the worlds of Reality, Meaning, and Play. It asserts that the aspects and criteria of each world need to be part of the evaluation process to get strong, reliable, and interpretable results. The reason for this is clear-cut: if all three worlds play an equally important role for creating a good game, variables from each world affect the eventual experience, and thus the results.

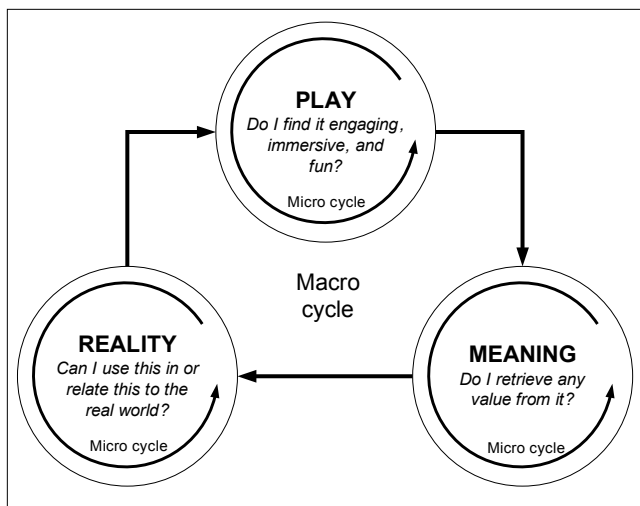


Figure 1. The sensemaking cycle

This means that when evaluating a game, aspects related to each of the worlds have to be looked into. These aspects can be decomposed into as many variables that the researchers find relevant for their evaluation. In doing this, the criteria of each world can be helpful. For example, for Reality the criteria could concern flexibility, validity, and fidelity; for Meaning motivation, relevance, and transfer; and for Play engagement, immersion, and fun.

Similar to De Freitas and Oliver [1], TGE also stresses that the player needs to be part of the evaluation. Unlike their framework, with TGE the player is not a separate dimension. The player can be seen as part of each of the three worlds as he or she can be conceived of having three different roles:

- *Player as person*: the player is a person in the real world. He or she has demographics, a personality, attitudes, and so on, that could affect how the game is experienced.
- *Player as interpreter (or learner)*: people interpret information differently and so do players. This depends, for example, on the existing knowledge, education, learning styles, and expectations.
- *Player as player*: players differ amongst each other. This means that amongst other things previous experience with games as well as game preferences can make a difference in the results.

What researchers want to consider depends very much on their game and the context in which they apply it. They only have to make sure that they structure their evaluation on the basis of Reality, Meaning, and Play.

The Way of Working

Similar to Kriz and Hense [5] and Winn [7], TGE insists that for evaluating a game, it is important to look into what happens “in-between.” It has a very specific reason for this. When playing a game, players go through what I call a “sensemaking cycle” (Fig. 1). This cycle is reminiscent of Kolb’s learning cycle [4]. It says that players continuously are “making sense” of their play experience by questioning and/or feeling consciously and/or unconsciously whether

they find the game engaging, immersive and fun, whether they learn from it or derive some other specific value from it, and whether this is something they can use or relate to in the real world.

The sensemaking cycle consists of three “micro cycles” and one large “macro cycle”. The micro cycles are tied to each world. It could be that the micro cycle in itself is enough for players to continue playing. This happens with most of the entertainment games. The macro cycle connects the three worlds and is critical for the success of a game.

The cycle happens iteratively and continuously. As a researcher it would be valuable to capture this process as the eventual results become better interpretable. To do this, methods should be employed that are able to retrieve insights intermittently. This should be done with care, because the methods themselves can severely disrupt the sensemaking cycle.

As for the methods in general, TGE does not prescribe any particular ones, but as long as no clear relationship exist, it is important to use various qualitative methods. It further recommends logging the game data. Game data enables to get a clearer picture of the play experience.

THE CASE OF LEVEE PATROLLER

To make the idea of TGE concrete, I will show how it is applied to a game I designed. Before I discuss the setup of the evaluation, I will first explain what the game is about.

The Game

Levee Patroller can be described as a “single-player 3-D first person game.” This means the game is solely played by one user from the perspective of the player character. In the game, the player’s role is that of a “levee patroller.” These are people who inspect “levees” regularly or in cases of emergency. Levees are the natural and artificial barriers that protect the land from flooding. Inspecting these levees is of high importance to the Netherlands due to the high risks involved in a possible levee failure. Therefore, it was desired that patrollers are able to increase their inspection knowledge and skills. This is difficult to achieve in reality, because failures occur rarely. To make it possible for patrollers to get experience, it was decided to develop a game.

In the eventual game, the goal is to find every virtual failure (see Fig. 2 for an example). After finding a failure, players need to fill out a report, possibly measure the failure, and contact the coordinating field office to discuss the severity of the problem. If players do well, they get a high score. If they do not do well, they get a low score. Additionally, a levee breach that floods the whole region may be the result.

The Setup

For the evaluatory study of *Levee Patroller*, I set up a training of three weeks. At the start and the end of the training a short meeting takes place. Aside from filling out the questionnaires, the participants play one exercise at both meetings. The start meeting further includes an introduction to the game, and the end meeting includes a

discussion. In between the meetings, the participants play the game at home. They are required to play two exercises per week.

With this setup in mind, I can explain some of the methods I am using for this study and how they relate to TGE.

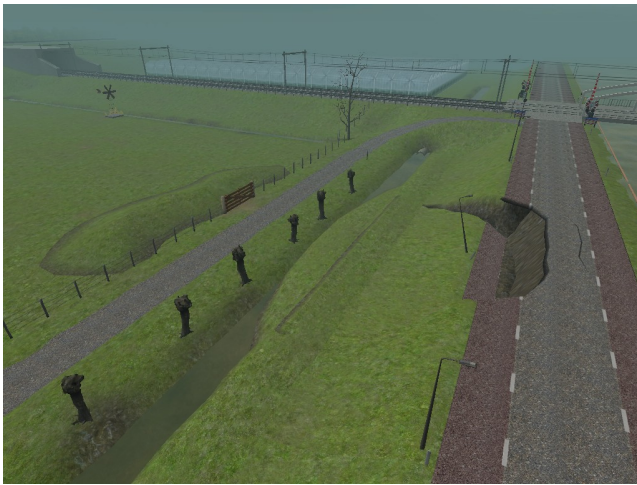


Figure 2. A screen shot of a levee failure

Pre- and post-questionnaire

The questionnaires are the prime method in this study. During the pre-questionnaire the background variables that relate to “player as person, interpreter, and player” are first asked. For example, one of the questions asks whether the participant has any experience with First Person Shooters (FPSs), a game genre very similar to *Levee Patroller*.

The post-questionnaire goes into whether the game fulfilled the expectations and deals with the design. The game has, for instance, no sound and one of the questions deals with whether participants missed this. This part deals with the world of Play.

Both questionnaires end with two parts that relate to the worlds of Reality and especially Meaning. One part deals with the perceptions the participants have about levee inspection and the other part consists of real and virtual pictures. For this latter part, participants need to write down what they see and how they would deal with it.

Game log

Every action, from accessing a help menu to measuring a failure, is logged. As such, it becomes possible to analyze how participants play the game. Such analysis could be helpful to improve the design, but is also helpful in understanding how participants go through the sensemaking cycle. The game logs are sent to a central server over the Internet. As a backup, the logs are also saved locally.

In-game evaluation

After each exercise, participants have to fill out a short evaluation form, which is based on the worlds of Reality, Meaning, and Play. This way, it is attempted to capture the sensemaking cycle as well. Similar to the game logs, these forms are saved locally and sent to a central server over the Internet.

CONCLUSION

Evaluatory research for games with a serious purpose is still at its infancy. Some initial work has been done, but we are still exploring how to assess these types of games. In this paper, I put forward a framework called Triadic Game Evaluation (TGE), which can be used for such evaluations. I illustrated its use with the game *Levee Patroller*.

The framework is still under development and the results with *Levee Patroller* will highlight to what extent it is helpful to consistently take the worlds of Reality, Meaning, and Play into account. Additionally, it will become clear what variables of each world matter in this regard. With these insights, the framework can be further elaborated.

ACKNOWLEDGMENTS

I thank Deltares, a research institute for water, soil, and subsurface issues, for supporting my research. I also thank everybody who has contributed to *Levee Patroller*.

REFERENCES

1. De Freitas, S., and Oliver, M. How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & Education* 46, 3 (2006), 249-264.
2. Hartevelde, C. *Triadic Game Design: Balancing Reality, Meaning and Play*. Springer, London, UK, 2010.
3. Ke, F. A qualitative meta-analysis of computer games as learning tools, in R.E. Ferdig (ed.), *Handbook of Research on Effective Electronic Gaming in Education* (Vol. I, pp. 1-32), Information Science Reference, Hershey PA, 2009.
4. Kolb, D.A. *Experiential learning: Experience as the source of learning and development*. Prentice Hall, Upper Saddle River NJ, 1984.
5. Kriz, W.C., and Hense, J.U. Theory-oriented for the design of and research in gaming and simulation. *Simulation & Gaming* 37, 2 (2006), 268-283.
6. Prensky, M. *Digital game-based learning*. McGraw-Hill, New York NY, 2001.
7. Winn, B.M. The Design, Play, and Experience framework, in R.E. Ferdig (ed.), *Handbook of Research on Effective Electronic Gaming in Education* (Vol. III, pp. 1010-1024), Information Science Reference, Hershey PA, 2009.