Master Thesis Business Informatics

Gamified tutorials: the case of operations research

Master Thesis
12 January, 2015

Author:
Horia Constantin
h.constantin@students.uu.nl
Master Business Informatics
Utrecht University

Supervisors:
dr. F. Dalpiaz (Utrecht University)
dr. ir. J.M. van den Akker (Utrecht University)
J.W. van Crevel (AIMMS B.V.)

Contents

Acknowledgements

In the development of this research project, besides all the people who have helped me indirectly, to whom I am deeply thankful, there are a couple of persons who contributed directly (with effort or time) that I wish to acknowledge.

Fabiano Dalpiaz, for being the best first coordinator that anyone can have. Thank you for guiding me all the way through this project: making sure that the project runs according to schedule, offering a huge amount of timely feedback, supporting me with the organization of the experiment, and, in general, being there for me whenever I needed help.

Jan Willem van Crevel for encouraging me to finish, keeping me focused and offering a lot of support in a very busy period of my life.

Marjan van den Akker, for providing valuable feedback on the OR side of the project and, by doing that, giving me the confidence to pursue a subject which was unfamiliar to me.

Koos Heerink, Marcel Hunting, and Ovidiu Listes, for being part of the expert panel and supporting me in understanding the basics of OR.

Mischa Bronstring, Marcel Roelofs, and Chris Kuip, for sharing their expertise and coming up with ideas that helped me create the non-gamified tutorial faster.

Rick Barneveld, Rares Sfirlogea, and Rick Hoving, for thoroughly testing both versions of the tutorial.

Gabriel Slot, for creating the icons for the badges used in the gamified tutorial

All the participants to the experiment. Without their help I would not have been able to validate the artifact that was generated by research.

1 Introduction

1.1 What is operations research (OR)?

Operations research (OR) (alternatively, operational research) is loosely defined as the discipline that deals with the application of advanced analytical methods to help make better decisions[1]. It is deeply rooted in mathematics, as it uses statistical analysis and mathematical optimization techniques. As a separate technology, OR emerged in 1937 when it was used to ensure the operational efficiency of the newly invented radar system (Cunningham, 1984). It was used industriously during the Second World War (Shrader, 2008); for example, it increased flight hours by 61 percent; by modifying the convoy system, it reduced Allies shipping losses in the Atlantic by more than a third.

OR, because of the limited number of experts and high costs of the required hardware, was originally available only to a selected few (e.g., the military). However, lower costs of computing power, availability of big data and an increase in the number of OR experts have recently made this technology available to a wider audience (Sashihara, 2011). Being rooted in mathematics, the benefits of OR are clear and measurable, for OR enables identifying better (faster, less expensive, less risky) plans to achieve an organization's objectives. The Franz Edelman Award[2], a competition that attests the contributions of OR and analytics in both the profit and non-profit sectors, was established in 1972. All the participants have achieved notable benefits through OR; some of the winners are the following:

- TNT Express saved 207 million euros and reduced CO_2 emissions by 283 million kilograms over the period 2008–2011 (Fleuren & Goossens, 2013). This was achieved by using three subprograms:
    - TRANS, at the operational level, enabled automatic network optimization analysis and quick analysis in schedule changes. The change of TNT Express analysts from spreadsheets to OR models allowed the analysis of the transport network and the allocation of workload towards improving their way of working.
    - SHORTREC, at the operational level, offered improvements in terms of ability to cope with volume fluctuations of pickup and delivery of packages, thus offering better service levels.
    - DELTA Supply Chain, at the strategic level, led to the decommissioning of 12 airports, the opening of a new one and to the elimination of six aircrafts. All this, with minimal impact on service levels. The program is also used by the board to develop the Vision 2015 strategy
- Netherlands Railways (NS) achieved an additional profit of 10 million euros in the first year (2007) of introducing OR based train scheduling. An additional annual profit of about 20 million euros came in 2008, due to increased average punctuality level (Kroon, Huisman, & Abbink, 2009). This was achieved by changes in schedules for rolling-stock and crews, which are major cost sources for NS.
- Midwest Independent Transmission System Operator, Inc. (MISO) has saved from 2007 through 2010 between $2.1 billion and $3.0 billion. (Carlson, Chen, Hong, & Jones, 2012). This was achieved by applying OR to define energy output levels and to establish the prices at which energy traded, as well as to determine when each power plant should be on or off.

[1] https://www.informs.org/About-INFORMS/What-is-Operations-Research
[2] https://www.informs.org/Recognize-Excellence/Franz-Edelman-Award

Implementing OR requires an understanding of the problem at hand and creating a mathematical model that describes the problem. Once the mathematical model has been created, statistical analysis, mathematical optimization and other specific techniques are used to solve it, gain insights from the results and further improve it (Bisschop, 2006). The author states that this flow (Figure 1) should not be seen as strictly sequential since some steps might be repeated, and the process could be iterated.

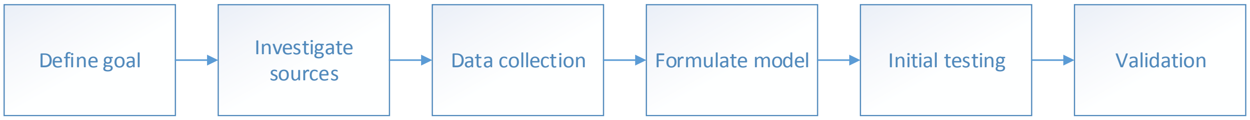


Figure 1 The modeling process (adapted from (Bisschop, 2006)*)*

In the *define goal* stage, the business problem is defined, analyzed and it is decided which of its aspects are relevant to its solution. The emphasis is on problem definition, rather than on mathematics.

In the *investigate sources* stage, the literature and industry experts are consulted in regard to the existence of a similar model.

In the *data collection* stage, the data required by the model is prepared for use. This is a tedious task because, in the worst case, data is not available or, in general, data needs to be transformed to fit the model needs.

In the *formulate model* stage, the optimization model is built. Unless a similar model has been found, this is a creative task since "there are no clear rules for making a choice".

In the *initial testing stage*, the initial (not finalized) versions of the model are verified to confirm the formulation of the model is correct.

In the *validation* stage, the results that the fully developed model produces are compared with already known situations. At minimum, the former should be at least as good as the latter.

1.1.1 Obstacles to OR adoption

Even with the potential to obtain clear benefits that is shown by the presented case studies, OR is still not widely adopted by businesses, not even in fields where the benefits would be immediate, such as supply chain management (Sashihara, 2011) (Fenn, 2013). One of the reasons why this is happening is that managers are not aware of OR and don't understand when and where OR can be applied effectively. Another reason for low adoption is that OR is rooted in advanced mathematics, making it inaccesible to some users. As a result, people are skeptical about the solutions it offers when applied to a specific problem (Sashihara, 2011).

OR can be contextualized in the broader scope of Business Analytics, which distinguishes between three main types of analyses: descriptive (reporting), predictive and prescriptive (Sharda, Asamoah, & Ponna, 2013), (Delen & Demirkan, 2013)).

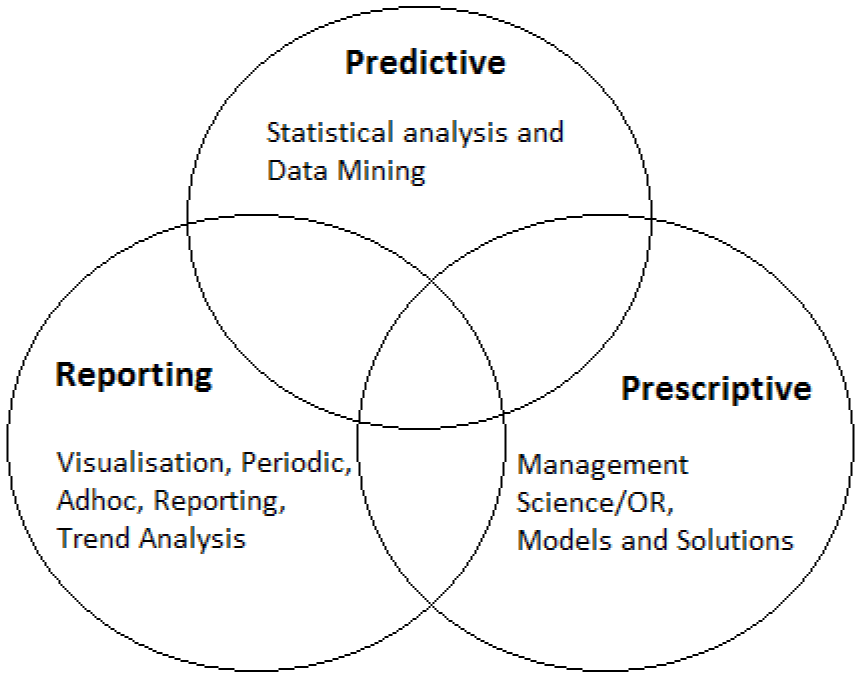


Figure 2 Types of business analytics (adapted from (Sharda et al., 2013)*)*

Descriptive analytics answer questions like: What happened/is happening in the organization? They are made possible by consolidating and aggregating all the data of the organization in a single place, business reporting and visualization tools (e.g. dashboards) (Delen & Demirkan, 2013).

Predictive analytics answer questions like: What will happen? Why will it happen? They are made possible through various mining techniques (e.g. data, text, web) and statistical forecasting (Delen & Demirkan, 2013).

Prescriptive analytics answer questions like: What should I do? Why should I do it? They are made possible through operations research, simulation modeling, decision modeling and expert systems (Delen & Demirkan, 2013).

Linden (2014) places business analytics related techniques such as prescriptive analytics ("a set of analytical capabilities that specify a preferred course of action") and optimization (a type of prescriptive analytics that uses mathematical algorithms to choose the "best" alternative(s) that meet specified objectives and constraints) in the Innovation Trigger of the Hype Cycle, close to machine learning ("a discipline that allows software components to be synthesized from data without being explicitly programmed"), and deep learning ("an increasingly popular variant of neural nets, with more than the typical two processing layers"). Their argument for this positioning is that, while optimization has existed for years in logistics, supply chain and manufacturing, it is starting to be broader used in business (e.g. customer next best offer, pricing optimization, call center agent assignment). Past usage was focused on finding the "best answer", which met specified business objectives and constraints. Newer applications focus more on the decision-making process, where the insights come less from finding the best answer, and more from the ability to do "what if" simulations, explore alternative decisions and understand trade-offs. (Linden, 2014)

A professor of Econometrics and Operations Research at a Dutch University (with more than 20 years of experience in the OR domain) shared the following situation during the scoping interviews for this thesis: One of the companies he was involved with, after going through a process of change, eventually embraced OR in their daily business operations and obtained clear profits as a consequence. After some time, the management team changed and the new one, even with the in-house proof that OR brings business value, did not accept OR as a profit-enhancing practice. Currently, the professor is again going through the process of changing the management team perception in regards to OR, with uncertain results, despite the previous success.

1.1.2 Diffusion of innovations

"An innovation is an idea, practice, or object that is perceived as new by an individual or other unit of adoption" (Rogers, 2010). Rogers makes an important remark: the characteristic of "new" is subjective. To expand that remark: an idea does not necessarily have to be new to be an innovation, it is enough that it is new in the perception of the unit of adoption or that the unit of adoption has not formulated an opinion about it.

As we have stated in the previous paragraphs, OR has been in use for some time. In industries that deal with very expensive assets it has long been adopted as a necessary business practice (e.g., fleet management, plane loads and prices by airline operators) (Sashihara, 2011). However, this is not the case for other domains, although the potential for savings and higher efficiency exist. Under these circumstances, we put forward the idea that, for most industries, OR is an innovation that is at its very early stages of adoption.

Rogers (2010) proposes an innovation decision process, shown in Figure 3, that describes the five stages that lead a decision-making unit (i.e. an individual or an organization) to adopt or reject an innovation.

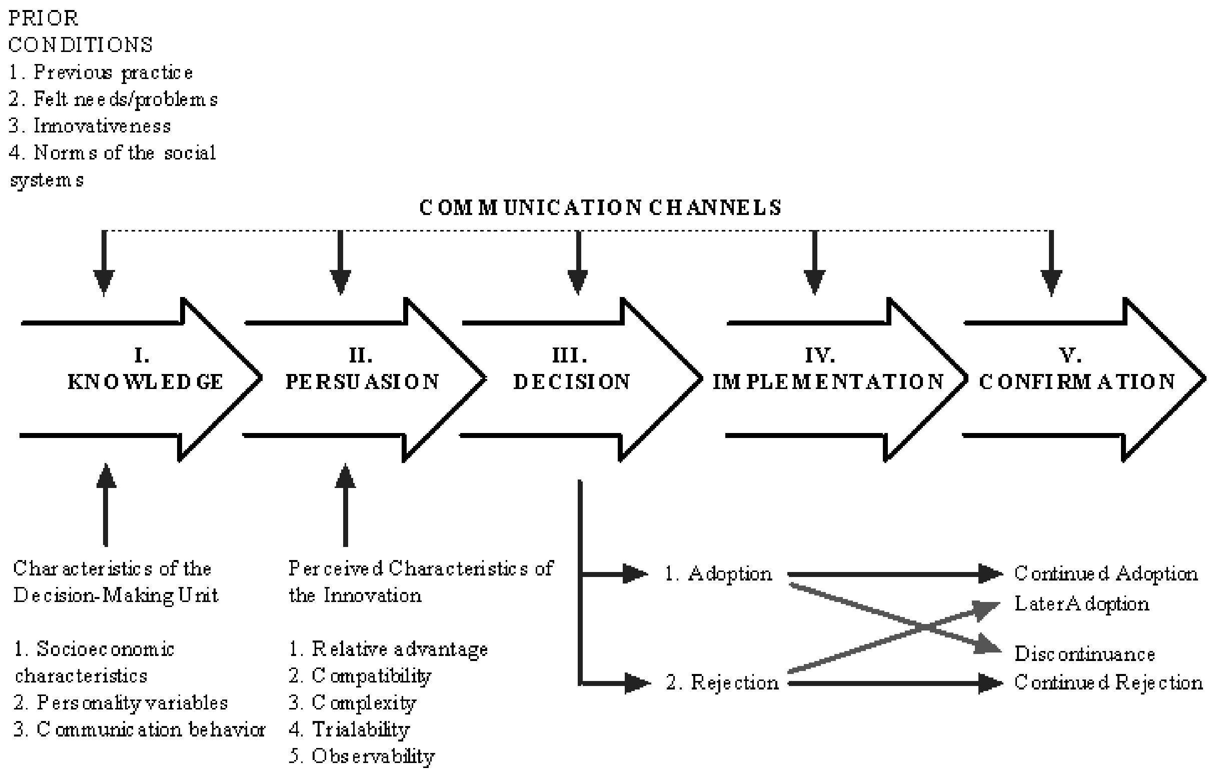


Figure 3 A Model of Stages in the Innovation-Decision Process, taken from (Rogers, 2010)

Knowledge is transferred when the decision-making unit is exposed to an innovation's existence and gains an understanding of how it functions. *Persuasion* occurs when the decision-making unit forms a favorable or an unfavorable attitude towards the innovation. Whereas the mental activity at the knowledge stage is mainly cognitive, the main type of thinking at the persuasion stage is affective. An individual seeks innovation evaluation information, messages that reduce uncertainty about an innovation's expected consequences. *Decision* takes place when the decision-making unit engages in activities that lead to a choice to adopt or reject the innovation. *Implementation* occurs when the decision-making unit puts a new idea into use. *Confirmation* takes place when an individual seeks reinforcement of an innovation-decision already made, but he or she may reverse this previous decision if exposed to conflicting messages about the innovation. (Rogers, 2010)

The speed at which a decision-making unit goes through this process depends on various factors that are represented with numbered lists in the model. The speed at which a target population for a specific innovation goes through this process is known as the rate of the adoption of the innovation. It makes sense to note that the same innovation can have different adoption rates for different populations. The adoption rate is also influenced by the penetration of the innovation in the target population. (Rogers, 2010)

By analyzing this model, we can conclude that, currently, the adoption of OR is in the first stage, *knowledge*. In this stage, the decision-making unit is exposed to the innovation's existence and gains an understanding of how it functions. The knowledge stage consists of *awareness-knowledge* (information that an innovation exists), *how-to knowledge* (information necessary to use an innovation properly) and *principles-knowledge* (information about the functioning principles underlying how an innovation works). When an adequate level of how-to knowledge is not obtained prior to the trial and adoption of an innovation, rejection and discontinuance are likely to result (Rogers, 2010). Since the knowledge step is so fundamental to the adoption of an innovation, we will not be investigating what are the other reasons for which OR is not being adopted by the industry. The OR community has been trying to increase the adoption rate of OR by establishing learning communities (Fleuren & Goossens, 2013) or by presenting a strategy of implementing OR in an organization (Sashihara, 2011). These initiatives are, until today, not able to help OR pass the gap between early adopters and early majority, as confirmed by interviews at AIMMS (a technological provider for OR).

1.1.3 Case study – AIMMS

AIMMS is a company that has been operating in the OR domain for 25 years. Their product, with the same name, is a software system designed for modeling and solving large-scale optimization and scheduling-type problems. It consists of an algebraic modeling language, an integrated development environment for both editing models and creating a graphical user interface around these models, and a graphical end-user environment (Figure 4). It provides developers (optimization specialists) with an interface for model building, data modeling and graphical user interface creation.

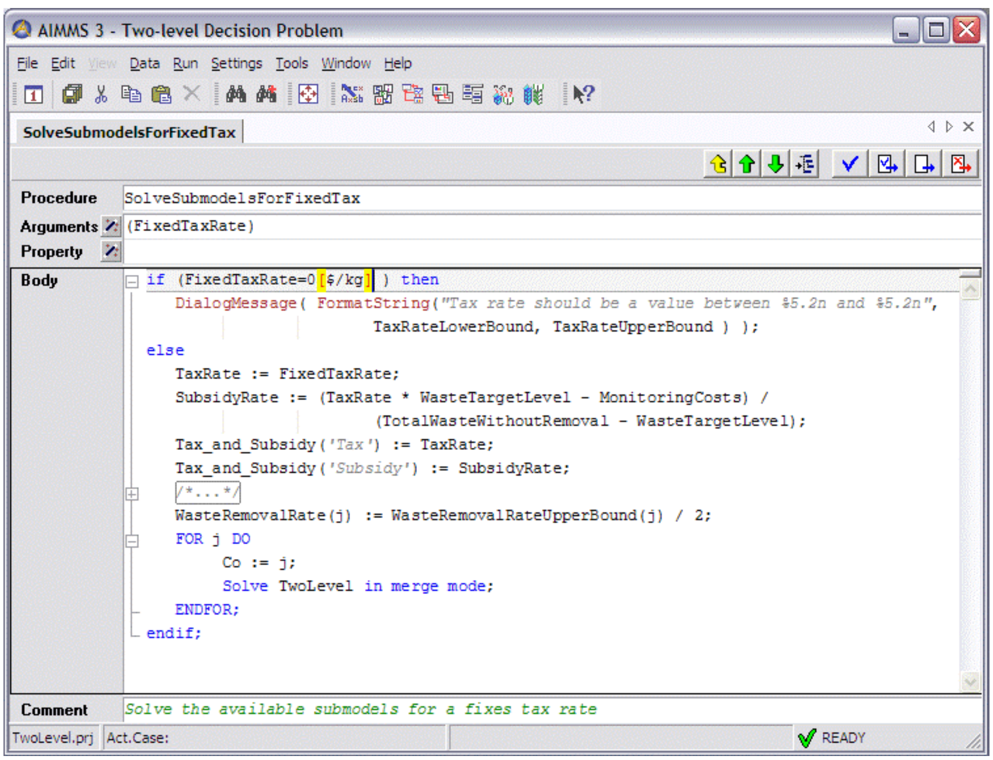


Figure 4 AIMMS IDE

After an OR application has been developed (Figure 5), it is delivered to end-users (supply chain managers, production planners etc.) and they make use of it in their daily operations (at strategic, tactical or operational levels).

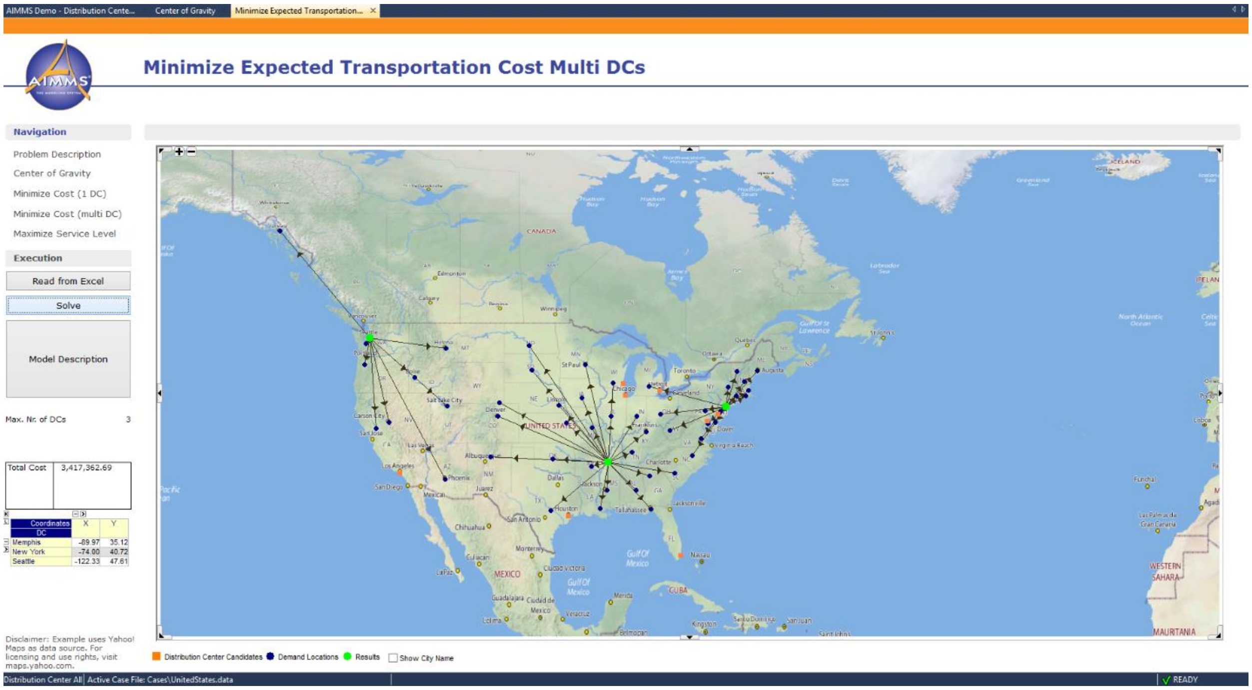


Figure 5 End-user application

1.2 Gamification

An emerging discipline that brings elements of games into business could be used for the purpose of making the OR innovation more appealing to users. Gamification, "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled, & Nacke, 2011), promises to help increase engagement in internal, external or behavior change contexts (Werbach & Hunter, 2012). The scientific literature is still not well established around this subject and, in particular, there is little known about the potential that gamification has to improve the adoption rate of a new technology. In other words, there is no scientific research for the non-game context of new technology adoption. According to the definition, it seems that gamification can bring benefits to any context, but it was shown by (Hamari, Koivisto, & Sarsa, 2014) that the benefits depend on the context to which gamification is applied. This research project studies whether gamification is well suited to the context of teaching OR and how gamification techniques can be employed in this context.

Gamification, in the context of learning, has been applied before for other disciplines. Li, Grossman, and Fitzmaurice, (2012) have created a gamified, in-product, interactive tutorial system for first time AutoCAD users. GamiCAD uses the following elements of gamification: fantasy, clear goals, feedback and guidance, progressive disclosure, time pressure, rewards and stimuli, to provide how-to knowledge. Dong, Dontcheva, and Joseph, (2012) have created a gamified application (named Jigsaw[3]) that teaches players how to use the basic functionality of Adobe Photoshop. Jigsaw also uses of the following elements of gamification: provides a clear goal, allows multiple paths to success, enables structured and guided exploration, gives feedback on progress and makes the application engaging to enable discovery-

[3] http://raschin.com/blog/?p=1431 (starting at minute 5:30)

based learning. Both these applications of gamification are the first of their kind in the fields of computer-aided design, respectively image editing.

1.3 Research question

The research question of this thesis is formulated as follows:

What role can gamification play in teaching operations research to potential new adopters?

This is further broken down into research sub questions:

1. *RSQ1: What are the intended learning outcomes of operations research for potential new adopters?*
   Operations research is a discipline with a broad reach that can and has been applied to various problem domains. The academic courses that teach operations research start by introducing students to different aspects of mathematics on which operations research relies. The challenge here is to find the basic principles that allow the learners to get started right away with using OR and to understand the concepts behind it.

2. *RSQ2: What are the guidelines for creating a digital interactive tutorial?*
   The role of this sub question is to establish the knowledge base that will enable us to build the application that can be used in teaching operations research to potential new adopters. We will find a proper definition from the Human Computer Interaction (HCI) field and establish the status of the research.

3. *RSQ3: What aspects of gamification can be applied on top of a digital interactive tutorial to increase its effectiveness?*
   As mentioned before, gamification uses game elements in its implementation. But gamification is still a young field at it is still unclear, out of the possible game elements, which ones are relevant to specific contexts. Barwood & Falstein (2006) have identified 400 game elements and their list is not exhaustive. Game elements, points, leaderboards and badges seem to be the most heavily used, but there are others that have been successfully used in gamified applications (Hamari et al., 2014). In addition to this, it appears that game elements have different conceptual levels (e.g., dynamics, mechanics and components) (Werbach & Hunter, 2012). Answering this question will clarify what aspects from games are most suited for the context of teaching operations research.

4. *RSQ4: How is a 'gamified digital interactive tutorial for operations research' designed?*
   Having found the basic principles of OR, the guidelines for designing digital interactive tutorials and the gamification guidelines, we are still a long way from a working gamified tutorial. During its design and implementation we expect to come across and solve various obstacles that we cannot foresee at this moment. The decisions made in the process will be recorded and will be used to answer this question.

5. *RSQ5: What are the benefits brought by gamification to a tutorial for operations research?*
   In order to find out whether gamification contributes or hinders the learning of operations research, we will compare the 'gamified digital interactive tutorial for operations research' with

its non-gamified version. We will measure aspects of the learners' behavior and the acquired operations research knowledge. This question will be answered by analyzing the collected data.

1.4 Research relevance

1.4.1 Business relevance

Business analytics/business intelligence, as a technology category, is currently seen as a top priority by C-level executives in various industries (Kalakota, 2011)(Elliott, 2012)(Kalakota, 2012)(Chen, Chiang, & Storey, 2012). Furthermore, companies are collecting a lot of data about their operations (colloquially known as "big data") and this data drives the need to apply analytical techniques to make sense of it (Harris, 2012)(Swoyer, n.d.). This technology category promises better business figures as a result of better/more informed decision making (Davenport, 2006)(Trkman, McCormack, de Oliveira, & Ladeira, 2010). The newest and most technically advanced category of business analytics, prescriptive analytics, is still in a low adoption state. To fulfill that promise, however, they need to be integrated in the organizational decision process. And for that to be possible, it is required that the decision makers understand how these technologies impact the existing decision process (Kowalczyk, Buxmann, & Besier, 2013). One of the prevalent prescriptive analytics techniques, operations research (OR), promises to enable companies to make better decisions (Sashihara, 2011).

We have identified a knowledge gap in the OR field: most of the available OR tutorials assume that the learners are familiarized with concepts of mathematical modeling. This means that, for a person that never encountered operations research and has only basic math skills (can use spreadsheets at a medium level), the available tutorials will be very challenging at best. Unfortunately, these are usually the persons (the decision makers) that need to be convinced of the usefulness of OR (Sashihara, 2011). Additionally, these people, because of the nature of their work, have a tendency to dismiss unfamiliar concepts that are not explained well enough.

This project will aim to use gamification to enhance the OR understanding of decision makers (thus increasing their how-to and principles knowledge). In other words, the objective is to give decision makers a flavor of what OR is or to trigger their curiosity in a fun way.

1.4.2 Social relevance

Beyond business relevance, OR can have a significant societal impact. Hutton, Brandeau, & So (2011) have employed OR to assist policy makers' decisions in improving health policies related to hepatitis B prevention. Shi (2014) has shown that OR can be used when a city reforms its school choice system by making the objects of discussion precise and quantifying trade-offs. The "Doing Good with Good OR" initiative has other relevant examples (Ergun, Keskinocak, & Swann, 2011). With this project, we hope to contribute towards lowering the obstacles that the managers of tomorrow have in understanding OR.

1.4.3 Scientific relevance

This research is scientifically relevant for two reasons:

- It will provide a preliminary assessment on whether gamification is an appropriate means to deliver digital training about OR;
- It will create the groundwork for future research of applying gamification for increasing the knowledge of an innovation in the context of the innovation diffusion theory.

1.5 Thesis outline

This chapter introduced the main problem, attempted to give a general overview of the OR domain and suggested gamification as a solution to the problem. Furthermore, it also positioned the research problem, defined the objectives and the implications of the research.

Chapter 2 presents the research methods that are utilized to answer the research questions and meeting the objective of this research.

Chapter **Error! Reference source not found.** presents and elaborates on how the research undertaken conformed to the research methods utilized to produce the desired results.

Chapter 4 describes in detail the design process that lead to the final version of the main artifact of this research project.

Chapter 5 explains in detail the steps that we took to evaluate the artifact: it presents how the experiment was delivered to the control and treatment groups, followed by a presentation of the collected data and a section in which the data is analyzed.

Chapter 6 concludes the document by presenting a summary of the research results, and provides the answers to the original research questions. In the same chapter we include various suggestions and proposals for future research, as well as the limitations of the current research.

2 Research approach

2.1 Research process

In terms on Information Science (IS), the work of (von Alan, March, Park, & Ram, 2004) delineates two possible paradigms for conducting research: behavioral science and design science. According to them, the former is concerned with studying IT artifacts that have already been implemented (in the organization) with the purpose of generating theories about behaviors. The latter is concerned with the creation and evaluation of IT artifacts with the purpose of solving organizational problems. According to the objective of this research project, the most suitable paradigm to follow is that of design science research.

The overall research process that was followed during this research project has been modeled by adapting the Information Systems Framework elements as proposed by (von Alan et al., 2004) and then overlaying the specific process steps of this research. The resulting research diagram is presented in Figure 6 and represents the core planning on how this research project consulted the existing knowledge base, built and evaluated the artifact that added to the knowledge base the necessary elements to tackle the existing environmental issues that have been the triggers for this research project. We found the concept of the IS Framework elements to be very helpful in charting the overall research process.

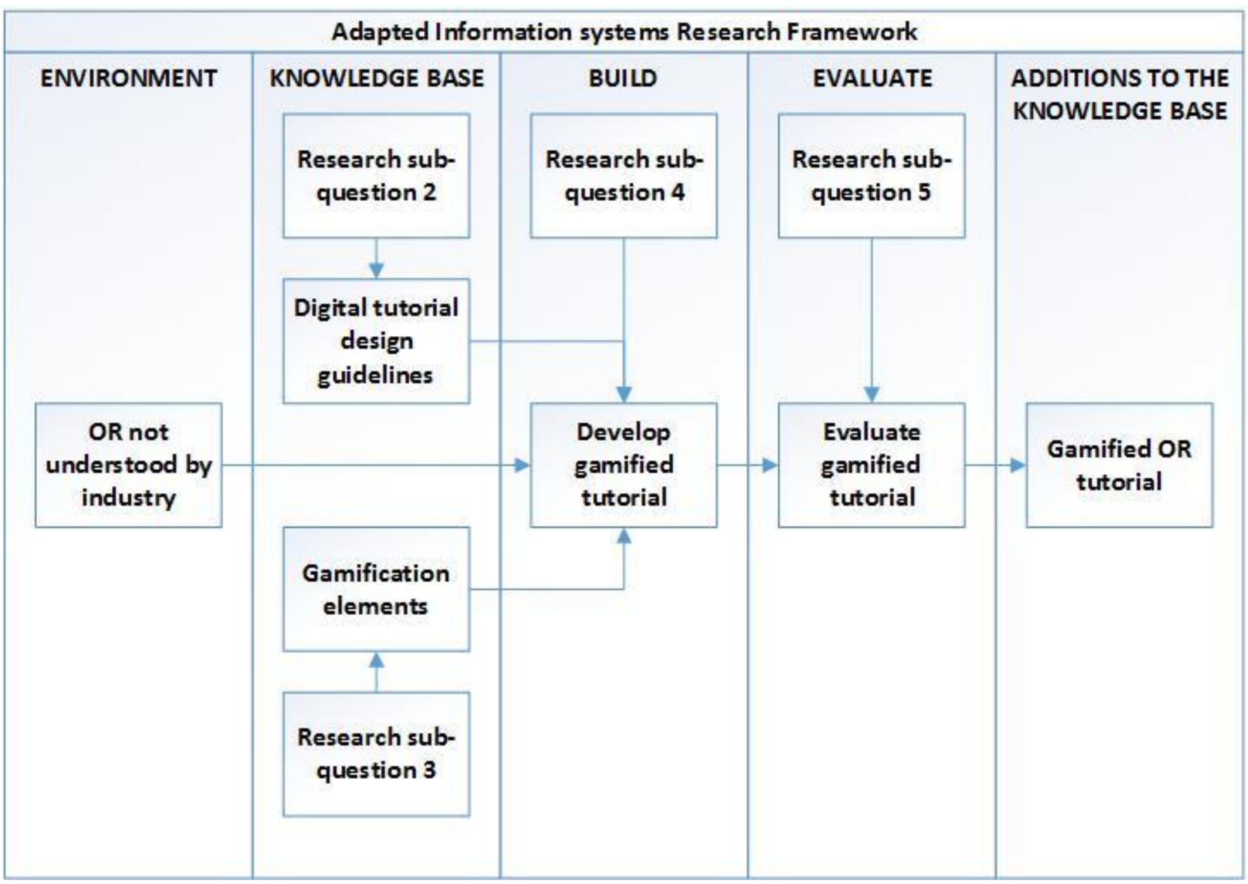


Figure 6 Information systems research framework (adapted from (von Alan et al., 2004)*)*

2.2 Research method

In this section we present the research approach that we took in order to answer the research questions and create the intermediate and final artifacts. We start by presenting, in Section 2.2.1, how we intended to obtain the content for the OR tutorial. Section 2.2.2 continues by presenting the process of obtaining the digital tutorial and gamification guidelines. We conclude, in Section 2.2.3, by detailing how the evaluation of the final artifact will be executed.

2.2.1 Defining the learning objectives

Since the artifact of this research project is a tutorial, we needed to clearly specify the objectives that we intended to accomplish with the tutorial (Mager, 1997). Based on these, we designed the learning experience and, later, evaluated if performance was according to the objectives. Therefore, it was very important that we started the creation of the tutorial by defining the learning objectives.

The work of Mager (1997) was used to define the learning objectives for operations research. In his work, Mager defines the concept of learning objective as a collection of words, symbols and/or pictures describing the intent of learning. He also provides guidelines for defining useful learning objectives. A correctly defined learning objective should clearly state (in the following description we use the term performance to refer to any action/behavior that the individual (learner) has acquired during the learning process):

1. What action/behavior should the learner be able to perform? Usually a verb, could be accompanied by a noun describing the product of the performance (e.g., write a 500 words essay on a specific topic).
2. Under what conditions (if any) should the learner be able to perform? These conditions usually have an impact on the performance (e.g. the essay should be written in ten minutes).
3. What is the acceptable performance level? Usually a check list of criteria that, successfully checked, deem the learner as competent (learning objective has been achieved) (e.g. the essay should treat the specified topic).

Biggs & Tang (2011) propose a learning outcome taxonomy, Structure of Observed Learning Outcomes (SOLO), which helped us to define the learning outcomes. According to SOLO, depending on the type and the complexity of knowledge, the learning outcomes should use different classes of verbs as illustrated in Table 1.

SOLO differentiates between two types of knowledge, (Biggs & Tang, 2011):

- Declarative: knowing about phenomena, theories, disciplines or specific topics.
- Functioning: requiring the student to exercise active control over problems and decisions in the appropriate content domains.

SOLO differentiates between five stages of complexity of knowledge that build on each other:

1. Pre-structural: students are simply acquiring bits of unconnected information, which have no organization and make no sense.
2. Unistructural: simple and obvious connections are made, but their significance is not grasped.
3. Multistructural: a number of connections may be made, but their significance in relation to the whole is missed.

4. Relational: the student is now able to appreciate the significance of the parts in relation to the whole.
5. Extended abstract: the student is making connections not only within the given subject area, but also beyond it, able to generalize and transfer the principles and ideas underlying the specific instance.

	Declarative knowledge	Functioning knowledge
Unistructural	Memorize, identify, recite	Count, match, order
Multistructural	Describe, classify	Compute, illustrate
Relational	Compare and contrast, explain, argue, analyze	Apply, construct, translate, solve a problem, predict within same domain
Extended abstract	Theorize, hypothesize, generalize	Reflect and improve, invent, create, solve unseen problems, extrapolate to unknown domains

Table 1 Typical declarative and functioning knowledge verbs by SOLO level (Biggs & Tang, 2011)

For the purpose of defining the learning objectives, an expert panel was assembled, which consisted of 4 OR experts: one from academia and three from industry.

The process through which the learning objectives have been defined and refined involved the expert panel at different stages:

- Finding resources treating introduction to operations research (by consulting the expert panel)
- From these resources, a list of basic operations research concepts was extracted
- The list was validated (by the expert panel)
- The list, the resources, the work of Mager and the SOLO framework were used to define the learning objectives
- The learning objectives were be validated (by the expert panel).

2.2.2 Snowballing

In the context of this research, for the literature review, we used the "snowball technique". This technique has been chosen because the thesis supervisors have provided valuable literature references that could be used as a starting point for the literature review. The technique has been compared with the structured literature review technique by (Jalali & Wohlin, 2012) without clear results in favor of one or the other. Webster & Watson (2002) describe the snowball method in three steps:

1. Finding one or more publications relevant to the research objectives.
2. Going backward by examining the references of the publications identified in step 1.
3. Going forward by examining the papers that cite the publications identified in step 1.

This process is iterated until there are no new concepts for the current iteration or until all publications identified in step two and three are exhausted (Webster & Watson, 2002). It is important to note that this technique does not stop at the first level of backward/forward references, but it applies recursively to any relevant references.

Relevancy to the research objectives was decided based on evaluating the title, the abstract or the text of the publication. The process was be recorded and the output is be available for review. Concerning the publications included in the recorded output: all publications that were included have been checked for relevancy, while obviously irrelevant (title review) publications were not included.

Considering the goals of the literature review, we decided to:

- Include just scientifically reviewed publications. An exception from this rule would be made in case the scholarly (peer reviewed) publications would not be enough to exhaust the research objective. In that case, we would also include grey literature (i.e. literature that has been less thoroughly reviewed): books, book chapters, theses and dissertations.
- Exclude references that cannot not be found in digital format in the Utrecht University library.
- Exclude references that present lessons learned. Recommendations that have not been validated experimentally (control-treatment), cannot be considered guidelines.

The literature review that answered RSQ2 had the following starting point: *"Gamicad: a gamified tutorial system for first time AutoCAD users"* (Li et al., 2012).

The literature review that answered RSQ3 had the following starting point: *Improving the adoption of software engineering practices through persuasive interventions* (Singer, 2013).

Once all the relevant references have been gathered, we proceeded towards a concept-centric analysis, recommended by (Webster & Watson, 2002): we extracted concepts from all the found references (based on their relevance to the literature review objectives) and we summarized them in a concept table (i.e. concept X -> author A, author B). The aim of this research project was not to review the state of literature regarding digital interactive tutorial guidelines or gamification guidelines. For that reason, we decided to keep the review as lightweight as possible (i.e. we extracted from the resources the concepts that were directly relevant to the review objectives).

2.2.3 Evaluation

In order to validate whether gamification influences the learning experience, we conducted a scientific experiment. This consisted of a pre experiment questionnaire, the controlled experiment and a post experiment questionnaire.

2.2.3.1 Pre experiment questionnaire

The pre experiment questionnaire allowed us to measure the values of the dependent variables.

According to (Prensky, 2005), one of the characteristics of play is that it is free, you should not be forced to do it. For the purpose of the experiment, we will not allow the participants to choose between using the gamified tutorial or the non-gamified tutorial. We took this approach to ensure that there was no participant bias towards choosing a tutorial that they prefer.

Prensky (2005) suggests the following constructs to be measured in order to have an understanding of the participants: age, gender, competitiveness, previous experience with games. In addition to this, we needed to measure the participants' experience with OR.

To measure competitiveness, we used the instrument developed by (Griffin-Pierson, 1990). They have shown that competitiveness can be broken down into two distinct constructs: Goal Competitiveness and Interpersonal Competitiveness. For Interpersonal Competitiveness, they use the definition given by (Spence & Helmreich, 1978): "the desire to do better than others, the desire to win in interpersonal situations, the enjoyment of interpersonal competition". For Goal Competitiveness, the definition

(Griffin-Pierson, 1990) give is: “the desire to excel, the desire to obtain a goal, the desire to be the best one can be”. They also created the devices to measure these constructs. It can be inferred from the paper of (Prensky, 2005) that the competitiveness he refers to is Interpersonal Competitiveness. Thus, the questions proposed by (Griffin-Pierson, 1990) that measure this concept will be used in the pre experiment questionnaire.

The construct previous experience with games was measured by adapting the questionnaire developed by Newcombe and Terlecki[4]. Upon inquiry, the authors stated that the questionnaire may be scored as it fits the purpose. In their usage scenario, points were allocated per question and the higher the score, the more experience the participant had. The participants where then divided into quartiles and only the highest and lowest quartiles were compared to each other. From their questionnaire, we used only the questions related to the previous experience with games concept: “How long have you been playing video games?” and “How often (approximately) do you currently play video games?” For the first question, we assigned points to the answers ascending (a-1; b-2; c-3; d-4; e-5, f-0). For the second question, we assigned points to the answers descending (a-5; b-4; c-3; d-2; e-1; f-0). These points were added together and, the higher the score, the more experience with games the participant had.

In order to measure the experience with OR of the participants, we asked the participants: how much experience (expressed in months) they have in general with operations research, if they could identify the correct definition of OR and if they could identify a linear programming problem. We chose not to measure their current knowledge level in depth, due to our chosen target population, which consists of students unfamiliar with the OR domain

2.2.3.2 Controlled Experiment

Concerning data collection, there were two possible designs that could have been used to obtain it (Gravetter & Wallnau, 2013):

- Independent-measures research design or a between-subjects research design, is one that uses a separate group of participants for each treatment condition
- Repeated-measures design, or a within-subject design, is one in which the dependent variable is measured two or more times for each individual in a single sample. The same group of subjects is used in all of the treatment conditions.

A comparison presenting the advantages and disadvantages of the between-subjects design and within-subject design is illustrated in Table 2.

	Between-subjects design	**Within-subject design**
Number of subjects	Requires double the number used in a within-subject design	
Study changes over time		Very well suited for longitudinal studies
Individual differences		Reduces or completely eliminates them, thus reducing variance and increasing the chance of finding a significant result.

[4] http://www.silccenter.org/resource-info/video_game_experience_survey2.pdf

Time-Related Factors	Not applicable	Could cause changes in the participants' scores.
Order effects	Not applicable	Could affect the participants' scores.

Table 2 Comparison of between-subjects design and within-subject design, adapted from (Gravetter & Wallnau, 2013)

In the post test of the experiment we compare the knowledge of the participants that were shown the simple tutorial with the participants that were shown the gamified tutorial. The experience that the participants gain during the first treatment will strongly affect their performance during the second treatment (strong order effects). For this reason, using a within-subject design is not feasible.

In this circumstance, we made note of the limitations of the between-subjects research design:

- It will require a large number of participants to the experiment. Considering the chosen population, we aimed at a number of at least ten participants per group. Similar studies by (Li et al., 2012), (Dong et al., 2012), (Grossman & Fitzmaurice, 2010) had similar number of participants.

The evaluation could have targeted one of two possible populations: students or people working in industry. Because of the requirements of the evaluation experiment (control group and treatment group), and limitation of the thesis, the sample has to have a low cost and a high availability. We concluded that students were the best choice and the limitations of this choice appear in the appropriate section. The representative group for this study were the students that follow a management education track. These students have a high possibility of becoming the managers that will adopt or reject OR.

Since this experiment aimed to measure the effect that gamification has in the education/learning context, we wanted to minimize the effect on the results of all the other variables. For that purpose, we decided to run the experiment in a controlled environment (i.e. all the participants would be simultaneously in the same room). The advantage of this approach is that all participants will spend the same time with the tutorial, in the same physical conditions (e.g. same level of distraction).

When designing the experiment, we were confronted with the decision of what to kind of learning experience to offer to the control group. It was possible to offer a classic (paper based) tutorial or a digital (computer based) tutorial. We chose the latter on the reason that the differences between a classic tutorial and a digital tutorial are fundamental (Kulik & Kulik, 1991). Since gamification is associated with the digital environment, the gamified tutorial would be a digital tutorial. If we would have administered a classic tutorial to the control group and a gamified digital tutorial to the treatment group, the measurement of the gamification effect would have been unreliable.

We explain below what conditions were administered to the participants in the experiment:

- The **control group** was presented and asked to go through a **(non-gamified) digital interactive tutorial** for teaching OR
- The **treatment group** was presented and asked to go through a **gamified digital interactive tutorial** for teaching OR

We decided to deliver the experiment with both tutorials in one go. We did not want to deliver the simple tutorial before we knew what changes gamification would bring to it (i.e. gamification might have required redesigning parts of the non-gamified tutorial, so that they were as similar as possible). If the simple tutorial would have already been delivered, these changes could have influenced, in unpredictable ways, the results of the experiment.

In order to make an interesting qualitative analysis we also wished to collect quantitative data existing during the experiment. Li et al. (2012) measured completion time and completion rate of the testing tasks. The system we developed contained functionality that measured quantitative data:

- time spent on every page
- times the show answer feature was used
- times an answer was submitted
- times a correct answer was submitted

2.2.3.3 Post experiment questionnaire

The design of the post experiment questionnaire was based on Kirkpatrick's (1994) four levels for evaluating training framework. Even if it is commonly used for evaluating training, this framework was used successfully for evaluating achievement of learning outcomes in games (O'Neil, Wainess, & Baker, 2005).

Kirkpatrick has developed a framework for the evaluation of training results that has become dominant since its creation (Kirkpatrick, 1994). The framework, presented in (Kirkpatrick, 1994), is structured in 4 levels:

- *Reaction*, "how those who participate in the program react to it". A positive reaction does not guarantee learning, but a negative reaction will reduce it.
- *Learning*, evaluated in the context of the learning program (related to the learning objectives).
- *Behavior*, measured after the training. This refers to the way the learned material is applied in out-of-training situations.
- *Results*, they are the reason for which the learner attended the training.

A literature review was conducted by (O'Neil et al., 2005) on the topic of learning outcomes in the context of video games. They reviewed the empirical research on video games and classified it according to the frameworks of (Kirkpatrick, 1994) and of (Baker & Mayer, 1999). They did not encounter any difficulty in adapting these frameworks, which were not designed to measure learning outcomes of games, to the video games context. By doing this, they reached the conclusion that it is also a good framework for evaluating learning from video games. Additionally, they also note that the frameworks are complementary, Kirkpatrick's taking a macro view of the evaluation and Baker & Mayer's taking a micro view (focusing only on the second level, learning, of Kirkpatrick's framework).

Kirkpatrick's framework was also successfully applied in their research of gamified interactive tutorials by (Dong et al., 2012). Therefore, we considered it appropriate for measuring the learning outcomes of the tutorial we designed.

Four levels for evaluating training	
Level 1: REACTION	Trainee's reaction to the program: level of satisfaction
Level 2: LEARNING	Trainee's attitude change, increased knowledge, and/or increased skill, due to the training
Level 3: BEHAVIOUR	On the job change in behaviour because of program participation, i.e. transfer of learning to the job setting
Level 4: RESULTS	How the organization benefited from the learner's participation in the program (e.g. increased production or profits, improved quality, decreased costs, fewer accidents)

Figure 7 Kirkpatrick's framework

The post questionnaire measured the extent to which the learning outcomes were achieved by the participants. For this purpose, questions were developed in close relation to the learning outcomes and their type and complexity of knowledge. The post questionnaire was validated by the expert panel.

The post questionnaire covered three aspects:

- Satisfaction with the tutorial system (measuring Kirkpatrick's Reaction)
- Operations research business value (measuring Kirkpatrick's Learning: ILO1). Each correct (according to the acceptance criteria) case was graded with 1 point.
- Operations research concepts (measuring Kirkpatrick's Learning: ILO2, ILO3, ILO4). Each correct response in this section was marked with 1 point. The points were summed up, to give a section score.

3 State of the scientific literature

This chapter will describe in detail the steps taken to answer RSQ1, RSQ2 and RSQ3.

3.1 Learning objectives of operations research

In this section we will be answering RSQ1: “What are the intended learning outcomes of operations research *for potential new adopters*?” based on the methodology described in Section 2.2.1.

The learning outcomes have been defined based on the curriculum that was found in five books treating the topic of optimization modelling (Bisschop, 2006), (Taha, 2007), (Vanderbei, 2001), (Jensen, Paul A and Bard, 2003), (Bradley, Hax, & Magnanti, 1977). After the learning objectives have been defined by the main researcher, the expert panel was asked to validate whether the learning objectives were suitable for an OR agnostic audience.

3.1.1 Defining the intended learning outcomes

The following sections provide a clear statement of the learning objectives for the operations research learning tutorial according to the guidelines of (Mager, 1997) and to the SOLO taxonomy. We have identified the content topic as introduction to OR.

3.1.1.1 Actions

The guidelines suggested by (Mager, 1997), were used to define the following actions that the learner should be able to perform:

1. Describe real-world cases in which operations research can be used to improve problems/situations that the learners are aware of (declarative, multistructural).
2. Describe the main concepts of operations research (Table 3) (declarative, multistructural)
3. Recognize each of these concepts when they appear in the text of a linear programming (LP) problem with 2 to 4 variables. (functioning, unistructural)
4. Recognize a LP problem with 2 to 4 variables. (functioning, unistructural)

3.1.1.2 Conditions

Each of the actions should be performed by the learner without any sources of information.

3.1.1.3 Acceptable performance

The acceptable performance levels for each of the actions identified in Section 3.1.1.1 are defined in conformance with the respective stage of complexity of knowledge:

1. The cases presented by the learner should be verifiable (a situation that is known in the operations research domain) and not an exact copy of the cases presented in the tutorial.
2. The learner should be able to correctly complete a “fill in the blanks” paragraph that contains all the concepts.
3. Given the text of a simple LP problem, the learner should be able to identify all of the concepts.
4. The learner should be able to select all the LP problems from a set of simple problems (LP and non-LP).

The acceptable performance levels will serve as the basis for the creation of the post experiment questionnaire.

3.1.2 Operations research concepts

In order to identify the learning objectives for the application we have looked for OR courses or books, those that were introductions to optimization modelling (the expert panel was also inquired about references to introductory resources on OR).

We have identified the following:

- (Bisschop, 2006), a general introduction to optimization modeling discussion basic and advanced concepts. It is aimed at users who are new to modeling or who have limited modeling experience
- (Vanderbei, 2001), a first introduction to constrained optimization. Treats topics like: linear programming, convex analysis, network flows, integer programming, quadratic programming, and convex optimization.
- (Bradley et al., 1977), introduces readers to the theory, algorithms, and applications of optimization. Treats topics like: linear programming, network optimization, integer programming, and decision trees.
- (Jensen, Paul A and Bard, 2003), is designed to bridge the gap between theory and practice by presenting the tools and techniques most suited for modern operations research. A principal goal is to give engineers, analysts, and decision makers a larger appreciation of the role of OR by providing examples of its applications and the basics of its theoretical development. It covers deterministic optimization and probabilistic systems.
- (Taha, 2007), provides a balanced coverage of the theory, applications, and computations of operations research. Aimed at beginning through advanced students.

Following that, we looked at the first part of the resource and extracted from there, for each, the concepts and their corresponding descriptions that appeared to be relevant: they would appear in text as emphasized, they would appear many times or they would be used as foundation for concepts introduced further along in the text. The results of the extraction are presented in Table 3.

The definition column contains the most detailed and clear definition out of all the resources available. Where it was the case, the definition was improved by merging definitions from multiple (usually two) sources. The notations used in Table 3, in the last five columns (the resources columns) are explained:

1. In case a field in the comparison table is left blank it means that a definition for the concept was not found in that particular resource.
2. An '=' symbol indicates that a concept's description was found to be similar in that particular resource.
3. A ">" symbol indicates that a description was found in the resource for that concept, but the reference description is more complete.
4. If a cell contains the name of a concept (e.g. C2), it indicates that the same description was found for that concept (C2) in the resource as the description of the concept in the first column (C1). In other words, C1=C2 according to the description, but not according to their naming.

The information contained in the table was validated with an expert panel.

The expert panel suggested the following improvements to the original version of the table that was shown to them:

- Adding an example column, and the mathematical notation where possible
- Using the term description in the second column instead of definition
- Upon inquiry by the main researcher, agreed that some terms are not relevant enough or are too advanced to be included in this version of the concepts table: “operations”, “research”, “Simple upper bound”, “Shadow Price”, “Parametric analysis”, “Unbounded problem”, “Simple bound”
- Adding a graphical representation of the relations between the concepts
- Unifying the different definitions into a more definition (introduction of the “ >” sign)
- Improving the definition of “Feasible region”

Term	Definition	Example	(Jensen, Paul A and Bard, 2003)	(Bisschop, 2006)	(Vanderbei, 2001)	(Bradley et al., 1977)	(Taha, 2007)
Operations research	A scientific approach to managerial decision making. It attempts to apply mathematical methods and the capabilities of modern computers to the difficult and unstructured problems confronting modern managers. Also known as: management science, operational research, systems analysis, cost–benefit analysis, and cost-effectiveness analysis.	http://pubsonline.informs.org/doi/abs/10.1287/inte.1080.0409 http://pubsonline.informs.org/doi/abs/10.1287/inte.1110.0601 https://www.informs.org/Recognize-Excellence/Franz-Edelman-Award/Franz-Edelman-Laureates2 http://www.scienceofbetter.co.uk/				=	
Mathematical programming	Branch of management science. Concerns the optimum allocation of limited resources among competing activities, under a set of constraints imposed by the nature of the problem being studied. In broad terms, mathematical programming can be defined as a mathematical representation aimed at programming or planning the best possible allocation of scarce resources.					=	
Model	An abstract representation of reality. In the current context, a representation of a decision problem related to the operations of the organization.		=	>		>	
Mathematical model	An abstract model that describes, in general mathematical terms, the relations contained in a model.			=			OR model
Optimization/ analytical	A mathematical model that contains a criterion or objective, which we seek to optimize (e.g.			>		=	

model	maximize or minimize), subject to a set of mathematical constraints that portray the conditions under which the decisions have to be made.						
Linear programming (models)	Linear programming consists of optimization models made up from linear equations and linear inequalities. The feasible decisions are compared using a linear objective function that depends on the decision variables.	http://www.me.utexas.edu/~jensen/ORMM/models/unit/linear/subunits/product_mix/index.html http://www.me.utexas.edu/~jensen/ORMM/models/unit/linear/subunits/resource_allocation/index.html http://www.me.utexas.edu/~jensen/ORMM/models/unit/linear/subunits/blending/index.html	=	=		>	>
Decision variables	The quantities that the decision makers would like to determine. They are the unknowns of a mathematical programming model.	Explained using the examples provided here: http://www.me.utexas.edu/~jensen/ORMM/models/unit/linear/subunits/workforce/index.html http://www.aimms.com/downloads/tutorials/tutorial-for-beginners/	=	>	>	=	>
Constraint	An inequality or equality defining limitations on decision variables.		=		=	=	=
Objective function	A function that consists of decision variables that needs to be maximized or minimized which specifies the criterion the decision maker will use to evaluate alternative solutions to the problem		>		>	>	
Parameters	The collection of coefficients, representing known data, which are used in the model.		>	=			
Non-negativity constraints	A special kind of constraint. In most practical problems most variables are required to be nonnegative.		>	=			
Solution	A proposal of specific values for the decision variables.			>	=		
Feasible	A solution that satisfies all of the constraints.				=		=

solution							
Optimal solution	A feasible solution that yields the best value (maximum or minimum) of the objective function.		>		>		=
Infeasible problem	A problem with no feasible solution.				=		
Graphical representation of a two variable LP							
Contour of the objective function	(inferred) Any line corresponding to a specific value of the objective function.	Shown on graphical example		=			
Feasible region	Region bounded by the lines corresponding to the constraints (where all the feasible solutions of the model lie). A constraint line separates any plane into a feasible plane and an unfeasible plane.	Shown on graphical example		=		=	=
Optimal corner solution	If a linear programming model has an optimal solution, an optimal solution is on a corner of the feasible region (intersection of two lines).	Shown on graphical example; also that it can be on a line		=		=	=

Table 3 Operations research concepts

Since operations research is a discipline with foundations in mathematics, the application of the concepts from Table 3 appears as an OR problem. We extracted the first problem that was shown/explained, the results are presented in Table 4.

Author	First operations research application
(Jensen, Paul A and Bard, 2003)	Product Mix Problem -> a LP problem
(Bisschop, 2006)	Linear programming is the simplest of the three main classes of constrained optimization models. Problem: potato chips problem. (Resource allocation problem)
(Vanderbei, 2001)	Resource allocation problem (maximize profit, constraint is the availability of raw materials.
(Bradley et al., 1977)	Production and Assembly (Resource allocation problem).
(Taha, 2007)	The first example consist of liner programming model (maximization): A company has 2 types of products and wants to maximize profit. They are constrained by the availability of the raw material.

Table 4 First application of operations research

By analyzing Table 4, we concluded that the first, easiest application of OR is a linear programming problem (with one to four variables).

In addition to identifying the easiest application of operations research, we needed to present the methodology of solving such an application (in this specific case, what steps are required to solve the problem successfully):

1. Define/identify the decision variables of the problem.
2. Determine the criterion the decision maker will use to evaluate alternative solutions to the problem. In mathematical-programming terminology, this is known as the objective function.
3. Define the constraints of the problem, which are the restrictions imposed upon the values of the decision variables by the characteristics of the problem under study.

3.2 Digital interactive tutorials

This section is concerned with answering RSQ2: "What are the guidelines for creating a digital interactive tutorial?" In order to do that, we will present the results of the literature review aimed at identifying guidelines for designing digital interactive tutorials. This section will end with the conclusions that can be drawn from performing the literature review.

3.2.1 Literature review process

Before starting the literature review it was necessary to identify the research objective. In the case of this research question we first needed to identify the publications that treat the domain of creation of digital interactive tutorials. Following that, we needed to identify the publications that offer guidelines for creating such a tutorial.

In order to identify the guidelines for creating a digital interactive tutorial, we have executed the method described in Section 2.2.2. The artifact in which the scientific papers reviewed were recorded is illustrated in Appendix 7.5.

While conducting the literature review we have encountered two new concepts that relate indirectly to the research objective. We decided (see Section 6.3, Further research) to consider these directions out of the scope of the current research:

- User interface design, user experience design and, in general, human computer interaction. “Human computer interaction is an area of applied cognitive science and engineering design. It is concerned both with understanding how people make use of devices and systems that incorporate computation, and with designing new devices and systems that enhance human performance and experience” (John M Carroll, 2009).
- Exploratory learning, discovery based learning, discovery learning, guided discovery or guided exploration. “In exploratory learning, instead of working through precisely sequenced training materials, the user investigates a system on his or her own initiative, often in pursuit of a real or artificial task” (Rieman, 1996).

3.2.2 Guidelines

After having studied how people learn from computer systems, John Millar Carroll (1990b) concludes that systematic instruction is not efficient and proposes a “minimalist” approach. In his research, he mentions that it is important to notice what minimalism should not (necessarily) be used for: advanced training and reference documentation. He presents the following principles for creating a digital interactive tutorial:

1. Use real tasks for the training exercises and let users select their own tasks. It enables people to use their prerequisite competence and engages a "powerful source of motivation."
2. Get the learner started on real tasks fast by eliminating almost all front-end orientation material. Extensive preambles can "obstruct meaningful activity."
3. Guide learners' reasoning, exploring and improvising with questions and other hints. The author sometimes recommends presenting incomplete training materials, so that learners have to explore. This is a core principle, because it directs the focus of learning activity and provides the most contrast with many conventional approaches. He also suggests presenting summaries in place of complete texts.
4. Design the materials so that they can be read in any order in so far as possible. This principle permits learners to "support their own goal-directed activities." Use "high degree of modularity" and "small, self-contained units."
5. Help learners to coordinate training materials and software by providing landmarks for normal or error situations. Illustrations which show what the screen should look like if everything is OK are the primary example given of this principle.
6. Focus early attention in the training materials on enabling the learner to recognize and recover from errors. Learners make many kinds of errors in learning computer systems. "Training materials must therefore explicitly support the recognition of and recovery from error both to make the materials robust with respect to user error and to train error recovery skills." Guided exploration of error possibilities is important to speed up initial learning and decrease the frustration resulting from making so many errors.
7. Engage the learner's prior knowledge in introducing novel concepts. Use familiar office tasks, language and metaphors. Highlight differences in operation of the system from what might be expected from the learner's background.

8. Consider using the learning situation, as opposed to practical on-the-job examples, for learning examples, exercises and explorations. Help the learner understand the "fine detail of the actual situations in order to create practical solutions."
9. Aim for optimizing learning designs by repeated testing and avoiding the temptation to systematize approaches into checklists. "There is no deductive theory of minimalist instruction; that is, given a set of minimalist principles, we cannot just crank out a training manual. Design never works this way."

The work by John Millar Carroll (1990a) brings in more principles, based on the observation that learners make mistakes and those should not just be ignored and avoided, but taken into consideration when developing the digital interactive tutorial:

1. "*Allow the user to get started fast*. Cut down overhead and repetition; cut down nonessential verbiage; reject the notion that every function must be covered, people never master every function even when every function is covered. Offer the learner meaningful activities as soon as possible.
2. *Rely on the user to think and to improvise*. Encourage and guide user inference; leave out material that can be inferred. Don't try to give the user an understanding when you can allow the user to create an understanding.
3. *Direct training at real tasks*. Introduce real work immediately. Instruction, no matter how well-organized, will fail if it fails to support the goals people bring to the learning situation.
4. *Exploit what people already know*. Even if it is possible to learn without analogy, it is too abstract and cumbersome.
5. *Support error recognition and recovery*. Errors cannot be avoided in learning, but they can confuse and frustrate Iearners. If they are properly managed they may play useful roles."

The research of Catrambone & Carroll (1987) has validated that the guideline of creating a training wheels (TW) interface enables learners to do basic tasks more quickly than those who use the full system (because they spend less time in error recovery mode). A TW interface is defined as routines which block the user from accessing various parts of a system, acquaint users with a system by letting them use it right away for meaningful tasks and at the same time prevent them from suffering the consequences of certain typical mistakes. An example of using a TW interface: the learner selects an operation that is inappropriate or unnecessary (at the current stage of the digital interactive tutorial), the TW interface returns a message indicating that the choice is unavailable.

The research of Grossman & Fitzmaurice (2010) investigated the use of animated or video documentation in addition to the textual contextual help (tooltips that appear on hovering over an element of the interface for a longer time) that the learner can access in digital interactive tutorial and have concluded that this technique increases task completion rate. There are also guidelines for designing the video content and presenting it, but those are out of the scope of the current research.

After they concluded their research in guided exploration tutorials, Vanderlinden (1988) found the following guidelines helpful in developing guided exploration tutorials:

1. "*Understand the user's task.* It is imperative to clearly define your learning objectives.

2. *Use naive users to identify possible interface problems*. Peer review is useful for first pass usability testing of a tutorial. When testing tutorials we need to find out not only how well the tutorials work, but also how users expect the system to work.
3. *Offer generic help table.* We found that we could generalize types of error conditions and provide recovery information in a central table.
4. *Use graphics as much as possible.* Use detailed graphics and text redundantly to show as well as tell how to do tasks.
5. *Use icons or other typographical elements (rules/boxes) as user signals.* We observed several subjects who followed every possible step even when it was inappropriate. These particular subjects were probably the die-hard proceduralists. To provide stronger clues, we used icons rather than text for the hint, checkpoint, and rescue signals in our final usability testing.
6. *Test final tutorial iteratively to get the bugs out.*
7. *Because users don't always read, use another medium to teach interface skills.* When we analyze our audiences we frequently assume a basic skill – the ability to read. Even those who can read well, frequently do not read carefully. We also need to remember that the more senses learners bring to a learning situation, the greater their learning retention."

3.2.3 Literature review conclusions

In Table 5 we present the guidelines that concern the design of digital interactive tutorials.

Reference	No	Guideline	Rationale
(John Millar Carroll, 1990b)	1	Use real tasks for the training exercises and let users select their own tasks	It enables people to use their prerequisite competence and engages a "powerful source of motivation."
(John Millar Carroll, 1990b), (John Millar Carroll, 1990a)	2	Get the learner started on real tasks fast by eliminating almost all front-end orientation material	Extensive preambles can "obstruct meaningful activity."
(John Millar Carroll, 1990b), (John Millar Carroll, 1990a)	3	Guide learners' reasoning, exploring and improvising with questions and other hints	Forces learners to explore. This is a core principle, because it directs the focus of learning activity and provides the most contrast with many conventional approaches. He also suggests presenting summaries in place of complete texts.
(John Millar Carroll, 1990b)	4	Design the materials so that they can be read in any order in so far as possible	Permits learners to "support their own goal-directed activities." Use "high degree of modularity" and "small, self-contained units."
(John Millar Carroll, 1990b)	5	Help learners to coordinate training materials and software by providing landmarks for normal or error situations	Illustrations which show what the screen should look like if everything is OK are the primary example given of this principle.
(John Millar Carroll, 1990b), (John Millar Carroll,	6	Focus early attention in the training materials on enabling the learner to recognize and recover from errors	Learners make many kinds of errors in learning computer systems. "Training materials must therefore explicitly support the recognition of and recovery from error

1990a)			both to make the materials robust with respect to user error and to train error recovery skills." Guided exploration of error possibilities is important to speed up initial learning and decrease the frustration resulting from making so many errors.
(John Millar Carroll, 1990b), (John Millar Carroll, 1990a)	7	Engage the learner's prior knowledge in introducing novel concepts	Use familiar office tasks, language and metaphors. Highlight differences in operation of the system from what might be expected from the learner's background.
(John Millar Carroll, 1990b)	8	Consider using the learning situation, as opposed to practical on-the-job examples, for learning examples, exercises and explorations	Help the learner understand the "fine detail of the actual situations in order to create practical solutions."
(John Millar Carroll, 1990b)	9	Aim for optimizing learning designs by repeated testing and avoiding the temptation to systematize approaches into checklists.	"There is no deductive theory of minimalist instruction; that is, given a set of minimalist principles, we cannot just crank out a training manual. Design never works this way.”
(Vanderlinden, 1988)	10	Understand the user’s task.	It is imperative to clearly define your learning objectives.
(Vanderlinden, 1988)	11	Use naive users to identify possible interface problems	Peer review is useful for first pass usability testing of a tutorial. When testing tutorials we need to find out not only how well the tutorials work, but also how users expect the system to work.
(Vanderlinden, 1988)	12	Offer generic help table.	We found that we could generalize types of error conditions and provide recovery information in a central table.
(Vanderlinden, 1988)	13	Use graphics as much as possible.	Use detailed graphics and text redundantly to show as well as tell how to do tasks.
(Vanderlinden, 1988)	14	Use icons or other typographical elements (rules/boxes) as user signals	There were particular subjects who followed every possible step even when it was inappropriate. To provide stronger clues, used icons rather than text for the hint, checkpoint, and rescue signals.
(Vanderlinden, 1988)	15	Test final tutorial iteratively to get the bugs out.	
(Vanderlinden, 1988)	16	Because users don’t always read, use another medium to teach interface skills	Users, frequently do not read instructions carefully. We also need to remember that the more senses learners bring to a learning situation, the greater their learning retention.
(Catrambone & Carroll, 1987)	17	Training-wheels interface	Enables learners to do basic tasks more quickly than those who use the full system (because they spend less time in error

			recovery mode)
(Grossman & Fitzmaurice, 2010)	18	Animated tooltips	By adding animated or video documentation in addition to the textual contextual, the task completion rate is increased.

Table 5 Selected guidelines for designing digital interactive tutorials

3.3 Gamification

This section is concerned with answering RSQ3: "What aspects of gamification can be applied on top of a digital interactive tutorial to increase its effectiveness?" In order to do that, we will present the results of the literature review aimed at identifying guidelines for applying gamification in the learning and education contexts. This section will end with the conclusions that can be drawn from performing the literature review.

3.3.1 Literature review process

In order to identify the guidelines for applying gamification in the learning and education contexts, we have executed the research method described in Section 2.2.2. The artifact in which the scientific papers reviewed were recorded is illustrated in appendix 7.6.

While conducting the literature review we have encountered new concepts that relate indirectly to the research objective. We decided (we discuss the implications of this choice in Section 6.3, Further research) to consider these directions out of the scope of the current research:

- User interface design, user experience design and, in general, human computer interaction
- Video game design
- Incentives, persuasion, motivation mechanisms, self-determination theory etc.
- The concept of flow, "a state of concentration or complete absorption with the activity at hand and the situation. It is a state in which people are so involved in an activity that nothing else seems to matter" (Csikszentmihalyi, 1991)
- Achievement systems. "From the perspective of the achievement system, an achievement appears as a challenge consisting of a signifying element, rewards and completion logics whose fulfilment conditions are defined through events in other systems (usually games). From the perspective of a single game, an achievement appears as an optional challenge provided by a meta-game that is independent of a single game session and yields possible reward(s)." (Hamari & Eranti, 2011)

We have also excluded studies that measured only the user experience for the gamified application, but not the achievement of the intended learning objectives. Studies were excluded from the literature review if they did not contain a comparison of the gamified application with a non-gamified version.

3.3.2 Guidelines

In order to identify the guidelines for creating a digital interactive tutorial, we have executed the method described in Section 2.2.2 with the amendments presented in Section 3.3.1 . The artifact in which the scientific papers reviewed were recorded is illustrated in Appendix 7.6.

3.3.2.1 Reviews of gamified systems

Denny (2013) evaluated a gamified version of PeerWise (a platform that allows students to create their own exam-style questions relevant to the course they are studying and share these via a central repository with other members of the class). Gamification consisted of a badge-based achievement system. They have uncovered that the badges have impact when the activity they are rewarding already has visible value; utilizing badges in similar contexts is a low-risk proposition for educators as they do not appear to negatively affect player participation.

Farzan & DiMicco (2008) evaluated a gamified social networking site. Gamification consisted of a point-based system. They have uncovered that the points system over the long term did not encourage more users to contribute to the site, but it boosted users' contributions when introduced. Additionally, players wanted the possibility to opt-out; another request was that the players wanted to be able to customize the system (assigning point values to the things they felt were most valuable).

Fitz-Walter & Wyeth (2013) evaluated a gamified Logbook application for driving learners. Gamification consisted of: narrative, virtual currency, content unlocking, player could chose artifacts to use, feedback, visual effects. They have uncovered that: gamification shouldn't come at the cost of utility and usability; an interactive tutorial could be included to teach the player the basics of gameplay; gamified systems should provide a way to opt-out of interacting with the game elements; potential for cheating in gamified systems needs to be identified early and addressed.

Hakulinen, Auvinen, & Korhonen (2013) evaluated a gamified online Learning Environment. Gamification consisted of achievement badges. They have uncovered that achievement badges have a significant impact on some aspects of participants' behavior (self-reflection, increased awareness of own studying habits), with a small group especially motivated to pursue them. They conclude stating that badges seem a good way of motivating students to study and to use desired learning practices.

Schutter & Abeele (2014) evaluated a gamified game design for educational purposes course. Gamification consisted of: quests, experience points, heroes (the player character), guilds (player groups), character levels, character skills and leaderboards. They have uncovered that certain game elements (being able to choose your side quest, special skills), could be related to higher engagement, motivation or enjoyment than others. Furthermore, they argue that the designer of the gamified application should: know who is the target population is; provide students with freedom of choice in how they want to show their mastery of the materials; provide extra structure (that allows the participants to learn the rules of gamification); evaluate their knowledge of the game (test, in the application that they understand the rules); create real-world quest names (so the quests are clear from their titles); communicate that the course will be challenging (people might expect a gamified application to be as easy as a game).

Domínguez et al. (2013) evaluated a gamified e-learning platform. Gamification consisted of: an achievement system, competition elements (possibility to compare to other and possibility to see what are all the achievements, and leaderboards). They have uncovered that gamification can have a great emotional and social impact on some students; cognitive impact of gamification over students is not very significant (similar results in tests). They recommend including an interactive introduction which not only explains, but also guides the participant through the gamification platform; tasks should be automatically evaluated by the e-learning platform.

Li et al. (2012) evaluated a gamified tutorial system for first time AutoCAD users (GamiCAD). Gamification consisted of: narrative, bonus levels, progressive disclosure, Clear Goals for each level, Speed Bonus Points, score counter, bonus points, 0 to 5 stars, Music and Sound Effects). They have uncovered that the participants using the gamified system finished tasks faster with a higher completion ratio and that the gamified condition was more enjoyable, fun and engaging. They posit that the performance gains were in large part due to unlocking mechanism present in GamiCAD, requiring users to repeat levels to improve their performance and that without enforcing a threshold level of performance, users may struggle through a tutorial without adequately learning the core concepts, as evident by the lower completion rates and higher completion time in the control condition.

Barata, Gama, Jorge, & Gonçalves (2013) evaluated a gamified college course. Gamification consisted of: experience points, levels, badges, challenges and leaderboards. They have uncovered that: students were participating and being more proactive in forums; the gamified course was perceived as more motivating and interesting; Students seem to score better and grade differences between them seem to decrease.

Charles, Bustard, & Black (2011) developed a Game-Enhanced Learning (GEL) Framework that defines a set of engagement factors: fun, social, identity, challenge, structure, feedback (Table 6). Even though they use the concept games in the framework, they are referring to gamification, since the framework is not designed to lead to the creation of a fully-fledged game. They also explain what stages are required to design a gamified application: Understand the Context (identifying requirements and operational constraints; extraction of a set of core and optional tasks from the existing way that the material is taught and the associated learning outcomes); Define the Challenge (design a set of hierarchical challenges; achievable and small, but with a level of difficulty that ensures a sense of achievement on completion; afterwards, develop the reward system; points can only be earned and never deducted). They evaluated this framework by gamifying two university courses. Gamification consisted of the five elements present in the framework. They have uncovered that: students enjoyed the courses more, attendance increased and more students have successfully finished the course; the gamified course requires the sponsorship (motivation and enthusiasm) of the teacher to increase awareness about the game.

Concept	Description
Fun	The game is essentially an additional voluntary activity (not graded). It should be fun but it should also have an educational benefit. In addition, it should have no negative impact on the course itself, or on any other course being studied at the same time. Engagement is easier if learning is enjoyable
Social	Students are encouraged to collaborate with fellow students to complete challenges. This cooperation helps provide a supportive environment in which students with mixed ability can all enjoy the game. It also implies that it is beneficial to make the game a group activity where that is practical.
Identity	the visibility of a student in the learning environment; greater visibility of students in a learning environment and a clear explanation of their role can encourage their engagement
Challenge	To provide a challenge, the game should include a selection of optional activities, with a suitable breadth of difficulty. Engagement can build on human competitive drive, enhanced by social pressure.
Structure	The basic game model assumes that students can earn game points by completing activities, some of which are considered ‘core’ and others that are ‘optional’. For example, points could be awarded for attendance at lectures, which is expected, and for participating in class, by answering questions, which is desirable but optional.

	Engagement is more likely if objectives and constraints are clear and acceptable.
Feedback	To maintain enthusiasm, it is assumed that the time gap between students earning game points and seeing the impact on their performance (as individuals or in groups) should be as short as possible. Can take a number of forms, including online graphical summaries, publicly displayed leaderboards and the awarding of achievement prizes.

Table 6 Description of GEL engagement factors (Charles et al., 2011)

Foster & Sheridan (2012) evaluated a gamified activity in a reverse engineering course. Gamification consisted of: achievements. They have uncovered that the gamified version of the activity: increased the achievement learning objectives of the low performing teams; impetus to learners to encourage them to want to understand more deeply.

Meyer, Crosby, & Ogawa (2014) evaluated a gamified course of introductory Computer Science. Gamification consisted of: badges, a progress bar for their collection, avatars that evolve based on the number of badges. They have uncovered that: practical assignments grades improved and there were more non-graded assignments submitted.

O'Donovan, Gain, & Marais (2013) evaluated a gamified an online learning management tool used in a university course in Computer Games Development. Gamification consisted of: storyline, leaderboards, Points, ranks, progress bars, badges, rewards, virtual currency. They have uncovered that the gamified application: improved the students' understanding, their engagement, had a significant impact on course marks and lecture attendance. They emphasize the importance of a greater integration of the storyline.

Singer (2013) has created a catalog of patterns that facilitate the adoption of software engineering practices (Figure 8). They evaluated it through a gamified application used the following patterns: Normative Behavior, Triggers, Points & Levels, Leaderboard, Challenge and Progress Feedback. They uncovered that the gamified application improves the adoption of software engineering practices. It should be noted that the patterns cannot be used directly, Singer gives specific examples of game elements for each pattern.

Pattern	Section	*suitable for* Starting Adoption	Improving Adoption	Creative Tasks	Routine Tasks	*requires* Existing Adopters	Meaningful Metric
Knowledge Stage							
Mass Medium	7.2.1	✓	✓	✓	✓		
Microblog	7.2.2	✓	✓	✓	✓	✓	
Voice for Help	7.2.3	✓	✓	✓	✓	✓	
Reputation for Help	7.2.4	✓	✓	✓	✓	✓	✓
Persuasion Stage							
Normative Behavior	7.3.1	✓	✓	✓	✓	✓	
Social Anchor	7.3.2	✓	✓	✓	✓	✓	✓
Peer Recommender	7.3.3	✓	✓	✓	✓	✓	
Motivation: KAP-gap							
Triggers	7.4.1	✓	✓	✓	✓		
Potential Value	7.4.2	✓	✓	✓	✓		✓
Appreciation	7.4.3		✓	✓	✓		
Reputation for Adoption	7.4.4	✓	✓	✓[3]	✓		✓
Points & Levels	7.4.5	✓	✓	✓[3]	✓		✓
Leaderboard	7.4.6	✓	✓	✓[3]	✓		✓
Relative Ranking	7.4.7		✓		✓	✓	✓
Decision Stage							
Challenge	7.5.1	✓	✓	✓	✓		
Embrace Examples	7.5.2	✓	✓	✓	✓		
Incremental Engagement	7.5.3	✓	✓	✓	✓		
Implementation Stage							
Progress Feedback	7.6.1		✓	✓	✓		✓
Performance Feedback	7.6.2		✓	✓	✓		✓
Supporting							
Tuned Activity	7.7.1	✓	✓	✓	✓	✓	
Review	7.7.2	✓	✓	✓	✓		
Meta-Review	7.7.3	✓	✓	✓	✓		
Automatic Badges	7.7.4	✓	✓	✓[3]	✓		✓
Peer Badges	7.7.5	✓	✓	✓	✓		

Figure 8 Overview of adoption patterns *(Singer, 2013)*

3.3.2.2 Reviews of gamification literature

Nah & Telaprolu (2013) have reviewed the literature on gamification for gamifying computer education and developed a framework that provides guidance in gamifying educational applications (Figure 9). The framework starts from five principles (Goal orientation; Achievement, Reinforcement, Competition, and Fun orientation) continues to explain how these are achieved by system design elements and how to evaluate whether the gamification goals have been achieved (engagement/cognitive absorption measures).

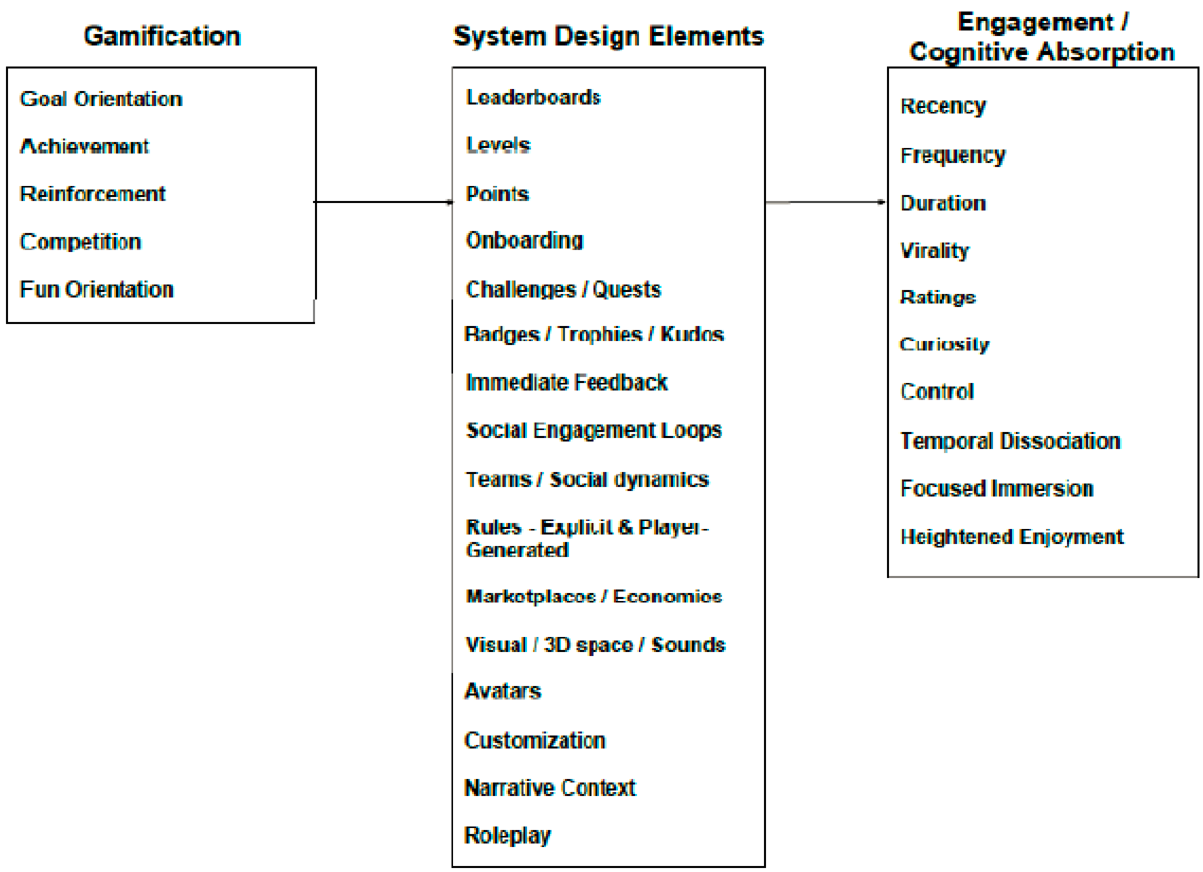


Figure 9 Framework for gamification (Nah et al., 2013)

Stott & Neustaedter (2013) conducted a literature review on application of gamification in the classroom setting. They uncovered that some concepts are consistently successful than others (Table 7). After presenting some case studies, they conclude that there is no once-size-fits all model for the successful gamification of a classroom and that education is already using (more or less) elements that gamification has to offer (sometimes under different names).

Concept	Description
Freedom to Fail	If students are encouraged to take risks and experiment, the focus is taken away from final results and re-centered on the process of learning instead. The effectiveness of this change in focus is recognized in modern pedagogy as shown in the increased use of formative assessment "Encouraging learners to explore content, take chances with their decision making, and be exposed to realistic consequences for making a wrong or poor decision".
Rapid Feedback	Feedback is already a key element in education, "continual feedback to learners in the form of self-paced exercises, visual cues, frequent question-and-answer activities, a progress bar, or carefully placed comments by non-player characters
Progression	Recognized in modern pedagogy as scaffolded instruction. In gamification, this means incorporating lower order thinking skills into the first stages (identifying, remembering, understanding), progressing to higher order thinking skills in subsequent levels (analyzing, evaluating, critiquing, summarizing) and finally arriving at the highest order thinking skills in the final levels.

Storytelling	Providing a unifying story throughout a curriculum can put the learning elements into a realistic context in which actions and tasks can be practiced, something that is considered extremely effective in increasing student engagement and motivation

Table 7 Gamification concepts (Stott & Neustaedter, 2013)

Nah, Zeng, & Telaprolu (2014) conducted a literature review and identified eight game design elements that are used extensively in the educational and learning contexts (Table 8).

Game element	Description
Points	The point system functions as a measure of success or achievement. These points may be used as rewards, as a form of investment for further progression towards the goals, or to indicate one's standing.
Levels/Stages	The level system is used in various game designs to give players a sense of progression in the game. Initial levels tend to require less effort and are quicker to achieve, whereas the advanced levels require more effort and skills
Badges	Badges are recognized as a mark of appreciation or task accomplishment during the process of goal achievement. In order to maintain learners' motivation, the use of badges is helpful for engaging the learners in subsequent learning tasks
Leaderboards	The objective of a leaderboard is to keep the learners motivated and create a sense of eagerness to advance their names for the achievements they have accomplished. Leaderboards are used to create a competitive environment among students
Prizes and Rewards	The use of prizes has been found to be effective in motivating learners
Progress bars	Progress bars are used to track and display the overall goal progression. In an educational game, progress bars are used as a display mechanism to motivate people who are close to achieving their educational goal or sub-goals.
Storyline	Storyline refers to the narrative or story in the game. A storyline also provides a context for learning and problem solving as well as helps to illustrate the applicability of concepts to real-life
Feedback	The more frequent and immediate the feedback is, the greater the learning effectiveness and learner engagement. Clear and immediate feedback has been shown to be important for attaining the flow state, which is a state of engagement and immersion in an activity

Table 8 Game design elements (Nah et al., 2014)

3.3.3 Literature review conclusions

In this section we will summarize the guidelines discovered in the previous sections and draw our main conclusions derived from the literature study.

At the end of the literature review it was very clear that the impact of gamification depends a lot on the context to which it is applied. Some contexts are not appropriate: Fitz-Walter & Wyeth (2013), noted that adding competition elements to their application could have generated dangerous driving behavior in the players and Hamari (2013) posits that, in a utilitarian service, gamification features will be mostly ignored by the users.

In Table 9 we present the guidelines about gamification elements along with their objectives and properties. In Table 10 we present the meta-guidelines related gamification in education. The way these guidelines will be used to gamify the digital interactive tutorial is detailed in Sections 4.2 and 4.3

Reference	Game element	Objective	Properties
(Farzan & DiMicco, 2008), (De Schutter & Vanden Abeele, 2014), **(Barata et al., 2013), (Charles et al., 2011),** (Nah et al., 2013), **(Nah et al., 2014), (O'Donovan et al., 2013), (Singer, 2013)**	Point-based system	Extrinsic motivators — even intangible ones like points or levels — are appropriate for jumpstarting a new behavior and for motivating routine work	boosted users' contributions when introduced
(Li et al., 2012)	Score counter	To increase the engagement and foster the use of the system when rewards the users for desired behaviors within the system.	Competition, instantaneous feedback, status in relation to the social group.
(De Schutter & Vanden Abeele, 2014), **(Barata et al., 2013),** (Nah et al., 2013), **(Nah et al., 2014), (O'Donovan et al., 2013), (Singer, 2013)**	Rankings/ leaderboards	To increase the engagement and the time of use through competition created among the users, besides providing status to the user in relation to the other members of the social network.	Competition, status in relation to the social group.
(Meyer et al., n.d.), **(Nah et al., 2014), (Singer, 2013)**	Progress Bar	To make the progression clear for the user when performing a certain activity, providing constant feedback and stimulating to complete the tasks	Constant feedback, feeling of progress, incentive to fulfill tasks
(De Schutter & Vanden Abeele, 2014), **(Li et al., 2012), (Barata et al., 2013), (Charles et al., 2011),** (Nah et al., 2013), **(Stott & Neustaedter, 2013), (Nah et al., 2014), (Singer, 2013)**	Challenges/ Missions/ Levels	To guide the user towards the best possible experience in the application, optimizing the use of the developed features	Objective, Sporadic Feedback, Optimization of the experience of use
(Denny, 2013), (Hakulinen et al., 2013), (Barata et al., 2013), (Foster & Sheridan, 2012), (Meyer et al., n.d.), (Nah et al., 2013), **(Nah et al., 2014), (O'Donovan et al., 2013)**	Badges/ achievement system	To increase the engagement through the reward in the fulfillment of small objectives, providing sporadic feedback to the user and building the status of the user in relation to the group	Objective, Sporadic Feedback, Status in relation to the group.

(Nah et al., 2014), (O'Donovan et al., 2013)	Gifting	To conquer new users and engage those already existent through the sharing/awarding of items and information.	Sharing of items and information, Increase the scope of the application, Collaborative users
(Fitz-Walter & Wyeth, 2013), (Li et al., 2012), (Nah et al., 2013)**, (Stott & Neustaedter, 2013), (Nah et al., 2014), (O'Donovan et al., 2013)**	narrative/ storyline	To make the experience more emotionally appealing to users.	
(Fitz-Walter & Wyeth, 2013), (Nah et al., 2013)**, (O'Donovan et al., 2013)**	virtual currency	To increase the realism of a game	Competition can be enhanced and made more realistic
(Fitz-Walter & Wyeth, 2013), (Li et al., 2012)	Progressive Disclosure	To ensure that the challenges in the gamified application match the player's skill levels.	The system could provide more strict guidance to a novice user or more freedom to proficient learners.
(Fitz-Walter & Wyeth, 2013)	player could chose artifacts to use		
(De Schutter & Vanden Abeele, 2014)	character level ups		
(De Schutter & Vanden Abeele, 2014)**, (Charles et al., 2011),** (Meyer et al., n.d.)**,** (Nah et al., 2013)	avatar	the visibility of a student in the learning environment; greater visibility of students in a learning environment and a clear explanation of their role can encourage their engagement	
(Li et al., 2012)	Time Pressure	To establish clear and challenging goals	
(Fitz-Walter & Wyeth, 2013), (Li et al., 2012), (Nah et al., 2013)	visual/ audio effects	To ensure high engagement levels	
(De Schutter & Vanden Abeele, 2014)**, (Charles et al., 2011),** (Nah et al., 2013)	social play	Students are encouraged to collaborate with fellow students to complete challenges. This cooperation helps provide a supportive environment in which students with mixed ability can all enjoy the game. It also implies that it is beneficial to make the game a group activity where that is	

		practical.	
(Nah et al., 2013)	Onboarding	to sustain user engagement	A scaffolding method that can help players progress and advance from a novice to an expert or master
(Stott & Neustaedter, 2013)	Freedom to Fail	To encourage players to explore content, take chances with their decision making, and be exposed to realistic consequences for making a wrong or poor decision	
(Singer, 2013)	Normative Behavior	Make explicit what normative behavior should be by continuously publishing the behavior of players, positively emphasizing desirable behavior.	
(Nah et al., 2013)**,** **(Singer, 2013)**	Triggers	Use notifications to cue players to do something by directing their attention to a task related to the practice.	To support motivation, associate triggers with positive feedback or a goal to be reached. Do not overload players with triggers.

Table 9 Selected effects of game elements in gamification of learning/education

Reference	General guidelines
(Farzan & DiMicco, 2008), (Fitz-Walter & Wyeth, 2013)	Offer players the possibility to opt-out from the gamified application.
(Farzan & DiMicco, 2008)	Offer players the ability to customize the application.
(Fitz-Walter & Wyeth, 2013)	Gamification shouldn’t come at the cost of utility and usability.
(Fitz-Walter & Wyeth, 2013), (De Schutter & Vanden Abeele, 2014)**, (Domínguez et al., 2013),** (Nah et al., 2013)	Include a tutorial that teaches the player the rules of the gamified application.
(De Schutter & Vanden Abeele, 2014)	Evaluate the player’s knowledge of the gamified application.
(Fitz-Walter & Wyeth, 2013)	Identify and address early the potential for cheating in the gamified application.
(De Schutter & Vanden	Know who is the target population is.

Abeele, 2014)	
(De Schutter & Vanden Abeele, 2014)	Provide students with freedom of choice in how they want to show their mastery of the materials.
(De Schutter & Vanden Abeele, 2014)	Create real-world quest names
(De Schutter & Vanden Abeele, 2014)	Communicate that the course will be challenging
(Fitz-Walter & Wyeth, 2013), (Domínguez et al., 2013), (Charles et al., 2011), (Nah et al., 2013)**, (Stott & Neustaedter, 2013), (Nah et al., 2014)**	Offer immediate feedback
(Charles et al., 2011)	Fun -> The game is essentially an additional voluntary activity (not graded). It should be fun but it should also have an educational benefit. In addition, it should have no negative impact on the course itself, or on any other course being studied at the same time.

Table 10 Selected general guidelines for gamification of learning/education

4 Design of a gamified digital interactive tutorial for operations research

In this section we will be answering RSQ4: "How is a gamified digital interactive tutorial for operations research designed?" We show how the guidelines identified in Chapter 3.2, the learning outcomes defined in Chapter 3.1.1, and the gamification document from Chapter 3.3 are combined in order to design the gamified digital tutorial.

The tutorial has been refined and improved by conducting user testing with three users from the target population. User testing consisted of presenting the application to one of the users and asking him to interact with all its features. The feedback of the user would be integrated in the application and the new version of the application would be presented to the next user. Considering that there were three users in our testing group, the application was reviewed in three iterations.

As mentioned is Section 2.2.3.2, in our design of the non-gamified tutorial and the gamified tutorial we aimed to make them identical, apart from the aspects of gamification. This gave us the confidence that the differences that we observed during the experiment were caused by gamification.

4.1 Operations research digital interactive tutorial design

Having a summary of the guidelines for the creation of digital interactive tutorials and the defined learning outcomes, we proceeded to the design of the OR digital interactive tutorial. During design, we followed the basic principle that the design choices should not be constricted by the technical possibilities, thereby focusing on what should be there, rather than what can easily be implemented.

As mentioned before, the tutorial is aimed at decision makers in companies. These people have generally little time, don't have a technical profile and are geared towards generating business value. Because of this, it is very important to balance the technical details given in the tutorial (how-to and principles knowledge) with the business value details. Too many technical details might scare away decision makers; too little might not be enough to take this innovation to the next stage (Persuasion) in the diffusion of innovations cycle. The tutorial was designed in such a way that it contains more technical knowledge than business value knowledge. We focused on simplifying the technical knowledge included, but not more than necessary to understand the OR concepts.

In Table 11 we have summarized how the design of the digital interactive tutorial maps to the guidelines in Table 5.

Guideline no	Digital interactive tutorial requirements
1	The tutorial focuses on the real task of solving operations research problems. Problems presented in the tutorial come directly from the resources on learning operations research and they are (simplified) but realistic problems a manager might encounter on the job (product planning, resource allocation, and transportation). The choice of a specific problem for each section was completely random. Navigation bar that gives easy access to the main pages of the tutorial. No section of the tutorial is mandatory.
2	Start immediately with an operations research problem. Keep introductions as short as possible. Offer information only when it is absolutely necessary.
3	Added a question after explaining every important concept. Added a hint and a reference for further reading for concepts that allowed it. Add a concept map that relates all the concepts together. As mentioned before, we decided that taking an exploratory learning approach is not applicable due to the time constraints of the project.
4	Break down tutorial in pages with specific purpose (explain the value of operations research, overview of operations research, problems that can be solved with operations research).
5	Not applicable due to the context (the users are going through a tutorial, they can't get the tutorial into an error state)
6	Not applicable due to the context (the users are going through a tutorial, they can't get the tutorial into an error state)
7	Because the assumption of this tutorial is that it will be presented to operations research agnostic users we did not aim to use familiar tasks or metaphors for the presentation of the content. We will use an informal tone to present the content. We made the assumption that users are well acquainted with modern websites, therefore we decide to make the tutorial behave and look like any other modern website.
8	The problems presented in the tutorial will also be used in the exercises.
9	Get user feedback at different stages of development with 3 users from the target population.
10	By this point we already have clearly defined user task (specified as learning objectives).

	We will make sure that the implementation respects them: there will be a section emphasizing the value of the operations research, have a specific section with the operations research concepts and have a section that explains how linear programming problems work.
11	Not necessary since there is user testing planned with users from the target population. User testing with naïve users would not add any value.
12	Not applicable due to the context (the users are going through a tutorial, they can't get the tutorial into an error state, hence no need for an error help table)
13	We will try to use images as much as possible throughout the tutorial.
14	Not applicable due to the context.
15	Test tutorial at different stages of development with 3 users from the target population.
16	Not applicable due to the context (interface is too simple to require any guidance).
17	Not applicable due to the time constraints of the project.
18	Animated documentation not applicable due to the time constraints of the project. Not applicable due time constraints of the experiment (videos would play without sound in the laboratory)

Table 11 Mapping of the guidelines to the design of the tutorial

There are some decisions that we made that are not related to any guidelines:

- We decided to put two problems in the tutorial because we wanted to make sure that the users get enough contact with OR and see how it is applied in various scenarios.

After the digital interactive tutorial design was ready, we proceeded to the implementation. During implementation we encountered a couple of issues that forced us to depart from the original design:

- Adding a question after explaining every important concept was implemented as a question related to one of the operations research problems. During user testing, it was demanded that the concepts would first be summarily presented separately and then, later, when the question would appear, the definition of the concept would be presented again. This differs from the guideline.
- During user testing, it came up that the user would get "lost" during the tutorial, not knowing what to do next. To prevent that, at the end of every page, we added buttons that would direct the user to the next logical section of the tutorial.
- The OR concepts related to the "Graphical representation of a two variable LP" (Contour of the objective function, Feasible region, Optimal corner solution) were too hard to represent visually in a way that would make sense and they would confuse the users more than they would explain a concept. We decided to leave them out of the tutorial

The rest of the feedback we obtained during user testing was mostly related to user experience. The other suggestions for improvement were in the lines of the guidelines. We can observe from this that the guidelines that we followed during design are a very good starting point for creating digital interactive tutorials.

Screenshots of the main sections of the resulting artifact can be found in Appendix 7.7.

4.2 Gamified operations research digital interactive tutorial design

Having a summary of the guidelines for the creation of gamified applications for learning/education, we proceeded to the design of the gamified version of the operations research digital interactive tutorial. During the design, we followed the basic principle that the design choices should not be constricted by the technical possibilities, thereby focusing on what should be there, rather than what one can easily implement.

The scientific literature does not offer a method/process that details the steps that one should take when starting to gamify an existing application. We had to turn to the industry to obtain that answer. By searching on gamification forums and mailing lists, we found three methods that explain what needs to be done when gamifying the application:

- 6D framework by (Werbach & Hunter, 2012), detailed in Section 4.3.
- The LevelUP Gamification Design Process[5], a seven step process: Love (for game design), Elaborate your strategy and get ready (get your team ready and think of the client's main problem), Visualize WHY, WHAT, WHO (understand your target players and what you want them to do), Explore a new world (make a small draft of the gamification theme and story), Level UP your mechanics (decide what gamification mechanics you will use), Upgrade your graphics (what kind of aesthetics and graphics you want to design and), Playtest, Playtest, Playtest (very important)
- 4×4 Framework[6], positing eight phases: WHAT is being gamified, WHY it is being gamified, WHO are the users, HOW is it being gamified, ANALYTICS are set up, TESTED with users, ACTED/ITERATED on feedback, RELEASED the solution. It also contains seven meta recommendations.

We could not find any good use case references for either, therefore we decided to use the 6D framework. This choice was influenced by a couple of factors:

- Personal preference, for the main researcher had already read the book written by the authors, and successfully completed an online gamification course[7] given by the same authors.
- Better scientific potential: the framework is supported by a book and a gamification course, being the most detailed in the set of gamification frameworks. Additionally, the authors are academics.

4.3 6D Framework explained and applied

4.3.1 Define Business Objectives

Which are the achievements of the project, of the system, of the campaign?

Process:

- *Make a list as concrete as possible and rank them*
- *Eliminate the things that are not a final business objective (go through your list and cross off anything that is a means rather than an end)*

[5] http://www.epicwinblog.net/2014/01/the-levelup-gamification-design-process.html
[6] http://marczewski.me.uk/gamification-framework/
[7] https://class.coursera.org/gamification-003

- *Justify objectives … why is that something to achieve?*

Motivate learners to assimilate all the OR how-to and principles knowledge as defined by the ILOs (Section 3.1.1).

4.3.2 Delineate target behavior

Specify the desired players' behaviors and how you'll measure them. They should promote the ultimate business objectives you previously defined, though the relationship may be indirect. Come up with as many possible behaviors as you can. Consider the win states.

Starting from the objectives defined at the previous point (motivate learners to assimilate all the operations research how-to and principles knowledge as defined by the ILOs), we have identified the following target behaviors (Table 12).

Target behaviors	Measured by
Players open all information pages of the tutorial	Players spend at least 4 minutes on each main page (pages of the top horizontal menu) and should open any other page.
Players solve all of the quizzes	Validating that a player has submitted responses to each quiz.
Players use other sources of information about OR during using the tutorial	Google searches on the general topic of OR performed during the tutorial.
Players interact with each other to clarify concepts	Activity in the instant messaging chat group (messages sent)

Table 12 Target behaviors and their measurements

Regarding the win states, we decide that, since this is a short lived/scope gamified application, we will not consider creating localized or temporal win states. And there would be no clearly stated winning condition. The player can choose to focus on earning all the points available in the application (thus reaching the ultimate level) or can choose to collect all the existing badges, but that will not make him/her a winner. There are certain badges and points that will only be awarded if the player successfully finishes a challenge (presents in the form of quizzes).

4.3.3 Describe yours players

What do you know about the players (demographics, age groups, psychographics, kind of behavior etc.)? Who are they? What is their relationship to the one that is offering the gamified application? What might motivate the players? Think about what demotivates the players. Consider splitting the players into player types (e.g. Bartle's player types). Create the images of the typical players. Consider the player lifecycle (newbie -> regular -> expert).

We will not be using Bartle's player types, as Bartle himself observes that they apply only to multi user dungeon games (Bartle, 2012). Furthermore, since the gamified tutorial has a very limited scope, we will create the gamification application for a single player type, which we will define in the following paragraphs.

The players are international and Dutch students (22-30 years old) of a master program at a Dutch University. Players are part of the millennial generation and have a high education.

We can further define the players by referring to target audience of the digital interactive tutorial: early adopters. Rogers (2010) states that this ideal category has the following characteristics:

- are integrated in the local social system
- has the highest degree of opinion leadership
- "the individual to check with" before adopting a new idea
- serve as a role model for many other members of a social system
- respected by his or her peers
- must make judicious innovation-decisions

Rogers (2010) also specifies what the differences are between earlier adopters, compared to later adopters:

- Are no different in age
- have more years of formal education
- are more likely to be literate
- have higher social status
- have a greater degree of upward social mobility (are on the move in the direction of still higher levels of social status)
- have larger-sized units (farms, schools, companies, and so on)
- have greater empathy. Empathy is the ability of an individual to project himself or herself into the role of another person.
- may be less dogmatic. Dogmatism is the degree to which an individual has a relatively closed belief system, that is, a set of beliefs which are strongly held.
- have a greater ability to deal with abstractions
- have greater rationality. Rationality is use of the most effective means to reach a given end.
- have more intelligence.
- have a more favorable attitude toward change.
- are better able to cope with uncertainty and risk.
- have a more favorable attitude toward science
- are less fatalistic. Fatalism is the degree to which an individual perceives a lack of ability to control his or her future.
- have higher aspirations (for formal education, higher status, occupations, and so on)
- have more social participation
- are more highly interconnected through interpersonal networks in their social. Connectedness is the degree to which an individual is linked to others.
- are more cosmopolite. Innovators' interpersonal networks are more likely to be outside, rather than within, their system
- have more contact with change agents.

- have greater exposure to mass media communication channels.
- have greater exposure to interpersonal communication channels.
- seek information about innovations more actively.
- have greater knowledge of innovations.
- have a higher degree of opinion leadership.

From these characteristics, we will infer the general image of the player of the gamified tutorial:

- is ageless
- has a high education
- finds social status modifiers important
- is up to speed with IT innovations (high expectations from IT)
- can grasp abstract concepts quite easy (they don't need a lot of hand holding)
- finds social participation important
- is an opinion leader

For the case of this gamified application, the player lifecycle will not be important. Our use case does not allow the user to become a regular user (with medium experience of operations research).

In our case, the players might lack desire to learn about optimization modelling or its business value, we will need to focus on an engagement-oriented approach combined with a progression system.

4.3.4 Devise activity loops

*There are two kinds of cycles to develop: engagement loops (*Figure 10*) and progression stairs. Engagement loops describe, at a micro level, what the players do, why they do it, and what the system does in response. Progression stairs give a macro perspective on the player's journey.*

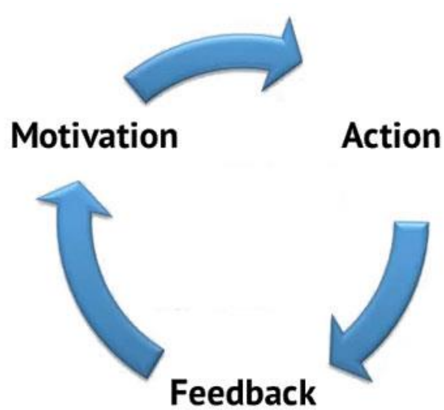


Figure 10 Engagement loop

We will be devising the activity loops based on the guidelines of gamification presented in Section 3.3.3. In addition to that, we need to refer to the target behaviors and to the characteristics of the player.

4.3.4.1 Engagement loops

The design of the engagement loops starts from motivation. The engagement loops for the gamified application are detailed in Table 13. The feedback is given through an experience points (EXP) based system.

Motivation	Action	Feedback	Max EXP
Open all information pages	Player clicks on button or link inside tutorial.	Awarded only once for the first visit: 10 points; 13 concepts; 4 levels on intro problem; 4 major pages; 2 LP examples, 1 concept map	250
Players solves a quiz	Player submits answer to quiz	Awarded only once: 5 points for first submission; 25 points for correct submission (if player did not look at solution already); 10 quizzes	300
Players use other sources of information about OR during using the tutorial	n/a – no means of measuring this from inside the tutorial.	n/a	0
Players interact with each other to clarify concepts	Player sends message (with more than 5 words) to chat group	Player receives a badge after 5 messages and a badge after 25 messages.	0

Table 13 Gamified tutorial engagement loops

4.3.4.2 Progression stairs

The progression stairs are represented by a level system (with a progress bar displayed to the user constantly in the lower right corner of the screen). All player actions will be connected to a points system and this, in turn, will be linked to the level system. These points are distributed in the following levels:

- Level 1: 0 XP
- Level 2: 50 XP
- Level 3: 160 XP
- Level 4: 320 XP
- Level 5: 540 XP

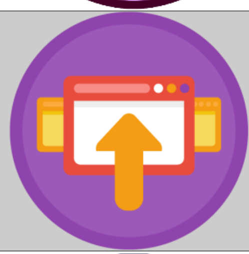
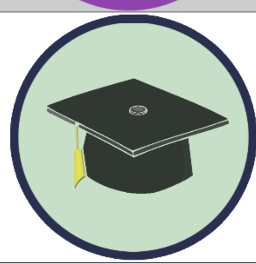
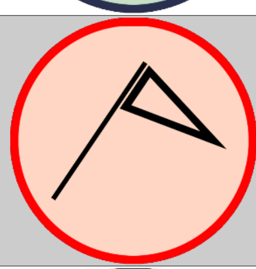
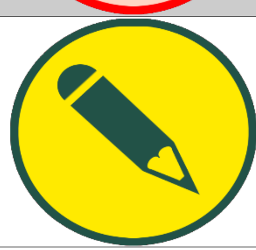
The current level, the score counter and the score bar are part of the user profile widget and will be present on the main pages of the tutorial.

The progression stairs are also visible on the “Why operations research page”: when a player submits a correct for a specific difficulty, the tab corresponding to that difficulty gets a green checkmark indicating that the player has completed that level.

4.3.4.3 Badges system design

The badge system is aimed at appealing to the social status characteristic of the players. The badges will always have a positive meaning (in order not to confuse the players). The order of the badges is also important, the first badges will be those that the user can earn in the close future, followed by the important badges, those that set higher level goals, followed by those that the user can earn whenever he wishes in time or are more for encouragement (and he will earn anyway as he progresses through the tutorial).

The badges use names that appeal to the players (early adopters).

Badge name	Badge icon	Badge description	Motivation for introducing the badge
Commentator		Comment (at least 5 words) in the chat room for 5 times	Aimed at encouraging the player to interact with other players. Appeals to the player characteristics of social participation and opinion leader.
Evangelist		Comment (at least 5 words) in the chat room for 25 times	
Feature explorer		Try all the features suggested on the “Using the tutorial” page. Hint: look for “try to…”	Strictly related to the tutorial explaining the rules of gamification, ensures that the players spend enough time to understand how the tutorial works
Academic		Read all the information in the tutorial. Hint: there are 25 windows with information in the tutorial	Encourages the target behavior “open all information pages”
Supporter		Spend at least 4 minutes, continuously, on the “Why operations research” page	Encourages a more specific part of the “open all information pages” target behavior.
Student		Spend at least 4 minutes, continuously, on the “Intro to operations research” page	

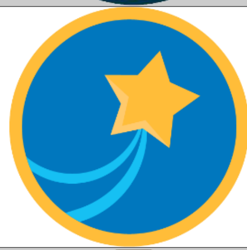
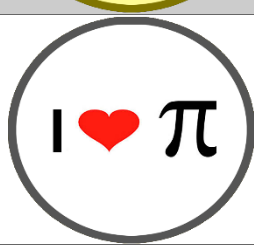
Enlightened		Spend at least 4 minutes, continuously, on the “Modeling a LP problem” page	
Rising star		Spend at least 4 minutes, continuously, on the “Modeling a LP problem 2” page	
Champion		Reach level 5	Encourages the player to reach the final level. Otherwise, reaching the final level would not be a special event.
Quiz whiz		Correctly answer all the 10 quizzes in the tutorial. Hint: you can submit your answer multiple times, as long as you don’t look at the answer	Rewards players that have engaged in the hard fun activity of solving the quizzes present in the tutorial.
Helping hand		You receive this badge when 2 other people have rated you as helpful in the chat room	Aimed at encouraging the player to interact with other players. Appeals to the player characteristics of social participation and opinion leader.
Math lover		Find the most profitable solution to all the versions of the problem on the “Why operations research” page	Rewards players that have engaged in the hard fun activity of solving the problem on the “Why operations research” page

Table 14 List of available badges in the tutorial

4.3.5 Don't forget the fun

Would players participate in your system voluntarily? If there weren't any extrinsic rewards offered, would they still be likely to play? If the answer is no, then you should think about what might make your system more fun. Use the 4 keys 2 fun by Nicole Lazzaro[8] as a starting point:

- *"Hard fun" is a challenge or puzzle, which is fun because of the pleasure of overcoming it.*
- *"Easy fun" is casual enjoyment, a way of blowing off steam without overly taxing yourself.*
- *"Altered states" we'll call experimental fun. It's the enjoyment of trying out new personas and new experiences.*
- *"The people factor" is essentially social fun: the kinds of fun that depend on interaction with others, even if competitive.*

Based on the previous descriptions and knowing the context of the gamified application, we recognize that this application will contain minimal (if any) easy fun and altered states fun. As a result, we will be focusing on expanding on the people factor and on the hard fun.

4.3.5.1 Hard Fun *(mastery)*

According to Lazzaro (2004), it creates emotion by structuring experience towards the pursuit of a goal. Keywords: frustration, fiero (personal triumph), relief, challenge, reward, accomplishment, focus, feedback on progress.

The hard fun is achieved through the vast amount of domain information that requires the player to understand it and then answer quizzes and eventually, solve a problem. The relief is ensured either when the player answers correctly to a question (fiero) or when the player decides to look at the provided correct answer (activity that is discouraged). The structured experience is offered by the progression stairs and the badges.

4.3.5.2 The people factor (relationships)

According to Lazzaro (2004), players also get enjoyment from playing with others inside or outside the game (they see games as mechanisms for social interaction). Keywords: player interaction (competition, cooperation, mentor), amici, amusement, amiero (social bonding), amidar (admiration).

The people factor fun is achieved through rewarding chat room activity, the livefeed, the Helping hand badge and dashboard. These features relate to player interaction, amidar and amiero.

4.3.6 Deploy appropriate tools

An extra page will be added at the beginning of the tutorial. It will explain how the badges can be earned and what the general rules of the gamified application are. It will also communicate that the content is going to be challenging, as this is not a video game. It will be made up of small videos on the following topics presenting & how to use: the badges; the point system; the levels and the dashboard.

One extra page will be added to the end of gamified tutorial. It will contain the user profile information, the badges drawer (enhancing mastery), a leaderboard, a live feed and a page where he/she can see other player's statistics (profile and badges) (enhancing competition).

---

[8] http://www.nicolelazzaro.com/the4-keys-to-fun/

There are two basic options for the technical implementation of gamified systems:

- build custom implementations yourself
- use one of the software-as-a-service offerings.

For this application we will be opting for the second choice as it allows for greater flexibility and for the separation of concerns between the systems.

By using the personal network of the main researcher and searching the web in general, we have identified and evaluated five platforms (that allow for the creation of a free developer account and offer an API), Table 15.

	Gioco	**Playlyfe**	**Behave.io**	**UserInfuser**	**Playbasis**
URL	http://gioco.pro/	https://playlyfe.com/	http://www.behave.io/	https://cloudcaptive-userinfuser.appspot.com/html/signup.html	http://www.playbasis.com/
Demo	no	no	no	no	http://demo.playbasis.com/
REST API	yes	no	yes	yes	yes
Other APIs	Ruby, JavaScript	JavaScript	JavaScript, php, ruby	Java, php, python, ruby	Android, iOS, php, ruby, python, .NET, JavaScript, Java
Developer docs	sufficient	insufficient	sufficient	sufficient	abundant
Gamification features	Badges, points, levels	Points, badges, virtual rewards, chance	Badges, points, levels, leaderboards	Only a container for gamification achievements	Quests, Levels, badges, leaderboards, points, virtual rewards
Widgets[9]	no	no	no	Trophy case, rank, points, notifier, leaderboard, milestones	Social login, leaderboard, livefeed, profile, userbar, achievement, notifier

Table 15 Free gamification platforms

Based on the results of our evaluation we have decided to use playbasis.com as the platform on which to implement gamification for this project.

Play-testing of the design, to see what might work and then see what actually does work, was done with 3 persons from the target group.

[9] Pieces of code that can easily be placed inside the gamified application and show various gamification features. Allow for very fast development of the gamified application.

In Table 16 and Table 17 we have summarized how the design of the gamified tutorial maps to the gamified guidelines in Table 9 and Table 10.

Gamification guideline	Relation to the gamified tutorial
Point-based system	Used across the application
Score counter	Part of the user profile widget
Rankings/leaderboards	Used in the player dashboard
Progress bar	Part of the user profile widget
Challenges/missions/levels	Levels are used in the progression stairs. The available challenges are described through the tutorial of gamification page (tooltips on badges).
Badges/achievement system	Used in the player dashboard.
Gifting	Not applicable due to the context
Narrative	Not applicable due to the time constraints of the project
Virtual currency	Out of project scope
Progressive disclosure	Not applicable due to player characteristics
Player could chose artifacts	Not applicable due to the context
Character level ups	Not applicable, not offering an avatar
Avatar	Not applicable due to player characteristics
Time pressure	Not applicable due to the context
Visual/audio effects	Out of project scope
Social play	Used in the chat feature (2 badges awarded based on number of messages sent)
Onboarding	Not applicable due to the context
Freedom to fail	Not applicable due to the context
Normative behavior	Used in the player dashboard (livefeed widget shows the behavior of the other players).
Triggers	Not applicable due time constraints of the experiment

Table 16

Meta gamification guideline	Relation to the gamified tutorial
Possibility to opt-out	Not applicable due to the experimental setup.
Possibility to customize the application	Not applicable due time constraints of the experiment
Don't sacrifice utility and usability	We will be aiming to keep the changes to the existing pages of the digital tutorial minimal.
Tutorial of gamified application	Present at the beginning of the tutorial. Additionally, the badge award notification contains the reason for which the badge was awarded. The level up notification contains a reference to the page on which the level ups are described.
Evaluate knowledge of the gamification rules	Not applicable, but a badge will be awarded after the user goes through all the tutorial (Feature explorer).
Know the target population	Used while applying the 6D framework ("Describe your players")
Freedom of choice for	Not applicable due to the context

showing mastery	
Real-world quest names	Not applicable as the tutorial will not contain quests.
Communicate that tutorial will still be challenging	Present in the gamification tutorial.
Immediate feedback	Used while applying the 6D framework ("Devise activity loops"). For every level up and badge earned the user will be notified of it via a popup.
Fun	Used while applying the 6D framework ("Don't forget the fun")
Identify and address potential of cheating	Used in the quizzes: keyword validation. Used during the badge system design.

Table 17

4.3.7 Implementation

After the gamification design document was ready, we proceeded to the implementation of the gamification system. We encountered a couple of issues that forced us to depart from the original design:

- The order in which the badges are shown to the user could not be respected (the gamification vendor did not support it).
- For the tutorial of the gamified application we wanted to create explanatory videos. However, the initial experimental conditions (not allowing sound) would make the videos useless. We have decided to add textual descriptions for the features and ask the players to perform actions related to those features.
- The feature that would allow players to see other players' achievements drawer and profile proved to be technically very hard to implement. Therefore, we have decided not to implement it.

The user feedback for the gamified tutorial was in general related to bugs in the application (the action taken here was to fix the bugs) or to suggestions for improving the usability (most of the time, the decision was not to improve since the same issues would be found in the non gamified application and this would not influence the results of the experiment). Some comments concerned ways in which the existing gamification mechanisms could be improved (these were integrated in the gamified tutorial with high priority).

Screenshots of the most important gamification features of the resulting artifact can be found in Appendix 7.8.

5 Evaluation

In this chapter we will discuss the details related to the evaluation of the tutorial for OR. We will be describing how the process of delivering the experiment went (Section 5.1), what data we collected (Section 5.2) and what conclusions can be drawn by looking at the available data.

5.1 Experiment delivery

Participants were selected for the experiment using convenience sampling: all the 111 active students of the Business Informatics master in the Utrecht University have agreed to participate in the experiment. In the end, 30 students participated in the experiment.

It would have been preferable to achieve a statistically significant sample size. However, due to the limited availability of the participants from the target population, this would have implied including subjects that were not part of the envisioned target population. As such, we decided to privilege the fit of the sample with the target population over sample size.

5.1.1 Onsite group

9 “onsite” participants did the gamified version of the tutorial. Besides the reason already provided (to better control variables), we decided to create this group in order to test the tutorial’s social gamification features (the chatroom and the livefeed), which required that participants are simultaneously logged in the tutorial.

Participants were instructed saying that if a section feels too hard they should leave it and move on. They should not worry too much about it. They were told that the experiment consists of a pre questionnaire, a tutorial and a post questionnaire. The aim of the post questionnaire is to measure how well the content of the tutorial was delivered. The participants were asked explicitly not to cheat.

The experiment started at 13:10 local time on 13th November 2014 on the premises of Utrecht University. At 14:05 the participants were asked to wrap up their work on the tutorial in the next 15 minutes, so that at 14:20 (the latest) everyone would be working on filling in the post-questionnaire. The delivery of the experiment went according to plan: we collected all needed data (responses to questionnaires and user action tracking data) for 9 participants doing the gamified tutorial in an “onsite” setting.

5.1.2 Online groups

After the onsite experiment was delivered, the participants (the two online groups mentioned before) unable to attend the onsite experiment were emailed (see appendix 7.9) the necessary information so they could start participating in the online experiment.

When almost all the participants had finished the experiment, we realized that, due to a flaw the user action tracking system, we were missing all the user action data. This meant that, for the 13 participants (6 non-gamified, 7 gamified) that we had up to that moment, we could only analyze the data from the pre-questionnaire and the post-questionnaire and we did not consider this sufficient. We decided to fix the user tracking defect, create a third group (9 participants) and deliver to it, online, the non-gamified tutorial. We managed to deliver it to only 8 participants (not 9 as planned) and this time, we collected all needed data.

We present a summary of the delivery of the experiment to the online groups:

- 6 participants did the non-gamified tutorial online, no user tracking data available
- 7 participants did the gamified tutorial online, no user tracking data available
- 8 participants did the non-gamified tutorial online, all data available

5.2 Summary of collected data

In this section we will present the data we have obtained by running the experiment. As mentioned in the previous section, the experiment was ran in multiple phases and data was collected from four distinct groups, Table 18. We decided that it makes most sense to compare the groups based on the availability of user tracking data (i.e. compare the group1 to group2 and group3 to group4). We will also look into common characteristics of the groups based on the tutorial type (i.e. compare group1 to group3 and group2 to group4).

	Group1	Group2	Group3	Group4
Participants	6	7	8	9
Delivery method	online	online	online	onsite
Tutorial type	non-gamified	gamified	non-gamified	gamified
Tracking data availability	no	no	yes	yes

Table 18 Experimental groups characteristics

5.2.1 Experiment without tracking data

Below we present all relevant data that was generated by running the experiment for which the participant tracking data was not collected, which includes only online participants. This data is analyzed further in Section 5.3.

Table 19 illustrates the answers of the participants to the first section of the post experiment questionnaire. This section measured the participants' satisfaction with the tutorial system and contained 8 questions that allowed answers on a Likert scale (minimum 1, maximum 5).

	Non-gamified (n=6)		Gamified (n=7)	
	$\bar{x}$	σ	$\bar{x}$	σ
I enjoyed using this system	3.17	1.47	3.29	0.95
Completing tasks was frustrating	3.33	0.82	2.86	1.07
It makes completing tasks fun	2.17	1.17	3.43	0.98
Completing tasks was difficult	3.67	0.82	3.29	0.76
It is an effective learning tool	3.17	0.98	3.71	0.49
Had to work hard to complete tasks	3.33	0.82	3.29	0.95
It is an engaging experience	3.50	1.52	3.29	1.11
I felt rushed when completing tasks	1.83	0.98	2.86	1.07

Table 19 Satisfaction with the tutorial system

Table 20 is an aggregated table. It presents the characteristics of the two groups of the experiment side by side.

- The column **Id** represents the username assigned to the participant for the purpose of the experiment.

- The column **Game XP** represents the score the participant obtained for the section measuring experience with video games from the pre experiment questionnaire (min: 0, max: 10).
- The column **Competitiveness** represents the score the participants obtained for the section measuring the interpersonal competitiveness from the pre experiment questionnaire (min: 1, max: 29).
- The column **OR knowledge** presents the results from the pre-questionnaire concerning the OR knowledge of participant before opening the tutorial. It allows only yes (the participant answered all control questions correctly) or no values. We decided to leave the participants that had a yes on this row out of the analysis because the scope of the tutorial and of the questionnaires did not account for them.
- The column **OR score** contains the results that the participants obtained by doing the OR knowledge test in the second section of the post experimental questionnaire (min 0, max 15).
- The column **Time (min)** expresses (in minutes) how much time the participants spent in the tutorial (from the moment they logged in, to the moment they started the post experimental questionnaire). N/A means that it was impossible to calculate time because the participant did not start the post questionnaire immediately. These participants were left out of the analysis.
- The final row (x̄) represents the average for some of the characteristics. It does NOT take into account the crossed-out participants.
- Note that the sample size is smaller than the one presented in Table 18. As mentioned before, we had to leave out some of participants because they had previous experience with OR (user5, user6, user32, user37) or because it was not possible to determine how long they spent in the tutorial (user38, user6, user32).

Non-gamified (n=4)						Gamified (n=4)					
Id	**Game XP**	**Competi-tiveness**	**OR knowledge**	**OR score**	**Time (min)**	**Id**	**Game XP**	**Competi-tiveness**	**OR knowledge**	**OR score**	**Time (min)**
user3	8	10	no	6	65	~~**user38**~~	~~9~~	~~12~~	~~no~~	~~5~~	~~N/A~~
~~**user5**~~	~~6~~	~~13~~	~~yes~~	~~9~~	~~27~~	**user34**	10	20	no	4	42
user7	9	12	no	7	57	**user33**	4	18	no	5	36
user10	1	14	no	6	50	**user36**	6	20	no	3	11
~~**user6**~~	~~0~~	~~21~~	~~yes~~	~~12~~	~~N/A~~	~~**user37**~~	~~0~~	~~17~~	~~yes~~	~~10~~	~~31~~
user2	9	16	no	5	57	**user31**	4	26	no	7	30
						~~**user32**~~	~~8~~	~~23~~	~~yes~~	~~3~~	~~N/A~~
x̄		13		6	57.25	**x̄**		21		4.75	29.75
σ		2.6		0.81	6.13	**σ**		3.5		1.70	13.42

Table 20 Characteristics of the two groups of participants

5.2.2 Experiment with tracking data available

Below, we present all the data that was generated by running the no-tracking experiment. This data is analyzed further in Section 5.3.

Table 21 illustrates the answers of the participants to the first section of the post experiment questionnaire. This section measured the participants' satisfaction with the tutorial system and contained 8 questions that allowed answers on a Likert scale (minimum 1, maximum 5).

	Non-gamified (n=7)		Gamified (n=8)	
	x̄	σ	x̄	σ
I enjoyed using this system	3.43	1.51	3.50	1.20
Completing tasks was frustrating	2.29	1.38	2.75	1.16
It makes completing tasks fun	2.86	0.69	3.50	1.07
Completing tasks was difficult	3.29	1.11	3.38	1.19
It is an effective learning tool	3.43	1.51	2.75	1.28
Had to work hard to complete tasks	3.29	1.11	3.25	1.16
It is an engaging experience	3.29	0.95	3.75	1.04
I felt rushed when completing tasks	2.14	1.07	3	0.93

Table 21 Satisfaction with the tutorial system

Table 22 is an aggregated table. It presents the characteristics of the two groups of the experiment side by side. Table 22 is the same as Table 20 in terms of structure, just the data values differ. For the explanation of the table headers, we refer the reader to the description of Table 20.

- The column **Id** represents the username assigned to the participant for the purpose of the experiment.
- The column **Game XP** represents the score the participant obtained for the section measuring experience with video games from the pre experiment questionnaire (min: 0, max: 10).
- The column **Competitiveness** represents the score the participants obtained for the section measuring the interpersonal competitiveness from the pre experiment questionnaire (min: 1, max: 29).
- The column **OR knowledge** presents the results from the pre-questionnaire concerning the OR knowledge of participant before opening the tutorial. It allows only yes (the participant answered all control questions correctly) or no values. We decided to leave the participants that had a yes on this row out of the analysis because the scope of the tutorial and of the questionnaires did not account for them.
- The column **OR score** contains the results that the participants obtained by doing the OR knowledge test in the second section of the post experimental questionnaire (min 0, max 15).
- The column **Time (min)** expresses (in minutes) how much time the participants spent in the tutorial (from the moment they logged in, to the moment they started the post experimental questionnaire). We had to exclude a participant (user42) from the data analysis entirely, since the time he spent in the tutorial (5 minutes) made him irrelevant for our study.
- The final row (**x̄**) represents the average for some of the characteristics. It does NOT take into account the crossed-out participants.
- Note that the sample size is smaller than the one presented in Table 18. As mentioned before, we had to leave out some of participants because they had previous experience with OR (user46, user12)

Non-gamified (n=7)						Gamified (n=8)					
Id	Game XP	Competi-tiveness	OR knowl-edge	OR score	Time (min)	Id	Game XP	Competi-tiveness	OR knowl-edge	OR score	Time (min)
user40	9	7	no	2	43	**user11**	8	17	no	0	62
user42	4	9	no	4	19	~~**user12**~~	~~10~~	~~12~~	~~yes~~	~~8~~	~~45~~

user43	9	20	no	6	18	**user13**	8	20	no	2	61
user45	7	24	no	6	42	**user14**	6	13	no	3	49
~~**user46**~~	~~8~~	~~13~~	~~no~~	~~3~~	~~5~~	**user15**	0	19	no	2	58
user47	2	17	no	14	42	**user16**	10	19	no	3	57
user49	4	16	no	7	17	**user17**	5	12	no	5	58
user51	6	25	no	4	70	**user18**	9	15	no	2	62
						user19	7	6	no	6	59
x̄		16.86		6.14	35.86	**x̄**		15.12		2.87	58.25
σ		6.91		3.85	19.37	**σ**		4.70		1.88	4.20

Table 22 Characteristics of the two groups of participants

Table 23 illustrates the time (expressed in minutes) that each participants spent in each major section of the tutorial:

- P1 refers to “Why optimization modeling”
- P2 refers to “Intro to optimization modeling”
- P3 refers to “Modeling LP problem 1”
- P4 refers to “Modeling LP problem 2”

One particularity about Table 23 is the presence of the * symbol next to some participants’ name (i.e. user13 and user15). The user tracking data collected from these participants had to be manually checked and cleaned because it contained erroneous entries (participants had more than one page of the tutorial was opened at the same time).

Non-gamified (n=7)					**Gamified (n=8)**				
	P1	**P2**	**P3**	**P4**		**P1**	**P2**	**P3**	**P4**
user40	11.73	2.37	12.10	16.25	**user11**	19.30	24.52	8.72	5.87
user42	14.13	1.82	2.12	0.43	~~**user12**~~	~~24.92~~	~~4.43~~	~~11.90~~	~~2.25~~
user43	1.90	4.43	6.95	4.57	**user13***	14.92	6.03	14.23	22.43
user45	9.80	12.98	11.08	7.63	**user14**	10.48	11.28	15.03	6.73
~~**user46**~~	~~2.02~~	~~1.48~~	~~0.67~~	~~0.43~~	**user15***	11.40	8.32	9.15	22.23
user47	8.43	10.00	18.30	4.88	**user16**	18.02	7.42	10.43	11.32
user49	11.78	1.05	0.75	3.55	**user17**	15.42	17.22	7.93	9.90
user51	54.97	7.53	5.63	0.98	**user18**	7.52	21.30	20.58	8.95
					user19	7.78	11.23	12.08	21.25
x̄	16.11	5.74	8.13	5.47	**x̄**	13.11	13.42	12.27	13.59
σ	17.57	4.55	6.14	5.34	**σ**	4.48	6.82	4.23	7.16

Table 23 Time (minutes) spent by each participant in each of the major sections of the tutorial

Table 24 illustrates the number of times each page of the tutorial was opened. There are two pages (dashboard and gamificationTutorial) that were not available in the non-gamified tutorial (we did not consider these pages in the computation of the percentage of total actions). The * symbol marks indicates the major section pages.

Page_id	**Non-gamified**	**% of total**	**Gamified (n=8)**	**% of total**

	(n=7)	actions		actions
busProblem	8	5.41	12	2.60
chairProblem	5	3.38	14	3.03
conceptMap	6	4.05	12	2.60
constraints	8	5.41	18	3.90
dashboard	N/A	N/A	59	N/A
decisionVariables	9	6.08	27	5.84
feasibleSolution	0	0.00	8	1.73
gamificationTutorial	N/A	N/A	45	N/A
introToOptimizationModeling*	26	17.57	54	11.69
linearProgramming	3	2.03	20	4.33
mathematicalModel	0	0.00	7	1.52
mathematicalProgramming	0	0.00	7	1.52
modelingLPProblem*	22	14.86	67	14.50
modelingLPProblem2*	15	10.14	64	13.85
nonNegConstraints	6	4.05	12	2.60
objectiveFunction	10	6.76	21	4.55
operationsResearch	1	0.68	12	2.60
optimalSolution	4	2.70	16	3.46
optimizationModeling	0	0.00	18	3.90
parameters	4	2.70	18	3.90
solution	0	0.00	7	1.52
whyOptimizationModeling*	21	14.19	48	10.39

Table 24 Number of times each page was opened

Table 25 illustrates the number of times participants attempted and managed to solve the various difficulties of the problem presented on the “Why optimization modeling” page.

	Non-gamified (n=7)		Gamified (n=8)	
	Total attempts	Correct attempts	Total attempts	Correct attempts
validateProdMixEasy	12	6	31	8
validateProdMixMed	2	2	10	5
validateProdMixHard	5	0	69	0
validateProdMixVeryHard	3	0	3	0

Table 25 Number of total and correct attempts to solve the problem

Table 26 illustrates the number of times participants attempted and solved the quizzes corresponding to the problems on pages “Modeling LP problem 1” and “Modeling LP problem 2”. The table also shows the number of times that the participants looked at the suggested answer.

	Non-gamified (n=7)			Gamified (n=8)		
	Total attempts	Correct attempts	Showed answer	Total attempts	Correct attempts	Showed answer
constraintsLP1	2	0	3	16	7	12

constraintsLP2	2	0	2	12	7	12
decisionVariablesLP1	4	2	7	18	8	17
decisionVariablesLP2	6	2	7	17	9	17
nonNegConstraintsLP1	2	0	3	13	5	15
nonNegConstraintsLP2	1	0	2	20	5	15
objectiveFunctionLP1	8	0	7	21	7	19
objectiveFunctionLP2	2	0	4	15	6	9
parametersLP1	0	0	2	12	7	10
parametersLP2	1	0	1	8	7	6

Table 26 Metrics concerning the pages Modeling LP problem 1&2

UserId	11	13	14	15	16	17	18	19	31	33	34	36
EXP earned	520	540	340	255	540	280	540	500	250	190	130	100
Commentator	x	x			x	x		x				
Evangelist	x	x			x			x				
Feature explorer		x	x		x	x	x	x		x	x	
Academic		x			x		x					
Supporter	x	x	x	x	x		x	x		x	x	
Student	x		x				x	x				
Enlightened			x						x	x	x	
Rising star		x										x
Champion		x			x		x					
Quiz whiz	x	x			x		x	x				
Helping hand	x	x			x	x	x	x				
Math lover												

Table 27 illustrates the achievements of the participants from the treatment group.

	Onsite								**Online**			
UserId	11	13	14	15	16	17	18	19	31	33	34	36
EXP earned	520	540	340	255	540	280	540	500	250	190	130	100
Commentator	x	x			x	x		x				
Evangelist	x	x			x			x				
Feature explorer		x	x		x	x	x	x		x	x	
Academic		x			x		x					
Supporter	x	x	x	x	x		x	x		x	x	
Student	x		x				x	x				
Enlightened			x						x	x	x	
Rising star		x										x

Champion		x			x		x					
Quiz whiz	x	x			x		x	x				
Helping hand	x	x			x	x	x	x				
Math lover												

Table 27 Achievements of all the participants to the gamified tutorial experiment

5.3 Data analysis

In this section we will proceed with a qualitative analysis of the data presented in the previous section. We will start by analyzing the data provided by the experiment without tracking data, we will continue with the analysis of the data provided by the experiment with tracking data and we will conclude with Section 5.3.3 that sums up all the observations we made.

5.3.1 Experiment without tracking data

There are a few interesting observations to be made related to the satisfaction with the tutorial system:

- There is a notable difference between the evaluation of the statement “it makes completing tasks fun” (non-gamified: $\bar{x}$=2.17, σ=1.17; gamified: $\bar{x}$=3.43, σ=0.98). We were expecting to see a difference here. It is unexpected that the participants rate both versions of the tutorial similarly in regard to “It is an engaging experience” (non-gamified: $\bar{x}$=3.50, σ=1.52; gamified: $\bar{x}$=3.29, σ=1.11). We were under the assumption that if a participant is having fun, he/she is also engaged with the application that is generating the fun. The section concerning the satisfaction with the tutorial system was adapted from the work of (Dong et al., 2012). Under these circumstances we have to wonder what construct validity tests did the authors ran when designing the questions of the original questionnaire. Furthermore, we believe that applying gamification to the context of teaching OR is either not a straightforward enterprise or that gamification is not well suited at all for this context. There is also the possibility that the concept of engagement was not properly explained in the post-questionnaire.
- A difference in the results for the statement “I felt rushed when completing tasks” (non-gamified: $\bar{x}$=1.83, σ=0.98; gamified: $\bar{x}$=2.86, σ=1.07) draws the attention because there were no gamification mechanisms related to time constraints. We intentionally left them out because we wanted to let the learners spend as much time as they need in the tutorial. Looking at the time participants spent on the gamified tutorial (non-gamified: $\bar{x}$=57.25, σ=6.13; gamified: $\bar{x}$=29.75, σ=13.42), it seems that the gamified version changed the behavior of the users in such a way that the participants spent a lot less time in the tutorial overall.

Participants rated the gamified tutorial as a slightly more effective learning tool than the non-gamified tutorial (non-gamified: $\bar{x}$=3.17, σ=0.98; gamified: $\bar{x}$=3.71, σ=0.49). However, their scores for the OR test indicate the opposite: the participants that did the gamified tutorial had considerably lower scores than the others (non-gamified: $\bar{x}$=6, σ=0.81; gamified: $\bar{x}$=4.75, σ=1.70). This could also be related to the time that each group spent in the tutorial. It would seem that the gamified tutorial discouraged the participants from spending time in the tutorial. This is contrary to what we were expecting: we were expecting that the participants doing the gamified tutorial would stay with it longer.

Regarding the experience of the participants with games, we expected that those with more video game experience would be more engaged with the gamified tutorial and vice-versa. The little data from Table 20 does not show any trend in any direction.

Regarding the interpersonal competitiveness score, the treatment group had higher scores (non-gamified: $\bar{x}=13$, $\sigma=2.6$; gamified: $\bar{x}=21$, $\sigma=3.5$). Considering that the gamified tutorial included elements designed to trigger interpersonal competitiveness, we were expecting to see these scores manifesting as more time spent in the tutorial. As mentioned before, the treatment group spent less time in the tutorial, so this is another unexpected discovery. We can only speculate that the competition elements present in the gamified tutorial were not relevant enough or placed in places that would not entice competition.

5.3.2 Experiment with tracking data available

Before going further in this section we would like to restate that there is a considerable difference between the participant groups that are involved in the comparison: the gamified version was delivered onsite and this meant that the participants had to work with the tutorial for at least 50 minutes. The non gamified version was delivered online and we had no direct control over how long the participants spent working with the tutorial. We asked them to spend at least 50 minutes, but the collected data (gamified tutorial time: $\bar{x}=32$, $\sigma=21$) shows that that did not happen. This is further discussed in Chapter 6.2.2.

There are a few interesting observations to be made related to the satisfaction with the tutorial system:

- The treatment group found the gamified tutorial to be slightly more engaging (non-gamified: $\bar{x}=3.29$, $\sigma=0.95$; gamified: $\bar{x}=3.75$, $\sigma=1.04$) and slightly more fun (non-gamified: $\bar{x}=2.86$, $\sigma=0.69$; gamified: $\bar{x}=3.50$, $\sigma=1.07$). We were expecting to see that these concepts are dependent on each other, but we were also expecting to see a more notable difference between the control and treatment groups. This could indicate that the fun aspect of the gamified tutorial was insufficient. This seems like a plausible explanation, since the gamification guidelines do not emphasize this aspect enough and don't suggest any method of how one can add fun to a gamified application.
- We notice a difference in the results for the statement "I felt rushed when completing tasks" (non-gamified: $\bar{x}=2.14$, $\sigma=1.07$; gamified: $\bar{x}=3$, $\sigma=0.93$). There were no gamification mechanisms related to time constraints. We intentionally left them out because we wanted to let the learners spend as much time as they need in the tutorial. Looking at the time participants spent on the gamified tutorial (non-gamified: $\bar{x}=35.86$, $\sigma=19.37$; gamified: $\bar{x}=58.25$, $\sigma=4.20$), it would seem that the gamified version generates this feeling despite the fact that the treatment participants spent more time in the tutorial. This difference might have appeared because the control participants were allowed to participate at the experiment at their own leisure, while the treatment group was simultaneously in the same room and the time limits were clearly stated at the beginning and repeated two times during the experiment (for the exact details see Section 5.1.1).

Participants rated the gamified tutorial as a slightly less effective learning tool than the non-gamified tutorial (non-gamified: $\bar{x}=3.43$, $\sigma=1.51$; gamified: $\bar{x}=2.75$, $\sigma=1.28$) and this is also reflected in their scores (non-gamified: $\bar{x}=6.14$, $\sigma=3.85$; gamified: $\bar{x}=2.87$, $\sigma=1.88$). Even with spending less time in the tutorial, the control group outperformed the treatment group (non-gamified: $\bar{x}=35.86$, $\sigma=19.37$; gamified: $\bar{x}=58.25$, $\sigma=4.20$). It appears that, for the treatment group, the tutorial became more geared towards the entertainment side, and it lost its educational qualities.

Regarding the experience of the participants with games, we expected that those with more video game experience would be more engaged with the gamified tutorial and vice-versa. The data from Table 22 does not show any trend in any direction.

Regarding the interpersonal competitiveness score, both groups had similar scores (non-gamified: $\bar{x}$=16.86, σ=6.91; gamified: $\bar{x}$=15.12, σ=4.70). Considering that the gamified tutorial included elements designed to trigger interpersonal competitiveness, we were expecting to see these scores manifesting as more time spent in the tutorial. However, due to the difference in the delivery of the experiment to the two groups, we cannot conclude if the difference in time spent in the tutorial can be explained by this score.

By looking at the times participants have spent in each major section of the tutorial (Table 23), we notice that, in general, the treatment group spent relatively equal amounts of time on each section (P1: $\bar{x}$=13.11, σ=4.48; P2: $\bar{x}$=13.42, σ=6.82; P3: $\bar{x}$=12.27, σ=4.23; P4: $\bar{x}$=13.59, σ=7.16), whereas the members of the control group were less interested in the second and fourth section (P1: $\bar{x}$=16.11, σ=17.57; P2: $\bar{x}$=5.74, σ=4.55; P3: $\bar{x}$=8.13, σ=6.14; P4: $\bar{x}$=5.47, σ=5.34). This could be explained by the fact that the gamified tutorial had similar incentives for each section of the tutorial.

5.3.2.1 Gamification mechanisms related to the opening of all tutorial pages

There were two mechanisms encouraging the participants to open all the pages of the tutorial:

- The “Academic” badge would be awarded when the participant would open all the 25 pages of the tutorial.
- Earning 10 EXP when opening a page for the first time.

The data available in Table 24 allows us to see which the most popular pages of the tutorial were. Considering that the control group spent little time in the tutorial, we can conclude that the design of the non-gamified tutorial favors the pages (we considered 3% as the threshold): busProblem, chairProblem, conceptMap, constraints, decisionVariables, objectiveFunction, nonNegConstraints. These pages contain very important content for the tutorial.

We notice that the treatment group had a tendency to open more pages of the tutorial (Table 24). This is in line with the intent of the gamification mechanisms. However, this came at the cost of opening the very important content pages less. This observation adds to the list of reasons for which the treatment group had considerably lower scores in the post experiment OR test. Another note that we need to make is that opening a page does not necessarily mean that a participant also absorbed the information in it.

The gamification mechanisms asked the participant just to open the page. We speculate that if the mechanism would have also asked the user to stay for a specific amount of time on that page (thus ecnouraging the participant to read the information on the page), the OR test score would have increased.

5.3.2.2 Gamification mechanisms related to solving the quizzes from the tutorial

There were two mechanisms encouraging the participants to open all the pages of the tutorial:

- The “Quiz whiz” badge would be awarded when the participant would correctly solve all the ten quizzes from the tutorial

- Earning 5 EXP for the first attempt to solve the quiz and an extra 25 EXP when the solution would be correct.

We note very few attempts to solve the quizzes by the control group. Because of the differences between the control group and the treatment group we can't proceed with further comparison. But we see that more than half of the treatment group got the right answers to all the quizzes. Considering the results in the final OR test, looking at how many times participants clicked the show answer and from the analysis of the chatroom messages, there is some evidence that the users took advantage of the fact that they could copy the answer and submit the response and they could still get the points.

5.3.2.3 Chat room feature

The chat room feature was created with the intent of allowing the participants to interact with each other to clarify concepts. This behavior was encouraged by:

- Two badges (Commentator, Evangelist) that would be awarded when the user would comment (5 words minimum) for 5, respectively 25 times.
- A Helping hand badge that the participant would earn if he would be voted by other 2 participants as helpful.

An analysis of the chat room logs shows that this feature failed to achieve its intended purpose: the participants did not help each other to understand the content of the tutorial better. Participants:

- shared tips about how to "game" the application: "I found we can skip the questions, still gain EXPs", "copy and paste the answers in the quiz, 300 EXP get""
- asked other to vote them so they could earn the "helping hand" badge,: "can somebody rate me helpful", "ratee meee", "rate me ,meeeeme <3"
- Wrote irrelevant comments: "t(T__Tt) no math plx", "Do they also make low fat chips?"
- commeted about their achievements: "awww yeah, got my first badge!!", "290 exp woo", "I got my student badge and proud of it"
- spammed the chat room to earn the Commentator and Evangelist badges: "This is a comment for the special badge exceeding 5 words", "s p a m m"

The content of the chat room at the end of the experiment is available on demand.

5.3.2.4 General gamification analysis

By analyzing Table 27 we can notice that five out of the eight onsite participants (user11, user13, user16, user18, user19) were engaged in earning points and badges (each of them had at least six badges). However, there are no notable differences between their results to the OR test and the results of participants that were not so engaged. We do not find support for any statement in the lines of "because the participants started chasing badges and points, they performed badly in the post experiment OR test". It seems that the low performance of the control group in the OR test is not determined by any particular game elements, but by the gamified experience taken as a whole.

We notice that some badges were not earned by all the participants:

- Supporter, Student, Enlightened, Rising star. Their intent was to encourage the participant to spend enough continuous time (at least four minutes) on the main sections of the tutorial. Looking at how many times were opened during the experiment, it would seem participants were just rushing through them and not spending enough time to absorb the information on those pages properly. It seems that most of them failed to achieve their purpose.

- Math lover. Its intent was to encourage participants to solve all the difficulties of the problem on the "Why optimization modeling page". This badge was not earned by anyone, the challenge it posed was too big.
- Commentator, Evangelist. Their intent was to encourage participants to comment in the chat room. As mentioned in Section 5.3.2.3, the chat room failed to achieve its purpose.

We conclude from the above that the badge system should have been designed better.

By looking at the achievements of the participants to the online experiment it seems that they were not interested at all in the gamification features of the tutorial, their achievement level is similar to that of the three participants from the onsite groups that were also not interested in the gamification achievements. It is possible that these participants fall in the category of people that are not interested in gamification, a category noticed by other authors as well (Farzan & DiMicco, 2008), (Fitz-Walter & Wyeth, 2013). This could also be a reason for which they spend little time in the tutorial.

5.3.3 Data analysis summary

It is not clear whether the gamified tutorial that we constructed in this project is more engaging and more fun than the non-gamified tutorial. The promise of gamification is that it increases user engagement. From this we were hoping to see an increase in the time spent in the tutorial and the OR scores for the treatment group. During the design of the gamified tutorial, we used a framework for applying gamification and the guidelines that other authors have found useful in the learning/education context. With these arguments in mind, it is hard to state that our approach in designing and implementing the gamified tutorial was so flawed that there would be no difference between the gamified tutorial and the non-gamified tutorial.

After-the-fact, we realized that we should have paid more attention to gamification concepts like: the fun aspect, the badges design, the elements related to competition, user testing. We found these elements mentioned in the literature, but their importance was definitely not emphasized enough. Therefore, we come to the conclusion that applying gamification to the context of teaching OR is either not a straightforward enterprise (something that the current gamification literature has discussed properly) or that gamification is not well suited at all for this context.

A very clear discovery is that the participants doing the gamified tutorial had considerably lower scores than those that did the non-gamified tutorial. The time spent in the tutorial did not have any effect on the score.

Another very clear discovery that this project has made is that the treatment group felt more rushed while doing tasks than the control group. This happened despite the fact that there were no gamification mechanisms related to time constraints and despite the fact that the treatment group spent more time overall in the tutorial than the control group.

In terms of the pre and post experiment questionnaires, although we adapted questionnaires that other authors have used, we found them to be slightly lacking. We did not find a link between game experience or competitiveness and the user experience with the tutorial. And we believe that the questions measuring the satisfaction with the tutorial system could be improved.

In a real-life scenario, gamification would be applied iteratively and would be monitored and improved as more information about the users and their behavior in the application becomes available. Usually, for scientific project similar to ours, redoing the experiment will be unfeasible and we have a couple of observations that will ensure a better chance for success. By looking at the impact that the gamification elements had on the behaviors of the participants, we note that, during gamification design, one should be fully aware of how the change in behavior that gamification brings affect the purpose of the underlying application (e.g. through gamification we encouraged participants to open more pages, but we did not take into consideration the fact that it will make the participants skip from page to page more and faster, thus decreasing quality of the information absorbed from those pages). Another thing to keep in mind during gamification design is not to leave any loopholes in the design of the application or in the gamification design. If they are obvious, participants will take advantage of them immediately.

6 Discussion and conclusion

This chapter presents a summary of the findings of this research project (Section 6.1), followed description of its limitations (Section 6.2) and ending with proposed directions for further research (Section 6.3).

6.1 Research results

In this section, we discuss and present the answers to the sub and main research questions.

6.1.1 RSQ1: What are the intended learning outcomes of operations research for potential new adopters?

When we started this research project we had little knowledge of the operations research field and, from preliminary research, it seemed that this domain is still missing clear definitions of the concepts it covers. Our task was made difficult by the fact that operations research is a wide and complex field using advanced mathematics. Keeping the target audience (managers) of the OR tutorial, we had to keep a balance between presenting business value and technical concepts.

To answer this research sub question, we assembled an expert panel (four industry experts) that helped us by suggesting sources of inspiration for the tutorial and, where it was the case, helping us decide whether a concept was fit for the purpose of the tutorial. We answered this question by identifying four intended learning outcomes that are relevant for a manager making first contact with operations research, thereby excluding advanced concepts that are more useful for OR modelers.

6.1.2 RSQ2: What are the guidelines for creating a digital interactive tutorial?

Gamification has been applied mostly in digital environments. Hence, we decided that the OR tutorial should also be a digital tutorial. Due to our lack of experience in creating digital interactive tutorials, we decided to find and follow the existing best practices available in scientific literature.

To answer this research sub question, we conducted a literature review with the objective of identifying the guidelines of designing digital interactive tutorials. We used the snowball method to find the relevant references and found eighteen guidelines. Having used them in our implementation of the artifact, we consider the guidelines we found as a very good place to start from in research projects that are concerned with designing digital interactive tutorials.

6.1.3 RSQ3: What aspects of gamification can be applied on top of a digital interactive tutorial to increase its effectiveness?

Gamification is a relatively new research field and there is much hype around it. There are various persons in the industry giving recommendations on how to apply gamification to various contexts. We needed to identify the gamification elements that others have applied successfully (and validated) to the learning/education context.

To answer this research sub question, we conducted a literature review with the objective of identifying the guidelines for applying gamification in learning and education contexts. We used the snowball method to find the relevant references and found twenty guidelines related to concrete game elements and twelve general gamification guidelines. These guidelines were found (through scientific experiments) beneficial by other authors. However, since the evaluation of the gamified tutorial showed negative results regarding the effectiveness of learning, we cannot make any statement about the value that they add to a gamified tutorial.

6.1.4 RSQ4: How is a 'gamified digital interactive tutorial for operations research' designed?

To create the envisioned artifact, we need to put everything into a coherent form. In order to ensure traceability and reproducibility of the design process, we documented all of our choices. Additionally, any changes that were induced in the design by obstacles encountered during implementation were also documented. We also documented how each of the elements of the guidelines and the intended learning objectives related to the design decisions.

To answer this research sub question, we first designed and implemented the 'digital interactive tutorial for operations research', keeping in mind the intended learning outcomes and the guidelines for creating a digital interactive tutorial. Having this baseline, we proceeded to design and to implement of the 'gamified digital interactive tutorial for operations research', keeping in mind the guidelines for applying gamification in learning and education contexts.

6.1.5 RSQ5: What are the benefits brought by gamification to a tutorial for operations research?

The next step in our research was to evaluate the gamified artifact (i.e. compare it with the baseline). The evaluation consisted of administering a pre questionnaire, allowing the participants to use of the tutorials (generating runtime data) and administering a post questionnaire.

To answer this research sub question, we analyzed all the data we gathered from the participants. Unfortunately, we found that gamification had a negative effect on the learning of OR: participants doing the gamified tutorial had considerably lower scores than those who did the non-gamified tutorial. Additionally, the gamified tutorial was not more engaging and more fun than the non-gamified tutorial (i.e., they seem to be equally engaging and fun) and the gamified tutorial users felt more rushed than the non-gamified tutorial users.

6.1.6 Main research question: What role can gamification play in teaching operations research to potential new adopters?

Summing up the answers to the research sub-questions will help us answer the main research question.

In this research we followed a rigorous method that guided us in the application of gamification to the context of teaching operations research. We compared this gamified application to its non-gamified version and observed, to our surprise, that gamification affected negatively the learning experience. It seems that applying gamification to the context of teaching OR is either not a straightforward enterprise or that gamification is not well suited at all for this context.

6.2 Limitations

The main author has the role of practitioner-researcher in a company involved in developing OR solutions. His role could be influenced by assumptions and preconceptions about the industry and the company. The issue related to the industry is not a concern since the author has a junior position and is not involved in the OR department of the company.

This research project has produced a gamified tutorial aimed at the OR domain. While the lessons learned might be applicable to other domains, there is no guarantee that they will have the same effect.

In other words, the results of this research are not generalizable. On the other hand, considering the results of the research and considering that we followed best practices that others have found to be useful in similar contexts, it seems that there is no guarantee that applying best practices of gamification to a new context will help in that context.

In the post experiment questionnaire we used the framework developed by Kirkpatrick (1994) for the evaluation of training results. The framework measures the results at four stages: reaction, learning, behavior, results. In this research we did not measure the behavior and results stages, since these required that the learner has a role in an organization and our context did not provide that.

During the literature review for digital interactive tutorial guidelines we decided to de-scope the 'human-computer interaction' (HCI) domain. This domain is very wide and outside the knowledge area of all the authors of this research and, considering the time constraints of this project, we decided not to follow it. The reason for which we see this as a limitation is that gamification is part of this domain. The same arguments apply to the domain of 'exploratory learning'. There are video games that are based on exploration to onboard users and gamification also makes use of it in the same way. Most of the user experience issues and unnatural presentation of the learning content was ironed out during user testing. Besides, since these guidelines have been used to create the "control" tutorial, they will also be missing from the "treatment" tutorial. For that reason, excluding them does not affect the results in any way.

6.2.1 Gamification limitations

Since gamification is a recent domain, we admit the possibility that scientific literature might not contain all the knowledge that is, at the moment, implicit. However, we decided not to assemble an expert panel that would review our design decisions as we did not believe that industry experts could provide us with objective, empirically-demonstrated-to-work advice.

Another issue encountered was the complexity of the gamification domain. Delving into the topic more, we discovered that it depends on many concepts that are research topics in their own right. We will treat these topics in Section 6.3.

During the design of the gamified tutorial, we made assumptions about the behavior/characteristics of the target group that we inferred directly from the characteristics of early adopters. It might not be fully the case that our target group (students) has these characteristics.

Another limitation related to the design of the gamified application is that we did not include all gamification guidelines we found in the gamified app. The time constraints of the project stopped us from including the following gamification guidelines:

- Design the application for multiple users
- Include a narrative
- Offer progressive disclosure
- Include character level ups
- Offer avatars
- Include triggers
- Allow the possibility to customize the application

6.2.2 Experiment

The original objective of this research project was to create an OR tutorial for decision makers. From the beginning, we recognized that we will not have access to decision makers. As a result, we had to design the artifact having students (that are on a career track that will most likely lead them to a decision making role) as a target group. Even with this decision, we have no guarantee that the students will understand the value of OR in the same way that seasoned decision makers would. Most of the users of OR are involved in the industry.

Not being able to gather enough participants to attend the onsite experiment, we were forced to deliver the experiment for the non-gamified tutorial online. This lead to not having enough control over the sample (gamified versus non-gamified, when we tracked data). For example, the onsite group was assisted by two experimenters and had a fixed time limit (noted twice as the experiment progressed) to do the experiment. This poses a serious threat to the validity of the collected data, even though we tried to minimize this risk during the data analysis phase.

Another limitation of the experiment is that we did not manage to find enough participants who would permit us to achieve statistical significance, hence we only created hypotheses about our observations. We leave it to others to reuse our research in order to validate our hypotheses.

6.3 Further research

In this section we propose four general research directions.

By looking at the results of this research we posit that gamification is not as easy to apply to a new context as it seems. We suggest that, before accepting that gamification can be successfully applied in other gamification/learning contexts, more research is conducted in these contexts, for various disciplines.

In this research project we created a gamified tutorial for teaching OR and it still needs an evaluation with a statistically significant sample in a more controlled environment. Future research could be concerned with redoing the experiment with: a statistically significant sample, keeping the differences between the control group and the treatment group to a minimum (i.e., they should be as homogenous as possible, with no OR experience; the only difference being the treatment condition) and making sure that the concepts that the questionnaires measure are proper (e.g., whether post questionnaire measures correctly the construct "satisfaction with the system"). We encourage this endeavor by offering the source code of the application freely[10].

During the literature review regarding the digital interactive tutorial guidelines, we intentionally excluded two domains that had the potential of offering further guidelines:

- User interface design, user experience design and, in general, human computer interaction
- Exploratory learning, discovery based learning, discovery learning, guided discovery or guided exploration

The main reason for considering these domains not in scope is that this research project did not aim to create the best possible user interface nor the best possible learning experience.

[10] https://github.com/treaz/treaz.github.io

During the literature review regarding the guidelines of gamification in the learning/education context, we intentionally excluded domains that had the potential of offering further guidelines:

- User interface design, user experience design and, in general, human computer interaction (e.g. the directions given by (Montola & Nummenmaa, 2009)). Since this domain is also excluded from the digital interactive tutorial guidelines, is makes sense to also exclude it from the gamification guidelines.
- Video game design. This is a domain with a lot research and from which gamification draws its main concepts. Gamification is a new domain and it is not clear yet whether or not is should be studied as a part of video game design or as a standalone domain. We have conducted this research starting from the assumption that gamification is a standalone domain. Therefore, we steered away from the video game design domain. Deterding (2011) makes a clear point that gamification cannot be decoupled from the game design discipline. Other gamification industry thought leaders share his view.
- Incentives, persuasion, motivation mechanisms, self-determination theory etc. We recognize these concepts as relevant to gamification, however, they have been treated by authors through standalone studies (i.e. making this direction out of the scope of the current research project). Furthermore, they fit the domain of social sciences, not the domain of natural sciences (which is the background of the authors of this research project).
- The concept of flow. Following this direction would require extensive user testing and fine tuning of the elements of the gamified application. Although this would be an interesting endeavor, it ~~is~~ requires going into detail at a level that this project does not allow.
- Achievement systems. There are some scientific studies treating this domain in detail. If we were to design a proper (validated with users) achievement system, we would be exceeding the time allocated for this project. According to (Galli & Fraternali, 2014), point systems, badges and leaderboards are also considered achievement systems. Future research could investigate the importance of achievement systems in gamified applications.

Further research could be concerned with clarifying the relation between gamification and these domains.

Finally, an interesting research direction is the behavior that players exhibit towards gamification. When it comes to video games, different game types demand different kinds of interactions from players (without going into details, casual games allow the player to be involved with the game for short periods of time, while massively multiplayer online role-playing games require the player's attention for longer spans of time). Would these differences also apply to gamification? Would the behavior that gamification creates fall under one of the existing video game player behaviors?

References

Baker, E., & Mayer, R. (1999). Computer-based assessment of problem solving. *Computers in Human Behavior*, *15*(3), 269–282.

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving participation and learning with gamification. *Proceedings of the First International Conference on Gameful Design, Research, and Applications - Gamification '13*, 10–17. doi:10.1145/2583008.2583010

Bartle, R. (2012). *Player type theory - uses and abuses*. Retrieved November 02, 2014, from http://mud.co.uk/richard/Hamburg.pdf

Barwood, H., & Falstein, N. (2006). *THE 400 PROJECT*. Retrieved May 28, 2014, from http://www.finitearts.com/Pages/400page.html

Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University* (pp. 113–131). McGraw-Hill International.

Bisschop, J. (2006). *AIMMS-optimization modeling*. Lulu.com.

Bradley, S., Hax, A., & Magnanti, T. (1977). *Applied mathematical programming*. Addison Wesley.

Carlson, B., Chen, Y., Hong, M., & Jones, R. (2012). MISO Unlocks Billions in Savings Through the Application of Operations Research for Energy and Ancillary Services Markets. *Interfaces*, *42*(1), 58–73.

Carroll, J. M. (1990a). An overview of minimalist instruction. In *System Sciences, 1990., Proceedings of the Twenty-Third Annual Hawaii International Conference on* (pp. 210–219). IEEE.

Carroll, J. M. (1990b). *The Nurnberg funnel: designing minimalist instruction for practical computer skill*. MIT press Cambridge, MA.

Carroll, J. M. (2009). *Human-Computer Interaction*. *Encyclopedia of Cognitive Science*. Wiley Online Library.

Catrambone, R., & Carroll, J. (1987). *Learning a word processing system with training wheels and guided exploration* (pp. 169–174). ACM.

Charles, T., Bustard, D., & Black, M. (2011). Experiences of Promoting Student Engagement Through Game-Enhanced Learning. doi:10.1007/978-1-4471-2161-9

Chen, H., Chiang, R., & Storey, V. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*(4), 1165–1188.

Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience*. New York: HarperPerennial.

Cunningham, W. (1984). OR Forum—Of Radar and Operations Research: An Appreciation of AP Rowe (1898–1976). *Operations Research*, *32*(4), 958–967.

Davenport, T. (2006). Competing on analytics. *Harvard Business Review*, *84*(1), 98.

De Schutter, B., & Vanden Abeele, V. (2014). Gradequest—Evaluating the impact of using game design techniques in an undergraduate course. In *9th International Conference on the Foundations of Digital Games*. Retrieved from http://fdg2014.org/papers/fdg2014_paper_07.pdf

Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359–363. doi:10.1016/j.dss.2012.05.044

Denny, P. (2013). The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 763–772). acm.

Deterding, S. (2011). *Meaningful play*. Retrieved November 03, 2014, from http://codingconduct.cc/Meaningful-Play

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). ACM.

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, *63*, 380–392. doi:10.1016/j.compedu.2012.12.020

Dong, T., Dontcheva, M., & Joseph, D. (2012). Discovery-based games for learning software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2083–2086). ACM.

Elliott, T. (2012). *2012: The Year Analytics Means Business*. Retrieved September 13, 2014, from http://smartdatacollective.com/timoelliott/45868/2012-year-analytics-means-business

Ergun, Ö., Keskinocak, P., & Swann, J. (2011). Introduction to the special issue on humanitarian applications: Doing good with good OR. *Interfaces*, *41*(3), 215–222. doi:10.1287/inte.1110.0578

Farzan, R., & DiMicco, J. (2008). When the experiment is over: Deploying an incentive system to all the users. In *Symposium on Persuasive Technology*.

Fenn, J. (2013). *Toolkit: My Hype Cycle*.

Fitz-Walter, Z., & Wyeth, P. (2013). Driven to drive: designing gamification for a learner logbook smartphone application, 42–49.

Fleuren, H., & Goossens, C. (2013). Supply Chain–Wide Optimization at TNT Express. *Interfaces*, *43*(1), 5–20.

Foster, J., & Sheridan, P. (2012). Gamification as a Strategy for Promoting Deeper Investigation in a Reverse Engineering Activity. In *American Society for Engineering Education*. American Society for Engineering Education.

Galli, L., & Fraternali, P. (2014). Achievement Systems Explained, 25–50. doi:10.1007/978-981-4560-26-9

Gravetter, F., & Wallnau, L. (2013). *Essentials of statistics for the behavioral sciences*. Cengage Learning.

Griffin-Pierson, S. (1990). The Competitiveness Questionnaire: A Measure of Two Components of Competitiveness. *Measurement and Evaluation in Counseling and Development*, *23*(3), 108–15.

Grossman, T., & Fitzmaurice, G. (2010). ToolClips: an investigation of contextual video assistance for functionality understanding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1515–1524). ACM.

Hakulinen, L., Auvinen, T., & Korhonen, A. (2013). Empirical Study on the Effect of Achievement Badges in TRAKLA2 Online Learning Environment. *2013 Learning and Teaching in Computing and Engineering*, 47–54. doi:10.1109/LaTiCE.2013.34

Hamari, J., & Eranti, V. (2011). Framework for designing and evaluating game achievements. *Proc. DiGRA 2011: Think Design Play*, *115*(115), 122–134.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. *2014 47th Hawaii International Conference on System Sciences*, 3025–3034. doi:10.1109/HICSS.2014.377

Harris, D. (2012). *IDC: Analytics a $51B business by 2016 thanks to big data*. Retrieved September 13, 2014, from https://gigaom.com/2012/07/11/idc-analytics-a-51b-business-by-2016-thanks-to-big-data/

Hutton, D. W., Brandeau, M. L., & So, S. K. (2011). Doing Good with Good OR: Supporting Cost-effective Hepatitis B Interventions. *Interfaces*, *41*(3), 289–300. doi:10.1287/inte.1100.0511

Jalali, S., & Wohlin, C. (2012). Systematic literature studies: database searches vs. backward snowballing. In *Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement* (pp. 29–38). ACM.

Jensen, Paul A and Bard, J. F. (2003). *Operations research models and methods*. John Wiley & Sons Incorporated.

Kalakota, R. (2011). *Gartner says – Business Analytics a $14.1 Bln market*. Retrieved September 13, 2014, from http://practicalanalytics.wordpress.com/2011/04/24/gartner-says-bi-and-analytics-a-10-5-bln-market/

Kalakota, R. (2012). *IBM CIO Study: BI and Analytics are #1 Priority for 2012/2013*. Retrieved September 13, 2014, from http://practicalanalytics.wordpress.com/2011/11/02/ibm-cio-study-bi-and-analytics-are-1-priority-for-2012/

Kirkpatrick, D. L. (1994). *Evaluating training programs: the four levels*. San Francisco: Berrett-Koehler.

Kowalczyk, M., Buxmann, P., & Besier, J. (2013). Investigating Business Intelligence and Analytics from a Decision Process Perspective: A Structured Literature Review. In *ECIS 2013 Completed Research*. AISEL.

Kroon, L., Huisman, D., & Abbink, E. (2009). The new Dutch timetable: The OR revolution. *Interfaces*, *39*(1), 6–17. doi:10.1287/inte.1080.0409

Kulik, C.-L. C., & Kulik, J. a. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, *7*(1-2), 75–94. doi:10.1016/0747-5632(91)90030-5

Lazzaro, N. (2004). Why we play games: Four keys to more emotion without story, 1–8. Retrieved from http://www.citeulike.org/group/596/article/436449

Li, W., Grossman, T., & Fitzmaurice, G. (2012). Gamicad: a gamified tutorial system for first time autocad users. In *Proceedings of the 25th annual ACM symposium on User interface software and technology* (pp. 103–112). ACM.

Linden, A. (2014). *Hype Cycle for Advanced Analytics and Data Science , 2014*.

Mager, R. (1997). Preparing Instructional Objectives: A critical tool in the development of effective instruction 3rd edition. *The Center for Effective Performance, Inc*.

Meyer, E., Crosby, M., & Ogawa, M. (n.d.). Augmenting Large-enrollment Instruction With Game-based Mechanics.

Montola, M., & Nummenmaa, T. (2009). Applying game achievement systems to enhance user experience in a photo sharing service. In *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era* (Vol. 4, pp. 94–97). ACM.

Nah, F. F.-H., Telaprolu, V. R., Rallapalli, S., & Venkata, P. R. (2013). Gamification of education using computer games. In *Human Interface and the Management of Information. Information and Interaction for Learning, Culture, Collaboration and Business,* (pp. 99–107). Springer.

Nah, F. F.-H., Zeng, Q., Telaprolu, V. R., & Ayyappa, Abhishek Padmanabhuni Eschenbrenner, B. (2014). Gamification of Education: A Review of Literature. In *HCI in Business: First ...* (pp. 401–409). Springer. Retrieved from http://link.springer.com/content/pdf/10.1007/978-3-319-07293-7_39.pdf

O'Donovan, S., Gain, J., & Marais, P. (2013). A case study in the gamification of a university-level games development course. *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference on - SAICSIT '13*, 242. doi:10.1145/2513456.2513469

O'Neil, H., Wainess, R., & Baker, E. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Cirriculum Journal*, *16*(4), 455–474. doi:10.1080/09585170500384529

Prensky, M. (2005). Computer games and learning: Digital game-based learning. *Handbook of Computer Game Studies*, *18*, 97–122.

Rieman, J. (1996). A field study of exploratory learning strategies. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *3*(3), 189–218. doi:10.1145/234526.234527

Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

Sashihara, S. (2011). *The Optimization Edge: Reinventing Decision Making to Maximize All Your Company's Assets*. McGraw-Hill.

Sharda, R., Asamoah, D., & Ponna, N. (2013). Business analytics: Research and teaching perspectives. *ITI*, 19–28. doi:10.2498/iti.2013.0589

Shi, P. (2014). *Guiding School Choice Reform through Novel Applications of Operations Research* (pp. 1–25). Massachusetts. Retrieved from http://www.mit.edu/~pengshi/papers/guiding-reform.pdf

Shrader, C. (2008). *History of Operations Research in the United States Army, V. 2: 1961-1973*.

Singer, L. (2013). *Improving the adoption of software engineering practices through persuasive interventions*. Gottfried Wilhelm Leibniz Universitat Hannover.

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity & femininity: Their psychological dimensions, correlates, and antecedents*. University of Texas Press Austin.

Stott, A., & Neustaedter, C. (2013). *Analysis of Gamification in Education*. *Human Interface and the …* (pp. 1–8). Surrey, BC, Canada.

Swoyer, S. (n.d.). *Demand for Analytics Keeps Growing*. *2012*. Retrieved September 13, 2014, from http://tdwi.org/articles/2012/09/18/Demand Analytics Grows.aspx

Taha, H. (2007). *Operations research: an introduction*. Pearson/Prentice Hall.

Trkman, P., McCormack, K., de Oliveira, M. P. V., & Ladeira, M. B. (2010). The impact of business analytics on supply chain performance. *Decision Support Systems*, *49*(3), 318–327. doi:10.1016/j.dss.2010.03.007

Vanderbei, R. J. (2001). *Linear programming Foundations and extensions*. *Foundations and extensions, International Series in Operations Research & Management Science* (4th ed., Vol. 37). Springer.

Vanderlinden, G. (1988). Designing tutorials that help users learn through exploration. In *Professional Communication Conference, 1988. IPCC'88 Conference Record. On the Edge: A Pacific Rim Conference on Professional Technical Communication., International* (pp. 295–299). Ieee. doi:10.1109/IPCC.1988.24054

Von Alan, R. H., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, *28*(1), 75–105.

Webster, J., & Watson, T. (2002). R., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), 3.

Werbach, K., & Hunter, D. (2012). *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Digital Press.

7 Appendix

7.1 OR terminology

Concept	Definition
Operations	The activities carried out in an organization related to attaining its goals and objectives.
Research	The process of observation and testing characterized by the scientific method. The steps of the process include observing the situation and formulating a problem statement, constructing a mathematical model, hypothesizing that the model represents the important aspects of the situation, and validating the model through experimentation
Model	An abstract representation of reality. As used here, a representation of a decision problem related to the operations of the organization. The model is usually presented in mathematical terms and includes a statement of the assumptions used in the functional relationships. Models can also be physical, narrative, or a set of rules embodied in a computer program.
Optimal solution	A solution to the model that optimizes (maximizes or minimizes) some objective measure of merit over all feasible solutions -- the best solution amongst all alternatives given the organizational, physical and technological constraints.
Linear programming	Linear programming is a widely used model type that can solve decision problems with many thousands of variables. Generally, the feasible values of the decisions are delimited by a set of constraints that are described by mathematical functions of the decision variables. The feasible decisions are compared using an objective function that depends on the decision variables. For a linear program the objective function and constraints are required to be linearly related to the variables of the problem.
Objective function	The objective function evaluates some quantitative criterion of immediate importance such as cost, profit, utility, or yield. The general linear objective function can be written as $z = c_1x_1 + c_2x_2 + \ldots + c_nx_n = \sum_{j=1}^{n} c_j x_j$ Here c_j is the coefficient of the *j*th decision variable. The criterion selected can be either maximized or minimized.
Decision variables	Decision variables describe the quantities that the decision makers would like to determine. They are the unknowns of a mathematical programming model. Typically we will determine their optimum values with an optimization method. In a general model, decision variables are given algebraic designations such as x1, x2, …, xn. The number of decision variables is n, and xj is the name of the jth variable. An assignment of values to all variables in a problem is called a solution.
Constraints	A constraint is an inequality or equality defining limitations on decisions. Constraints arise from a variety of sources such as limited resources, contractual obligations, or physical laws. In general, an LP is said to have m linear constraints

	that can be stated as: $\sum_{j=1}^{n} a_{ij} x_j \begin{Bmatrix} \le \\ = \\ \ge \end{Bmatrix} b_i, \text{ for } i = 1 \ldots m$ One of the three relations shown in the large brackets must be chosen for each constraint. The number a_{ij} is called a "technological coefficient," and the number b_i is called the "right-hand side" value of the *i*th constraint. Strict inequalities (< and >) are not permitted. When formulating a model, it is good practice to give a name to each constraint that reflects its purpose.
Non-negativity restrictions	In most practical problems the variables are required to be nonnegative (e.g. production values); This special kind of constraint is called a non-negativity restriction. Sometimes variables are required to be non-positive or, in fact, may be unrestricted (allowing any real value).
Simple upper bound	Associated with each variable, x_j, may be a specified quantity, u_j, that limits its value from above; $x_j \le u_j, \text{for } j = 1 \ldots n$ When a simple upper is not specified for a variable, the variable is said to be unbounded from above.
Parameters	The collection of coefficients (cj, aij, bi, uj) for all values of the indices i and j are called the parameters of the model. For the model to be completely determined all parameter values must be known.

Table 28 OR terminology extracted from (Jensen, Paul A and Bard, 2003)*. 11 concepts extracted*

CONCEPT	DEFINITION
Model	A prototype of something that is real. Such a prototype can be concrete (a teddy bear) or abstract (information of how to produce different types of teddy bears in a factory)
Mathematical model	A description of some part of the real world expressed in the language of mathematics. In other words, an abstract model that describes in general mathematical terms the relations contained in an abstract model. E.g. the equations used to determine the production price of a teddy bear.
Optimization models	One class of mathematical models
Parameters	Symbols representing known data
Variables	Unknowns of a mathematical model.
Non-negativity constraints	Constraints defining the lower bounds of the variables
Solver	Mathematical algorithms that is used to obtain solutions of mathematical models
Linear programming Models	Mathematical models made up from linear equations and linear inequalities in one or more unknowns. These are characterized by the restriction of employing only “+” and “−” operations on the terms (where a term is defined as a coefficient times a variable) and no power terms.
Solution	A set of values for the variables, consistent with the linear inequalities and/or equations.

Graphical representation of a two variable LP	
Contour of the objective function	(inferred) Any line corresponding to a specific value of the objective function.
Feasible region	Region bounded by the lines corresponding to the constraints
Optimal corner solution	If a linear programming model has an optimal solution, the optimal solution is on a corner of the feasible region (intersection of two lines)

Table 29 OR concepts extracted from (Bisschop, 2006)*. 12 concepts extracted.*

CONCEPT	DEFINITION
Decision variables	Variables whose values are to be decided in some optimal fashion
Objective function	A linear function that consists of decision variables that needs to be maximized (e.g. profit) or minimized (e.g. raw materials costs)
Constraints	An equality or an inequality associated with some linear combination of the decision variables
Solution	A proposal of specific values for the decision variables
Feasible solution	A solution that satisfies all of the constraints
Optimal solution	A feasible solution that attains the desired maximum
Infeasible problem	A problem with no feasible solution
Unbounded problem	A problem that has feasible solutions with arbitrarily large objective Values

Table 30 OR terminology extracted from (Vanderbei, 2001)*. 8 concepts extracted*

CONCEPT	DEFINITION
Decision variables	Elements under the control of the decision-maker, and their values determine the solution of the model
Management science	A scientific approach to managerial decision making. It attempts to apply mathematical methods and the capabilities of modern computers to the difficult and unstructured problems confronting modern managers. Also known as: operations research, operational research, systems analysis, cost–benefit analysis, and cost-effectiveness analysis.
Mathematical programming	Branch of management science. Concerns the optimum allocation of limited resources among competing activities, under a set of constraints imposed by the nature of the problem being studied. In broad terms, mathematical programming can be defined as a mathematical representation aimed at programming or planning the best possible allocation of scarce resources.
Linear programming	Form of mathematical programming that uses linear functions exclusively
Objective function	The criterion the decision maker will use to evaluate alternative solutions to the problem
Constraints	Restrictions imposed upon the values of the decision variables by the characteristics of the problem under study
Model	Simplified representations of the real world
Analytical model	The problem represented completely in mathematical terms, normally by means of a criterion or objective, which we seek to maximize or minimize, subject to a set of mathematical constraints that portray the conditions under which the decisions have to be made. The model computes an

	optimal solution, that is, one that satisfies all the constraints and gives the best possible value of the objective function.
Parametric analysis	Allows to investigate the behavior of the solution as other parameters of the problem (for example, minimum allowed silicon content) are varied.
Shadow Price	The shadow price on a particular constraint represents the change in the value of the objective function per unit increase in the right-hand-side value of that constraint.
Graphical representation of a two variable LP	
Feasible region	The area that contains all the feasible solutions
Optimal corner point	If a problem has an optimal solution, there is always a corner point that is optimal

Table 31 OR terminology extracted from (Bradley et al., 1977). *12 concepts extracted*

CONCEPT	DEFINITION
OR model	Expresses in amendable manner the mathematical functions that represent the behavior of the assumed real world.
Constraints	Restrictions that affect the decision making, in algebraic form.
Variable	The purpose of the model is to determine their values.
Feasible solution	Solution of the model that satisfies all the constraints
Optimal solution	A feasible solution that yields the best value (maximum or minimum) of the objective function.
Linear programming	The most prominent OR technique.
Graphical representation of a two variable LP	
Feasible solution space	The area in which all the constraints are satisfied simultaneously.
Optimal corner solution	The optimum LP solution is always associated with a corner point of the solution space (where 2 lines intersect)

Table 32 OR terminology extracted from (Taha, 2007). *8 concepts extracted*

7.2 Pre-experiment questionnaire (adapted from (Prensky, 2005))

Demographics

What is your gender?

What is your age?

What is your highest level of education (the latest completed or the one currently in progress)?

What is your current principal occupation?

Interpersonal Competitiveness (5-point Likert scale, Strongly Agree (5) to Strongly Disagree (1))

I perform better when I am competing against someone rather than when I am the only one striving for a goal. (Indicates competitiveness)

I do not feel that winning is important in both work and games. (Indicates lack of competitiveness)

For me, winning an award or a game means that I am the best compared to everyone else that was playing. It is only fair that the best person win the game. (Indicates competitiveness)

In school, I always liked being the first one to be finished with a test. (Indicates competitiveness)

I have always wanted to be better than others. (Indicates competitiveness)

When nominated for an award, I focus on how much better or worse the other candidates' qualifications are as compared to mine. (Indicates competitiveness)

I would want the highest grade possible because that means that I did better than other people. (Indicates competitiveness)

Because it is important that a winner is decided, I do not like to leave a game unfinished. (Indicates competitiveness)

Experience with games[11]

Have you ever played video games? Yes No

Do you currently play video games? Yes No If no on both accounts, skip further questions.

How long have you been playing video games?

a. 6 months

b. 1 year

c. 2-5 years

d. 5-10 years

e. 10 or more years

How often (approximately) do you currently play video games?

[11] Adapted from http://www.silccenter.org/resource-info/video_game_experience_survey2.pdf

a. daily

b. weekly

c. once a month

d. once in 6 months

e. once a year

f. less than once a year or never

Experience with operations research

How much experience (expressed in months) do you have with operations research (loosely known as operations research, decision support systems, prescriptive analytics or management science)?

What is a correct definition of operations research (OR)?

1. The administration of business practices to create the highest level of efficiency possible within an organization. It is concerned with converting materials and labor into goods and services as efficiently as possible to maximize profit.
2. The management of the flow of goods. It includes the movement and storage of raw materials, work-in-process inventory, and finished goods from point of origin to point of consumption.
3. A set of brief descriptive coefficients that summarizes a given data set, which can either be a representation of the entire population or a sample.
4. *An approach to managerial decision making. It attempts to apply mathematical methods and the capabilities of modern computers to the difficult and unstructured problems confronting modern managers.*
5. None of these.
6. I don't know.

What type of operations research problem is the following problem?

A company produces plain and mexican chips. The chips have to go through three processes: slicing, frying, and packing.

The time required by each type of chips for each process is presented in Table 1. There is a limit on the amount of time available for each process (Table 2).The sell price of the chips is presented in Table 3.

You want to determine what combination of chips to produce in order to yield the highest profit possible.

	Slicing(min)	Frying(min)
plain	2	4
mexican	4	5

Table 1, 1kg of potatoes processing time

Available slicing time	345
Available frying time	480

Table 2, machine availability(min)

	Sell price(euro)
plain	2
mexican	1.5

Table 3, chips sell price

1. mixed integer linear programming
2. nonlinear programming
3. linear programming
4. linear system
5. none of these
6. I don't know.

7.3 Post-experiment questionnaire

Satisfaction with the tutorial system

(5-point Likert scale, Strongly Agree (5) to Strongly Disagree (1))

I enjoyed using this system

Completing tasks was frustrating

It makes completing tasks fun

Completing tasks was difficult

It is an effective learning tool

Had to work hard to complete tasks

It is an engaging experience

I felt rushed when completing tasks

Operations research business value

Based on your own experience/knowledge, can you describe cases in which operations research can be used to improve the situation?

Operations research concepts

1. Fill in the blanks by selecting the appropriate words (for every blank space, three words were offered, including the correct one).
    1. Operations research is a discipline that deals with the application of mathematics to help make better decisions. Also known as: management science.
    2. In broad terms, mathematical programming can be defined as a mathematical representation aimed at programming or planning the best possible allocation of scarce resources.
    3. A mathematical model is an abstract model that describes, in general mathematical terms, the relations contained in a model.
    4. An optimization model is a mathematical model that contains a criterion or objective, which we seek to optimize (e.g. maximize or minimize), subject to a set of constraints that portray the conditions under which the decisions have to be made.
    5. A solution is a proposal of specific values for the decision variables.
    6. Linear programming consists of optimization models made up from linear equations and linear inequalities. The feasible decisions are compared using a linear objective function that depends on the decision variables.
    7. The objective function consists of decision variables that need to be maximized or minimized. It specifies the criterion the decision maker will use to evaluate alternative solutions to the problem.
    8. A constraint is an inequality or equality defining limitations on decision variables.
    9. A feasible solution is a solution that satisfies all of the constraints.
    10. An infeasible problem is a problem with no feasible solution.

11. Parameters are the collection of coefficients, representing known data, which are used in the model.
12. Decision variables are the quantities that the decision makers would like to determine. They are the unknowns of a mathematical programming model.
13. An optimal solution is a feasible solution that yields the best value (maximum or minimum) of the objective function.

2. In the following text, highlight all the operations research concepts that you recognize:

"An iron foundry has a firm order to produce 1000 pounds of metal mixture containing at least 0.45% manganese and between 3.25% and 5.50% silicon. As this is a special order, the foundry has no stock of metal with these properties. The mixture will sell for $0.45 per pound. The foundry has three types of iron available in essentially unlimited amounts, with the following properties:

	Iron A	Iron B	Iron C
Silicon	4%	1%	0.6%
Manganese	0.45%	0.5%	0.4%

Further, the production process is such that pure manganese can also be added directly to the melt. The producer does not want to keep any supply of the metal in stock. The costs of the various possible inputs are presented in the table below. To keep the model simple, other production costs will be ignored.

Metal	Price/unit
Iron A	$26/thousand pounds
Iron B	$30/thousand pounds
Iron C	$20/thousand pounds
Manganese	$ 8/pound

Out of what selection of raw materials should the foundry produce the metal mixture in order to maximize profits?"

3. Out of the following problems, check the ones that are LP problems:
   - A firm produces goods at 2 different supply centers, *i* = 1, 2. The supply produced at supply center *i* is *Si*. The demand for the good is spread out at 3 different demand centers, *j* = 1, 2, 3. The demand at the *j*th demand center is *Dj*. The problem of the firm is to get goods from supply centers to demand centers at minimum cost. The cost of shipping from supply center *i* to demand center *j* is a linear function of the volume shipped: cost = units_shipped * 0.7.
   - An investor has $5000 and two potential investments. Let Xi for i = 1 and i = 2 denote his allocation to investment i in thousands of dollars. From historical data, investments 1 and 2 have an expected annual return of 20 and 16 percent, respectively. Also, the total risk involved with investments 1 and 2, as measured by the variance of total return, is given by 2*X1*X1 +X2*X2 + (X1+X2)*(X1+X2), so that risk increases with total investment and with the amount of each individual investment. The investor would like to maximize his expected return and at the same time minimize his risk.
   - A store wants to liquidate 200 of its shirts and 100 pairs of pants from last season. They have decided to put together two offers, A and B. Offer A is a package of one shirt and a pair of pants which will sell for $30. Offer B is a package of three shirts and a pair of pants, which will sell for

$50. The store does not want to sell less than 20 packages of Offer A and less than 10 of Offer B. How many packages of each do they have to sell to maximize the money generated from the promotion?

7.4 Post-experiment questionnaire expected responses

7.4.1 ILO: Operations research concepts

1. Operations research is a discipline that deals with the application of mathematics to help make better decisions. Also known as: management science.
2. In broad terms, mathematical programming can be defined as a mathematical representation aimed at programming or planning the best possible allocation of scarce resources.
3. A mathematical model is an abstract model that describes, in general mathematical terms, the relations contained in a model.
4. An optimization model is a mathematical model that contains a criterion or objective, which we seek to optimize (e.g. maximize or minimize), subject to a set of constraints that portray the conditions under which the decisions have to be made.
5. A solution is a proposal of specific values for the decision variables.
6. Linear programming consists of optimization models made up from linear equations and linear inequalities. The feasible decisions are compared using a linear objective function that depends on the decision variables.
7. The objective function is a function that consists of decision variables that needs to be maximized or minimized which specifies the criterion the decision maker will use to evaluate alternative solutions to the problem.
8. A constraint is an inequality or equality defining limitations on decision variables.
9. A feasible solution is a solution that satisfies all of the constraints.
10. An infeasible problem is a problem with no feasible solution.
11. Parameters are the collection of coefficients, representing known data, which are used in the model.
12. Decision variables are the quantities that the decision makers would like to determine. They are the unknowns of a mathematical programming model.
13. An optimal solution is a feasible solution that yields the best value (maximum or minimum) of the objective function.

Operations research concepts that the user should recognize:

Objective function: “maximize profits”

Decision variables: any reference to Iron A and any reference to Iron B and any reference to Iron C and any reference to manganese.

Parameters: any references to the tables, sell for $0.45 per pound

Constraints: containing at least 0.45% manganese and between 3.25% and 5.50% silicon, 1000 pounds

7.4.2 ILO: Recognize a simple LP problem

Problem 3

7.5 Digital interactive tutorials

The literature review was executed between 01.08.2014 and 02.08.2014.

In Figure 11 we present the mind map showing the output of the snowball process for digital interactive tutorials guidelines. For readability purposes, the image has been split into Figure 12, Figure 13, Figure 14, and Figure 15.

- Crossed out publications were not relevant to the research objectives (highlighted in blue in the middle).
- The dashed blue arrows represent connect the same publication referred in different sources (the arrow points to the first mention of the paper, the source will always appear as crossed out)
- The green checkmarks indicate the completion of a particular step of the snowballing process
- The relations exist only for the researcher's reference, they do not reflect the level of saturation.

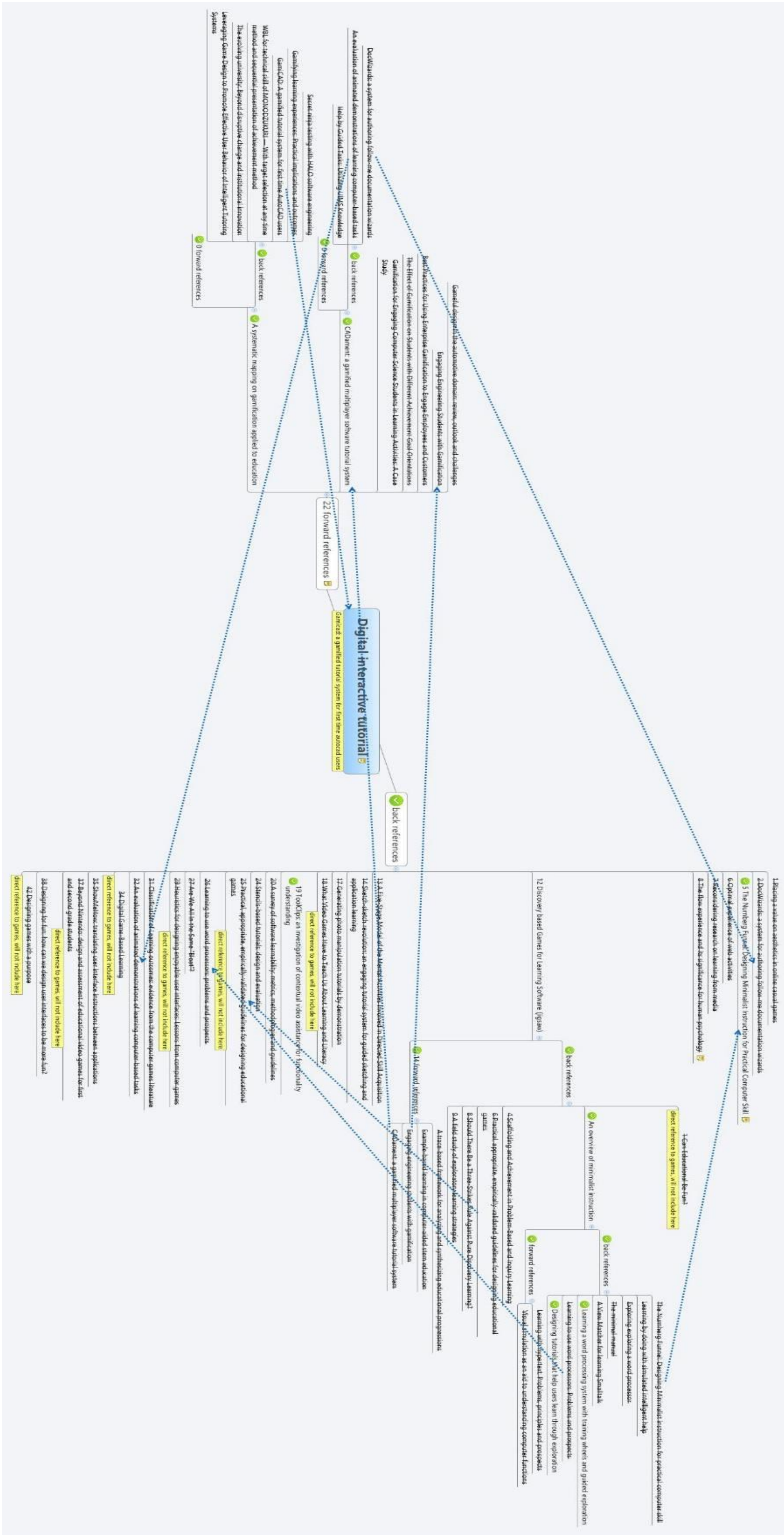


Figure 11 Overview of the snowball process searching for guidelines for designing digital interactive tutorials

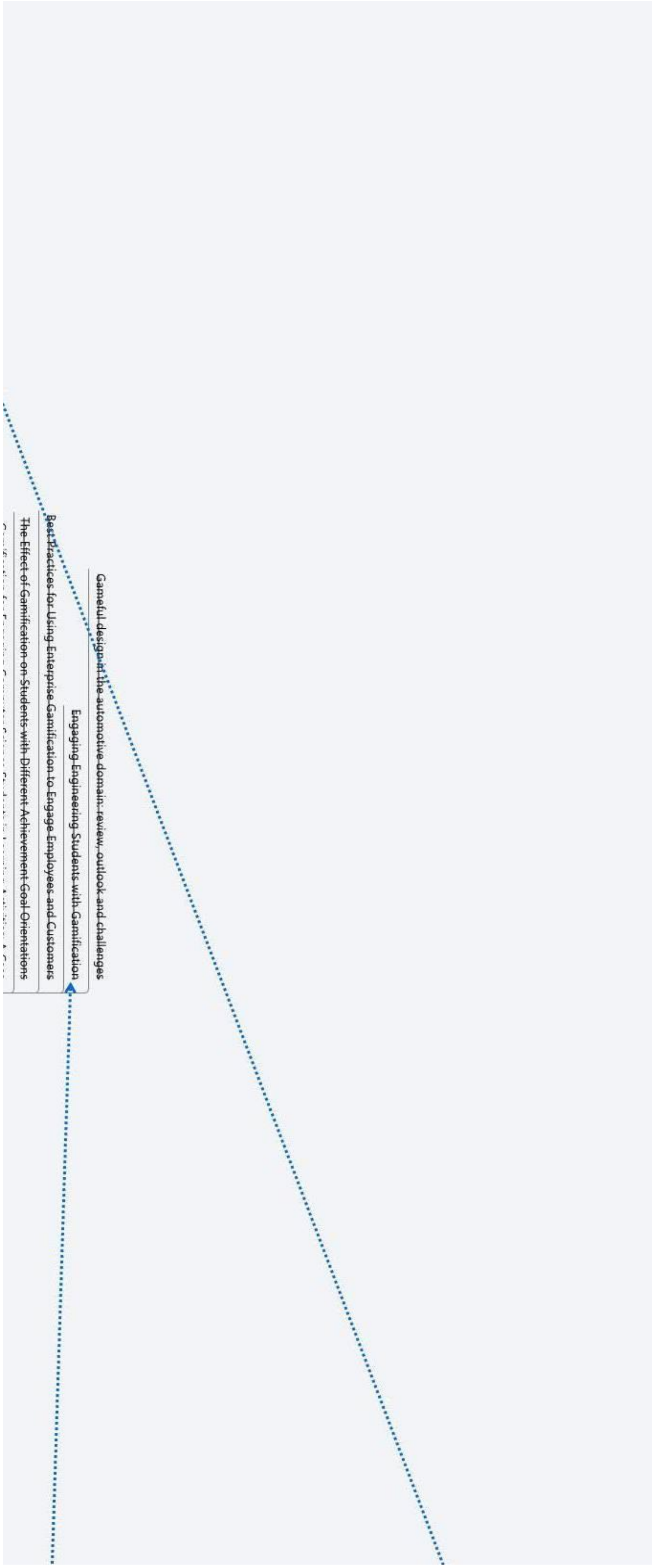


Figure 12 Upper left

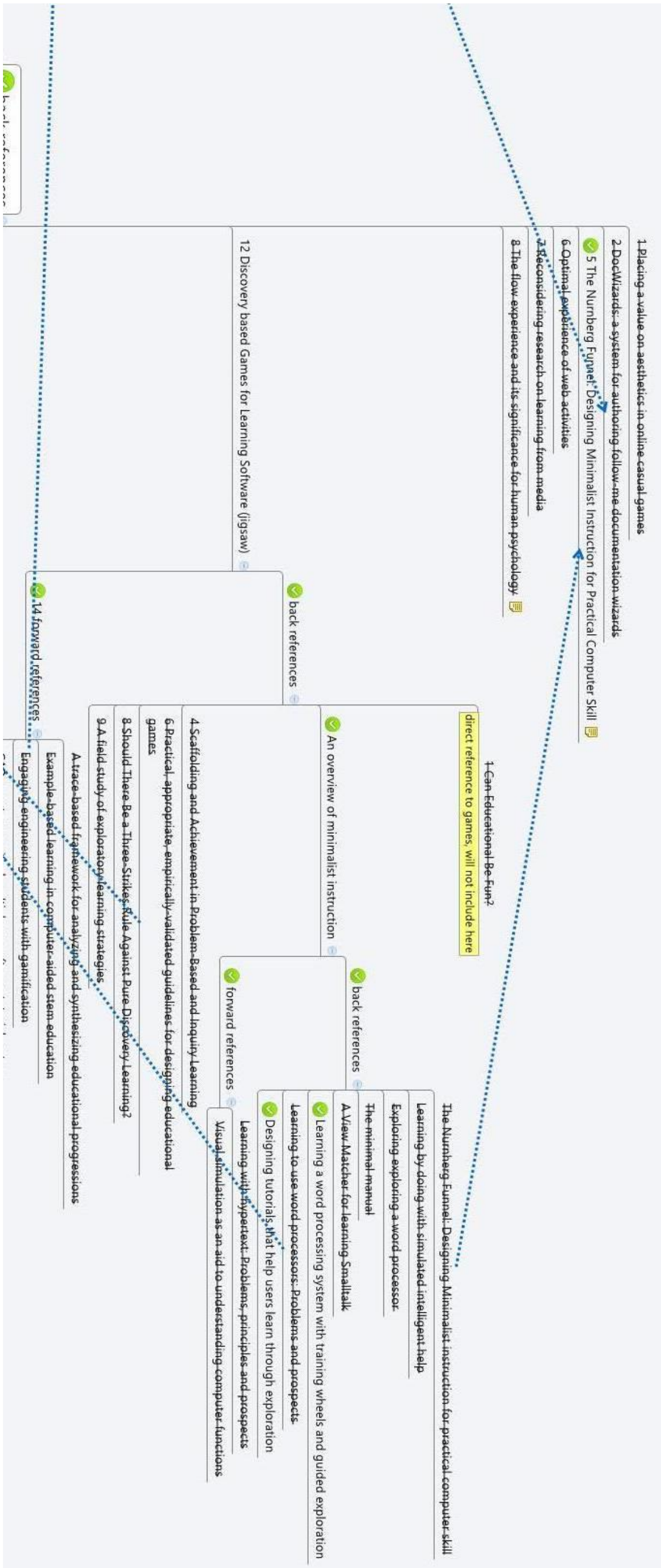


Figure 13 Upper right

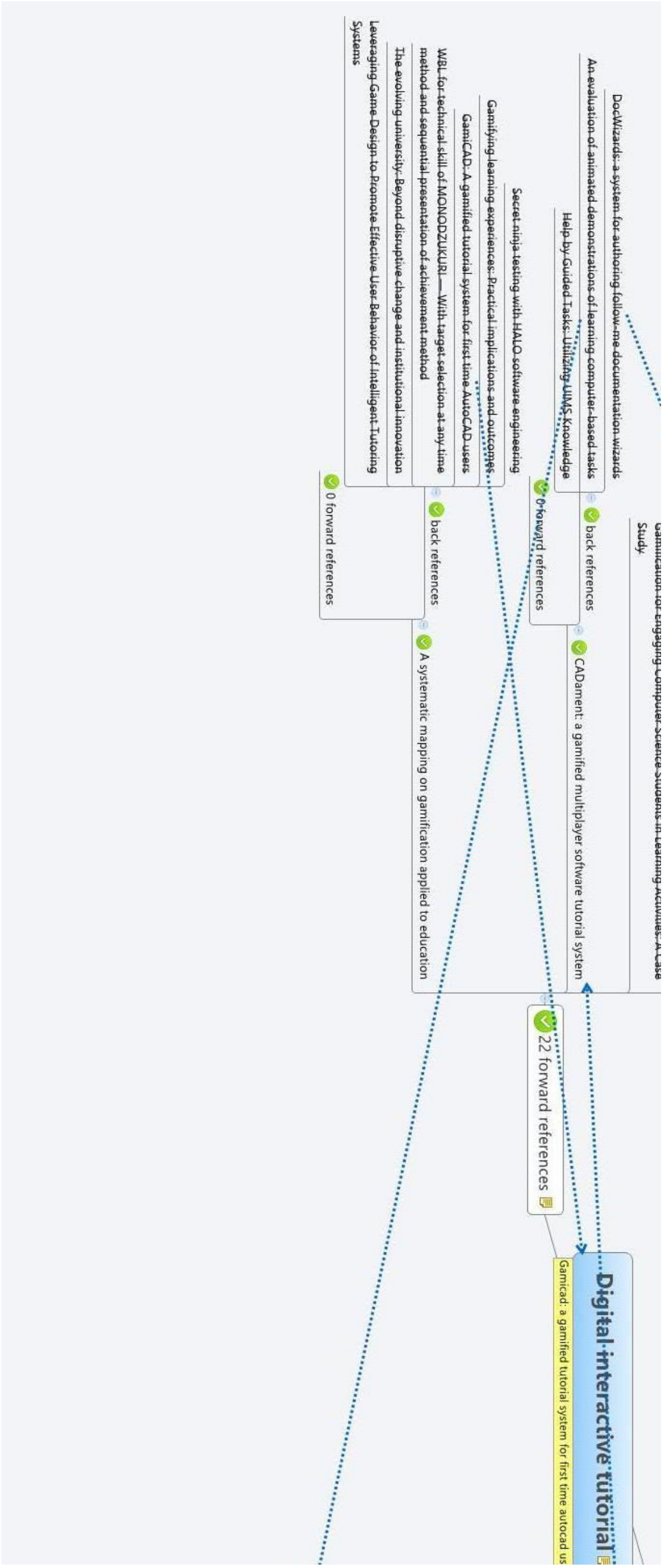


Figure 14 Lower left

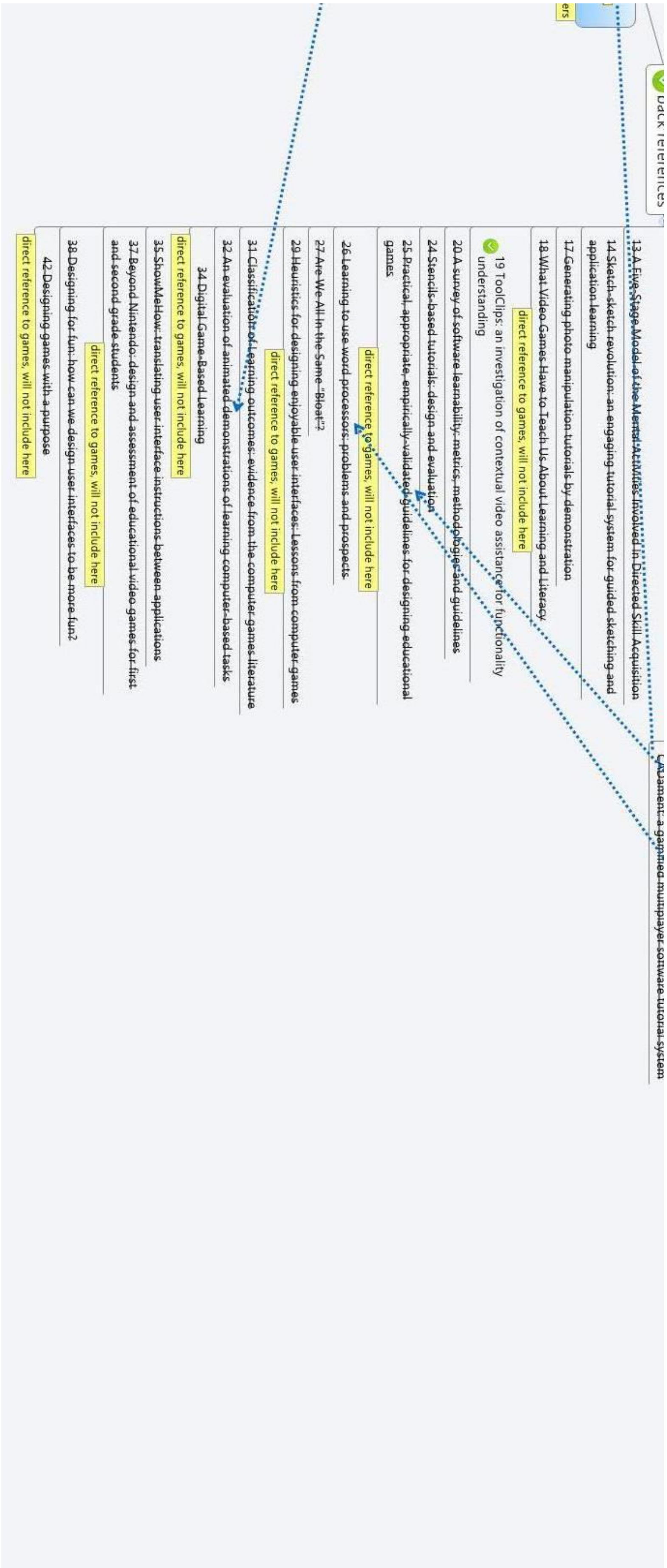


Figure 15 Lower right

7.6 Gamification in the learning context

The literature review was executed between 14.09.2014 and 26.09.2014.

In Figure 16 we present the mind map showing the output of the snowball process for Gamification in the learning context guidelines. For readability purposes, the image has been split into Figure 17, Figure 18, Figure 19, Figure 20, Figure 21, Figure 22 and Figure 23 (we omitted the upper left and lower right sections because they did not contain anything).

- Crossed out publications were not relevant to the research objectives (highlighted in blue in the middle).
- The dashed blue arrows represent connect the same publication referred in different sources (the arrow points to the first mention of the paper, the source will always appear as crossed out)
- The green checkmarks indicate the completion of a particular step of the snowballing process.
- The relations exist only for the researcher's reference, they do not reflect the level of saturation.

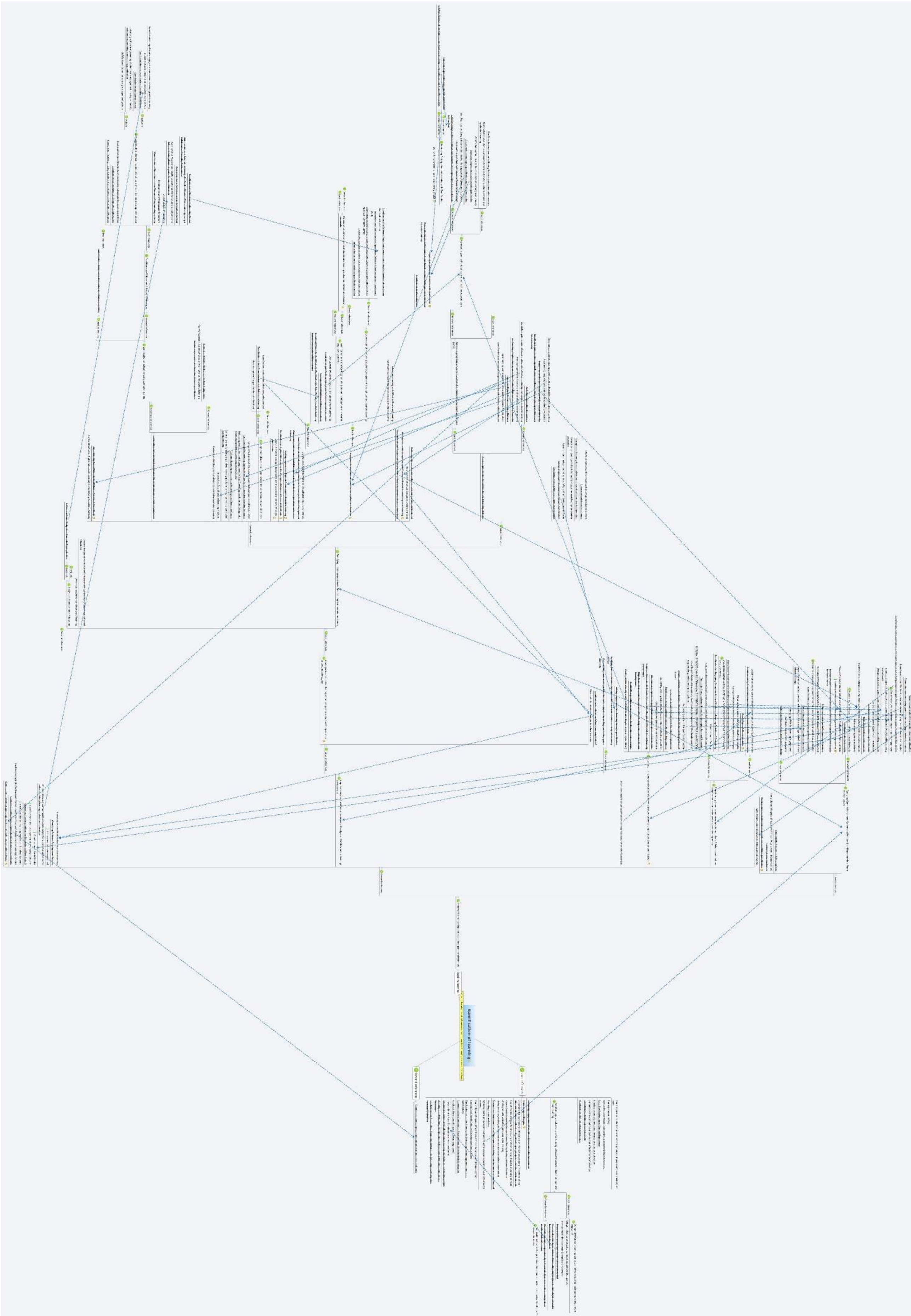


Figure 16 Overview of the snowball process searching for guidelines for applying gamification to education/learning contexts

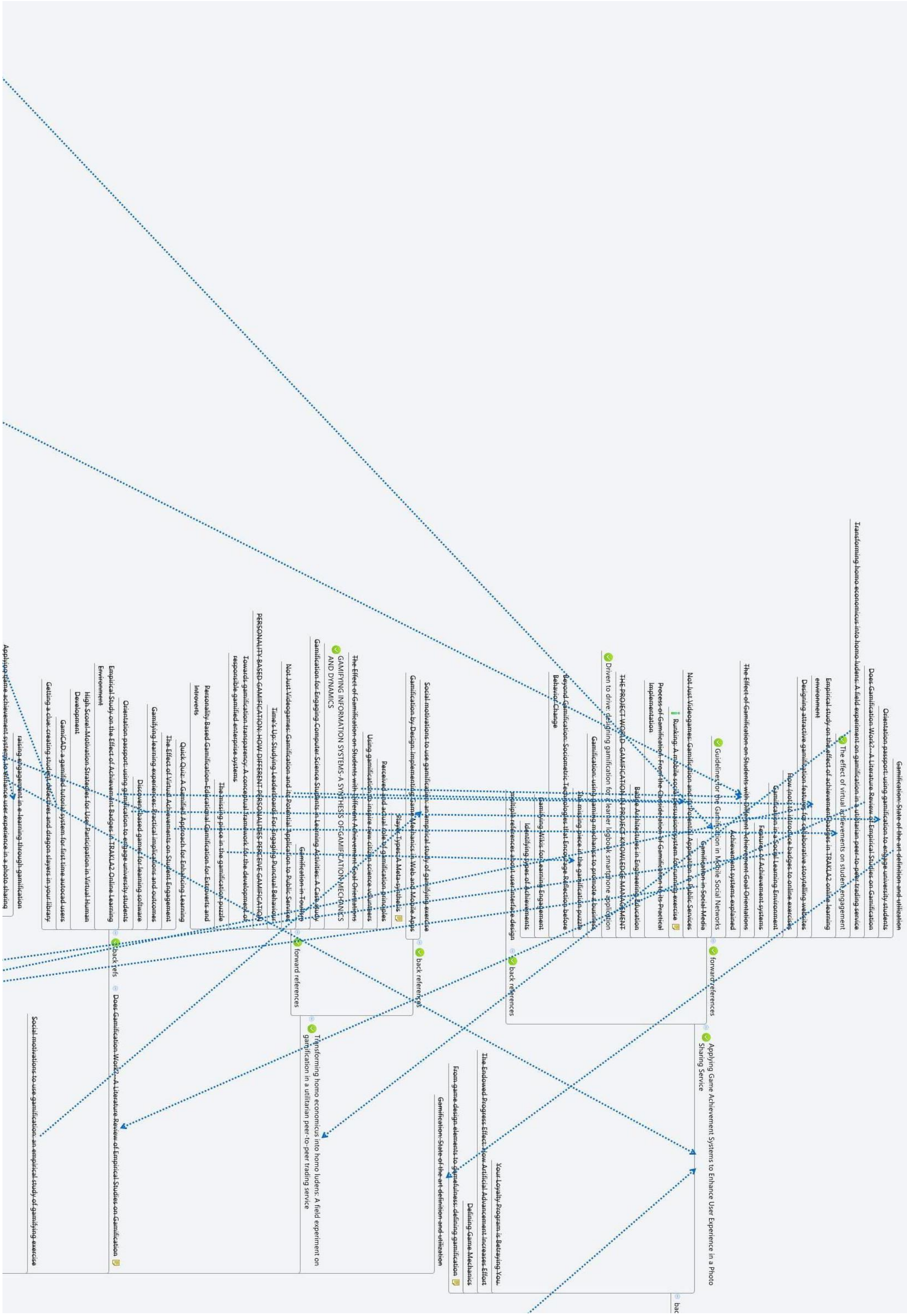


Figure 17 Upper middle

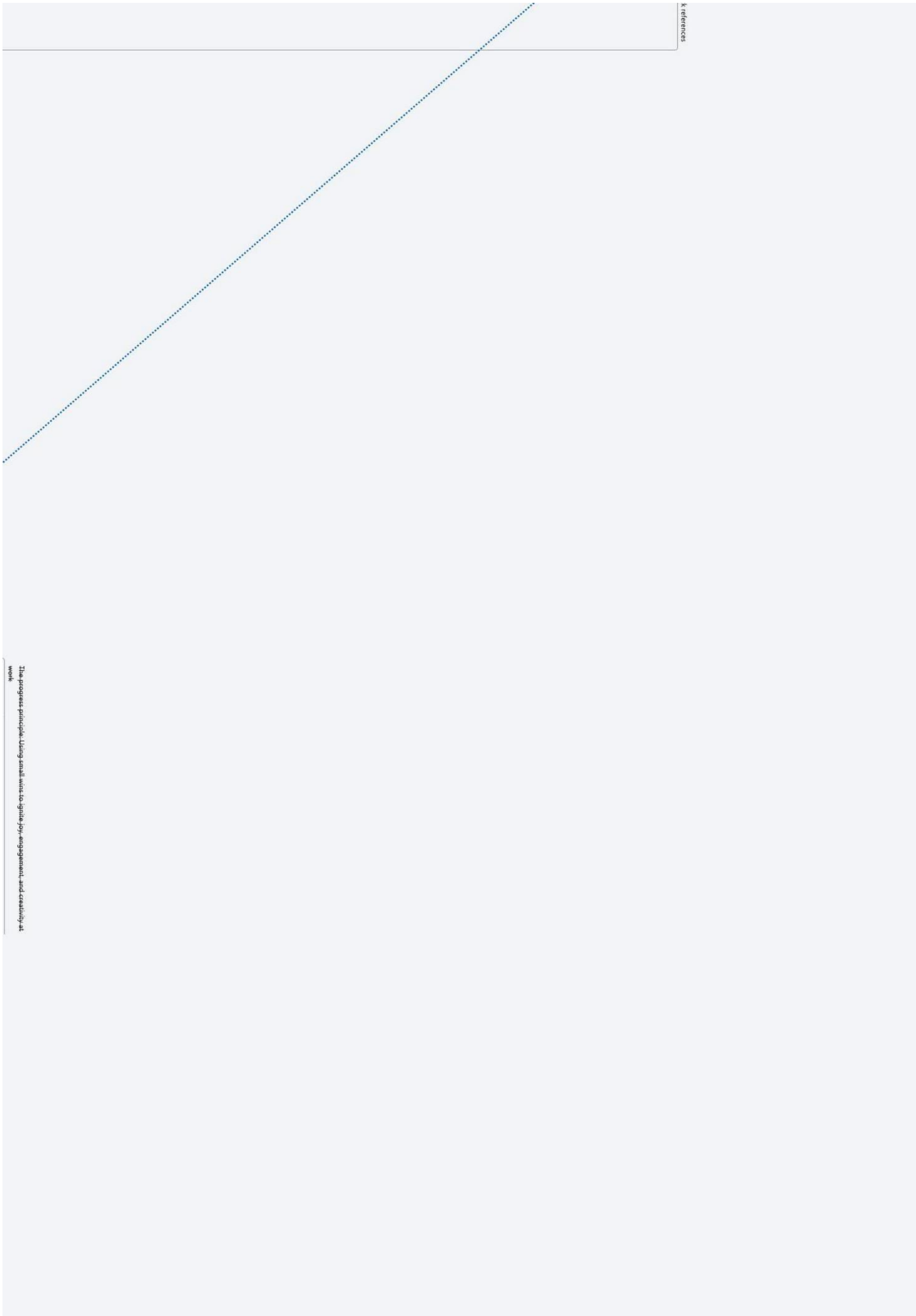


Figure 18 Upper right

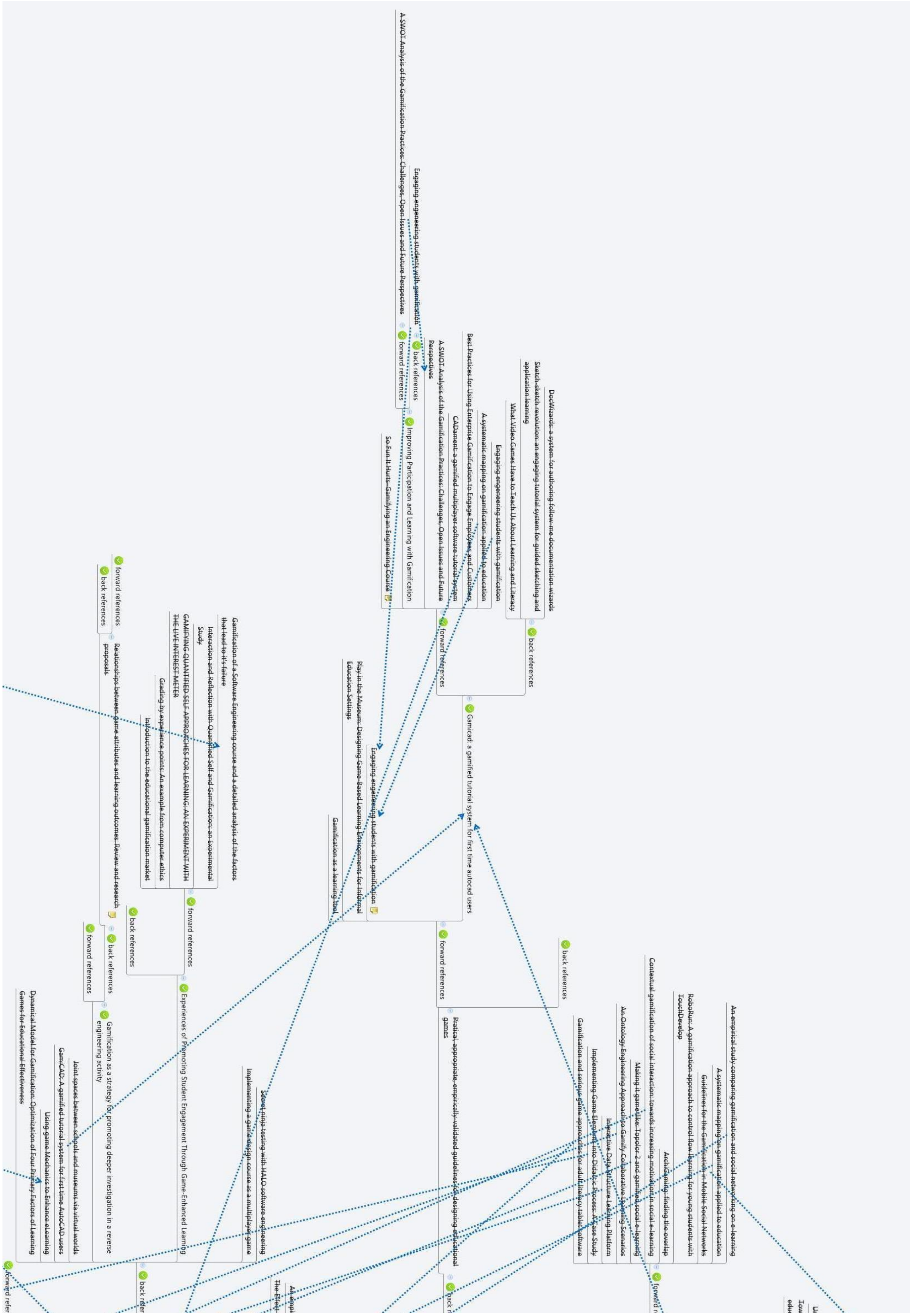


Figure 19 Middle left

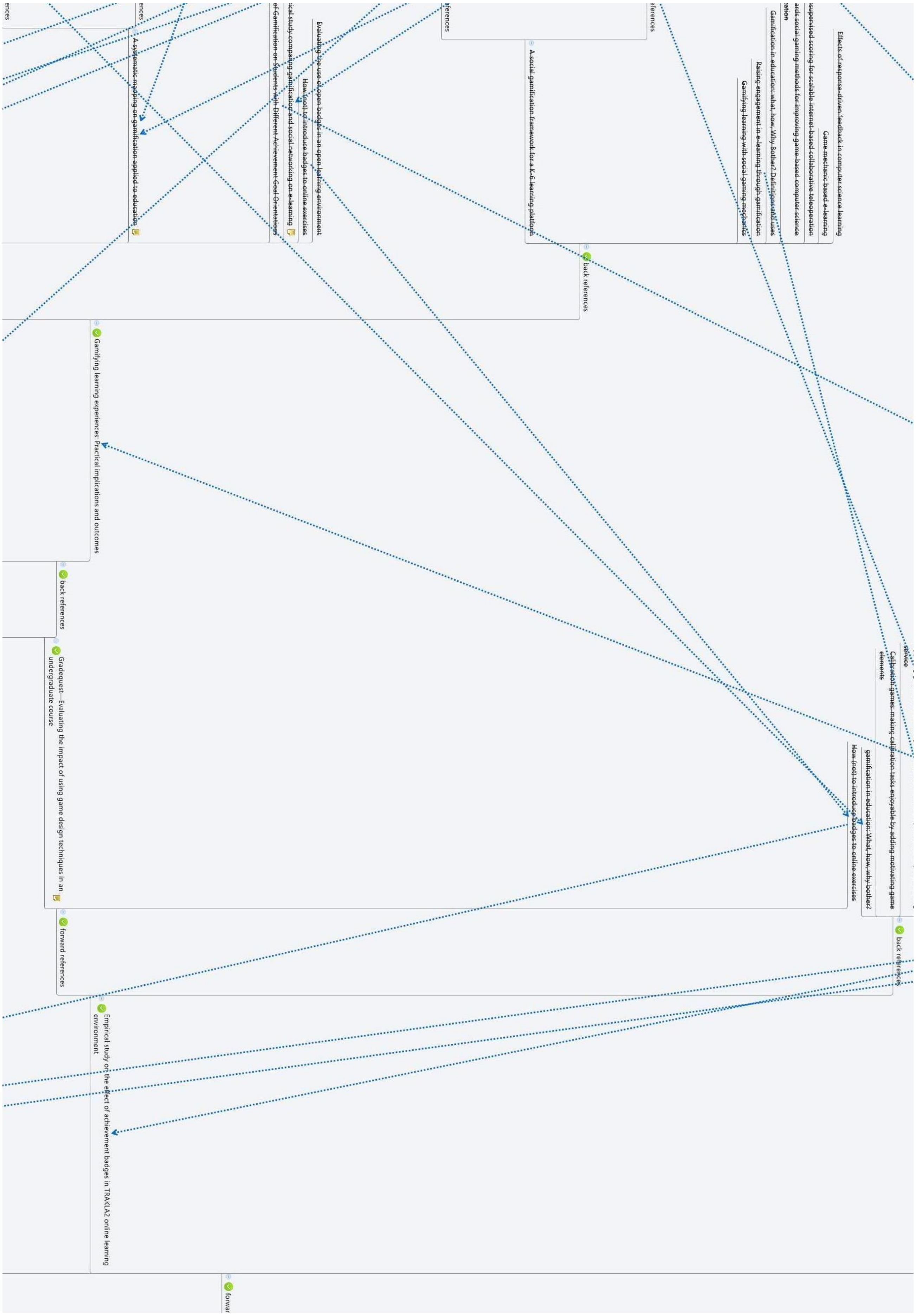


Figure 20 Middle middle

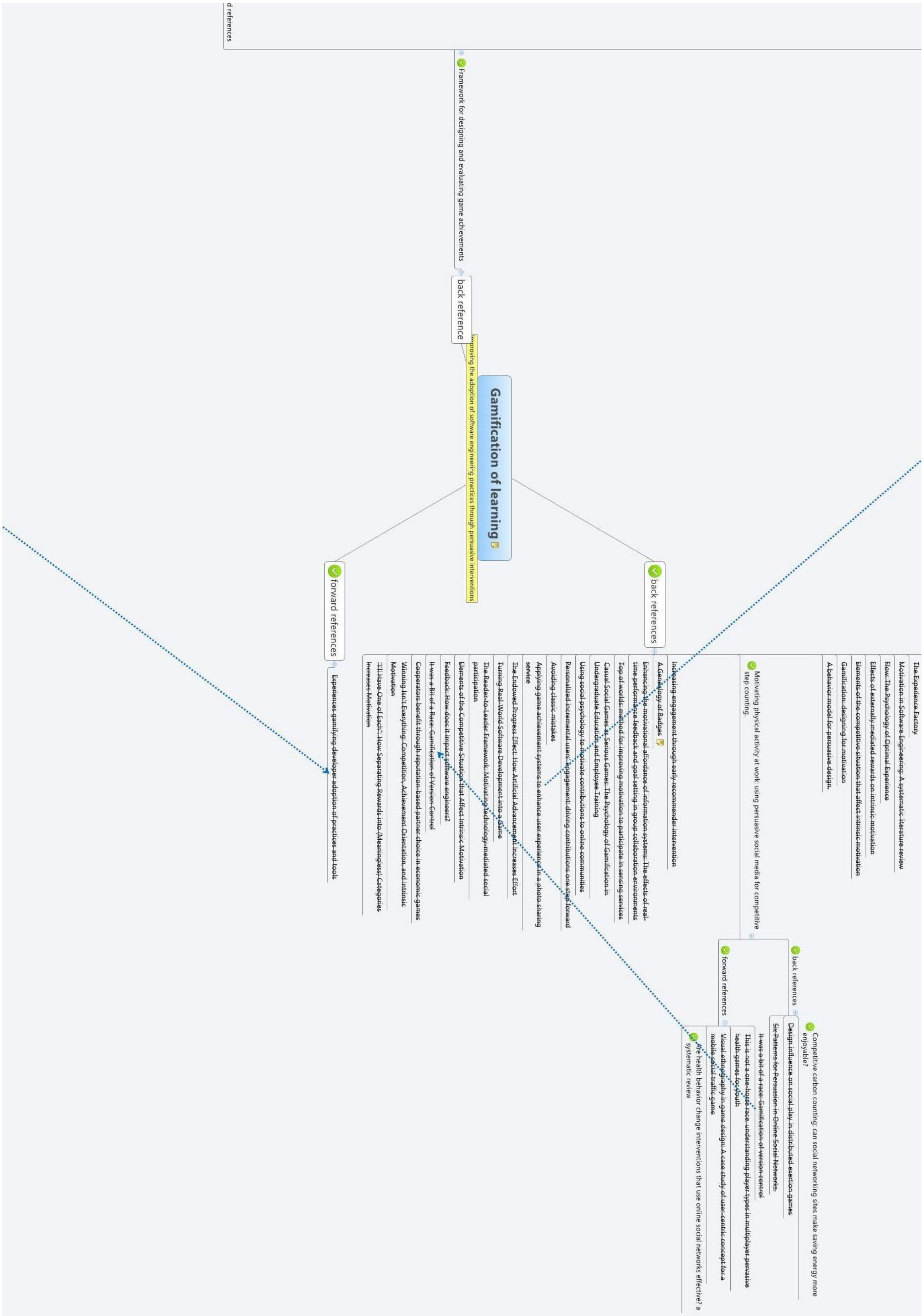


Figure 21 Middle right

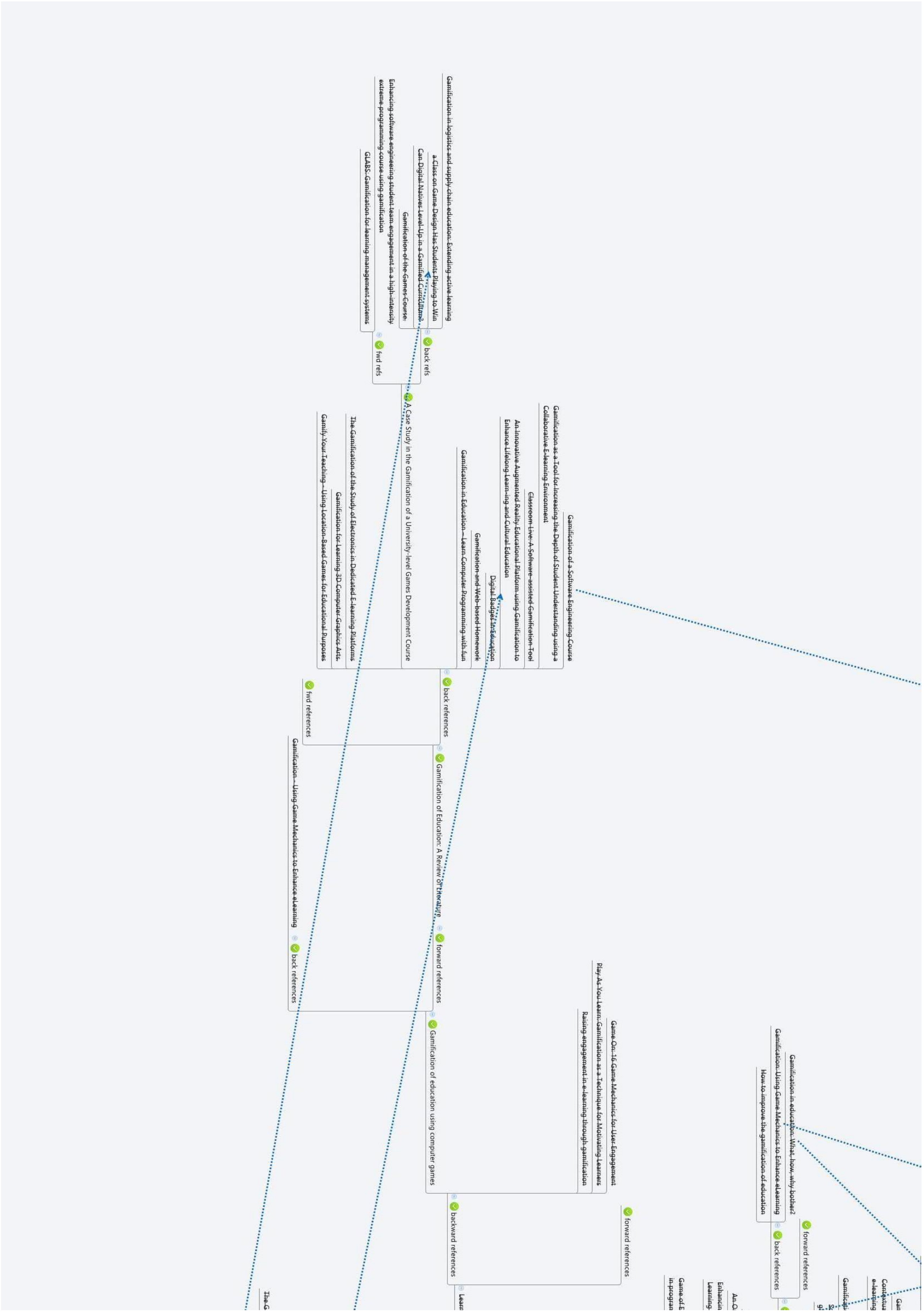


Figure 22 Lower left

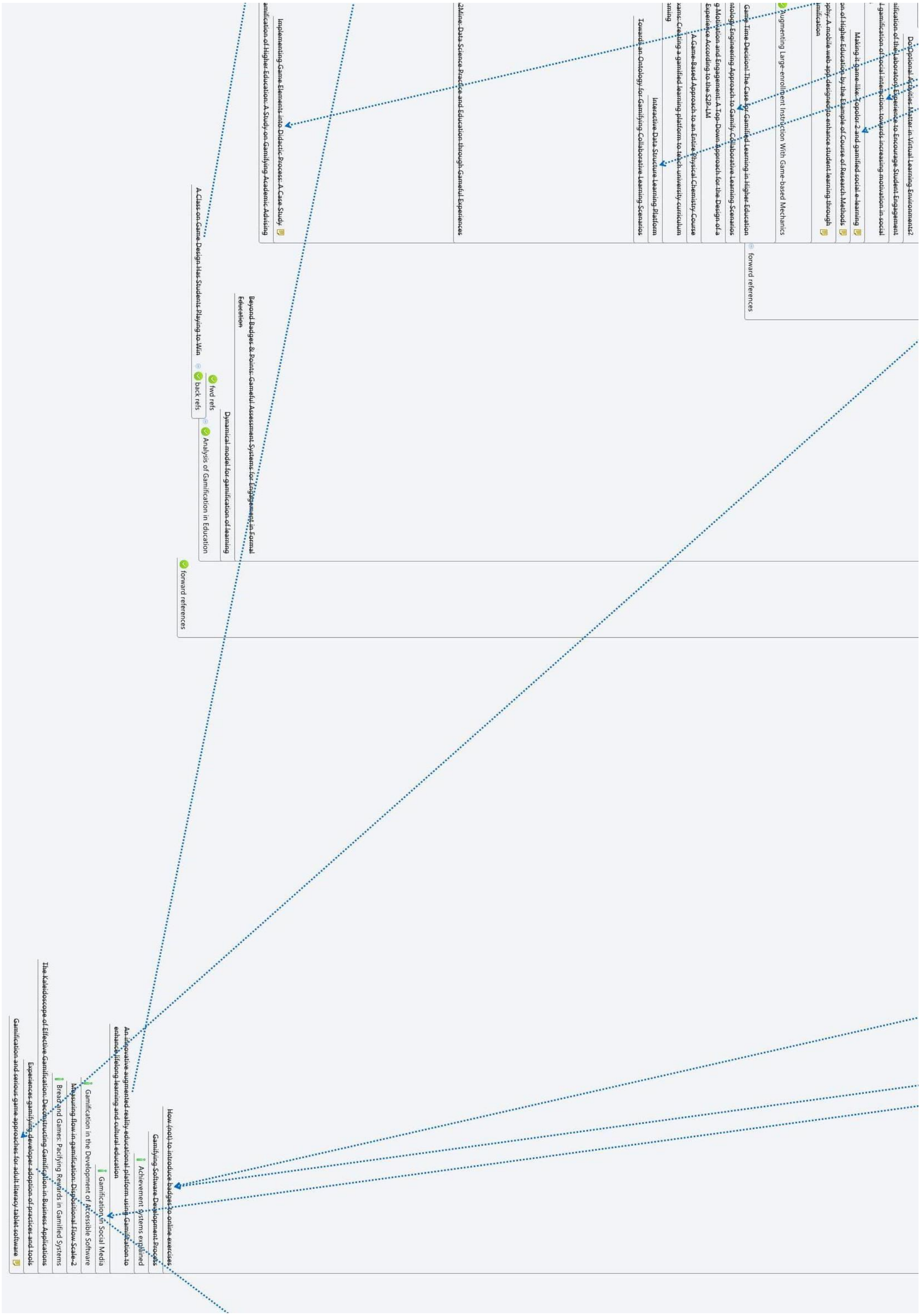


Figure 23 Lower middle

7.7 Screenshots of the OR digital interactive tutorial

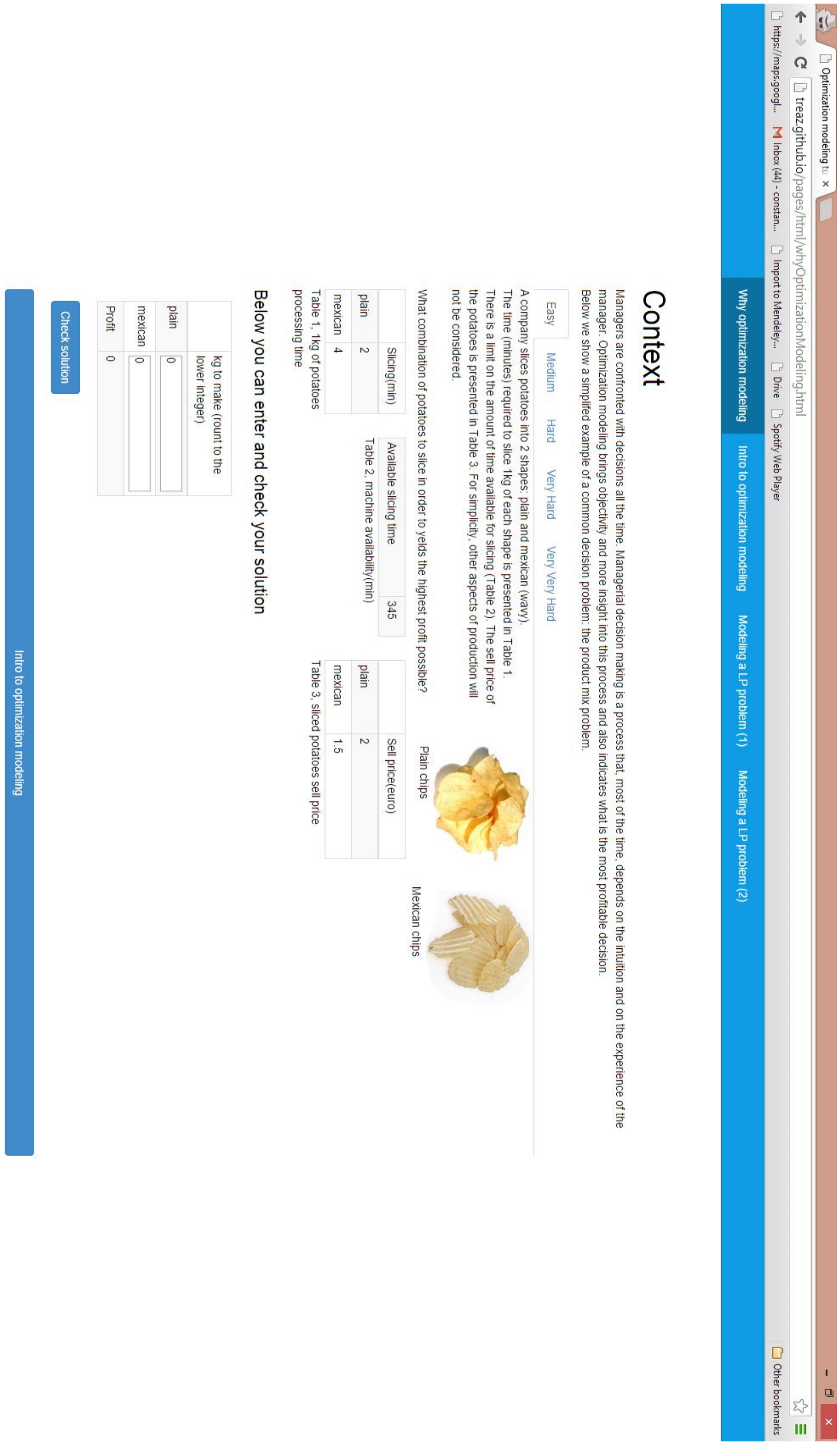


Figure 24 Why operations research page

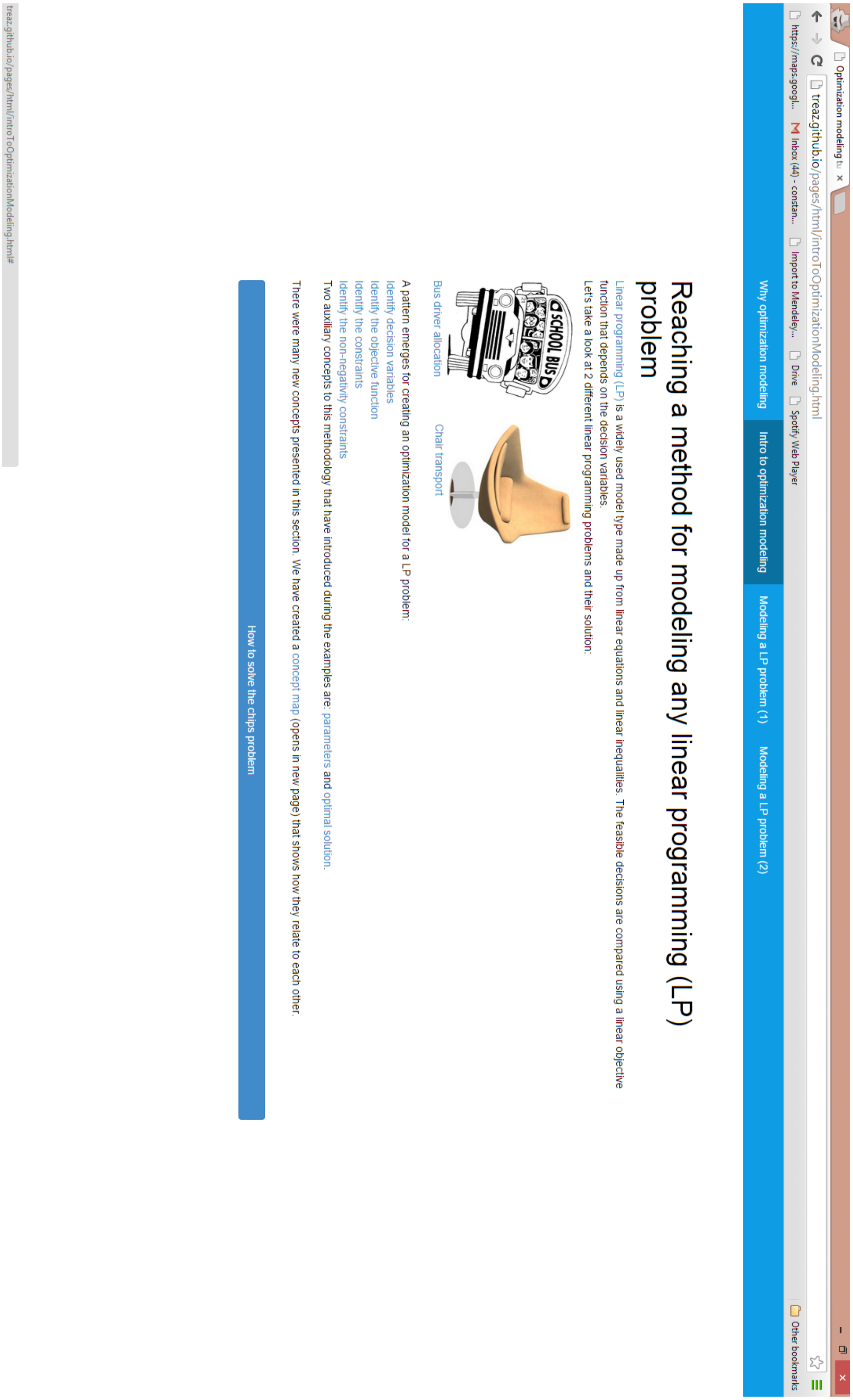


Figure 25 Intro to operations research page

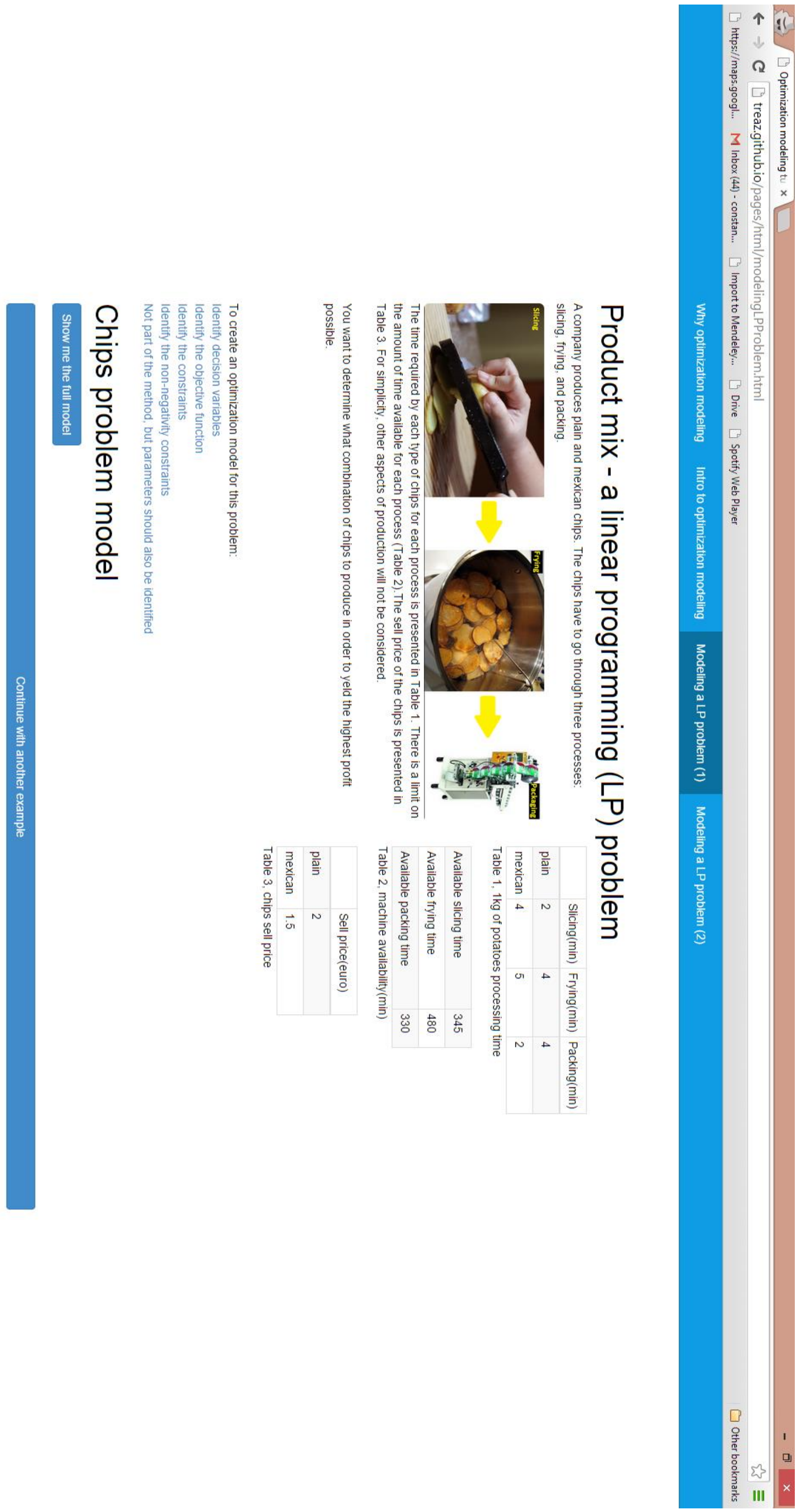


Figure 26 Modeling a LP problem (1) page

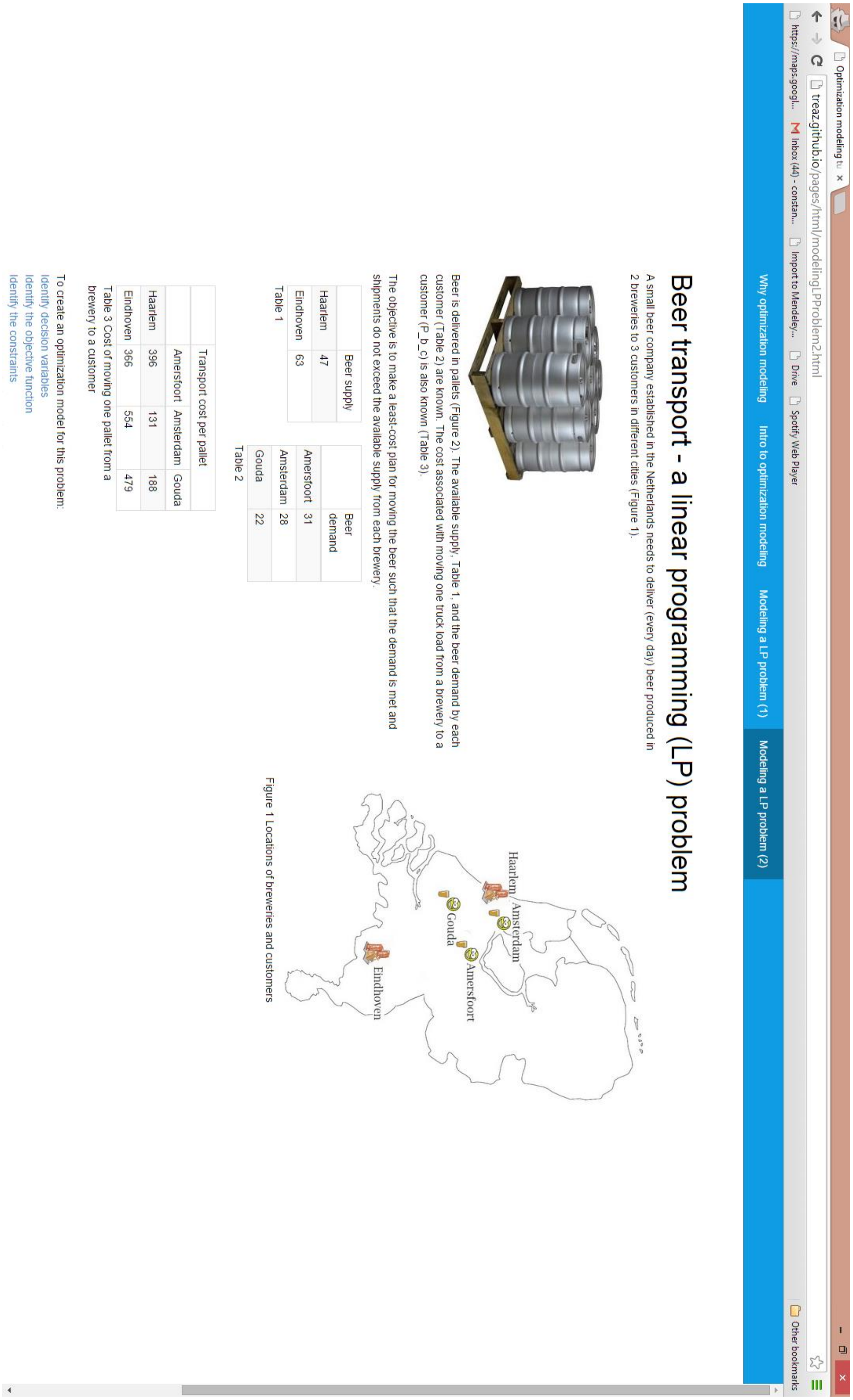


Figure 27 Modeling a LP problem (2) page

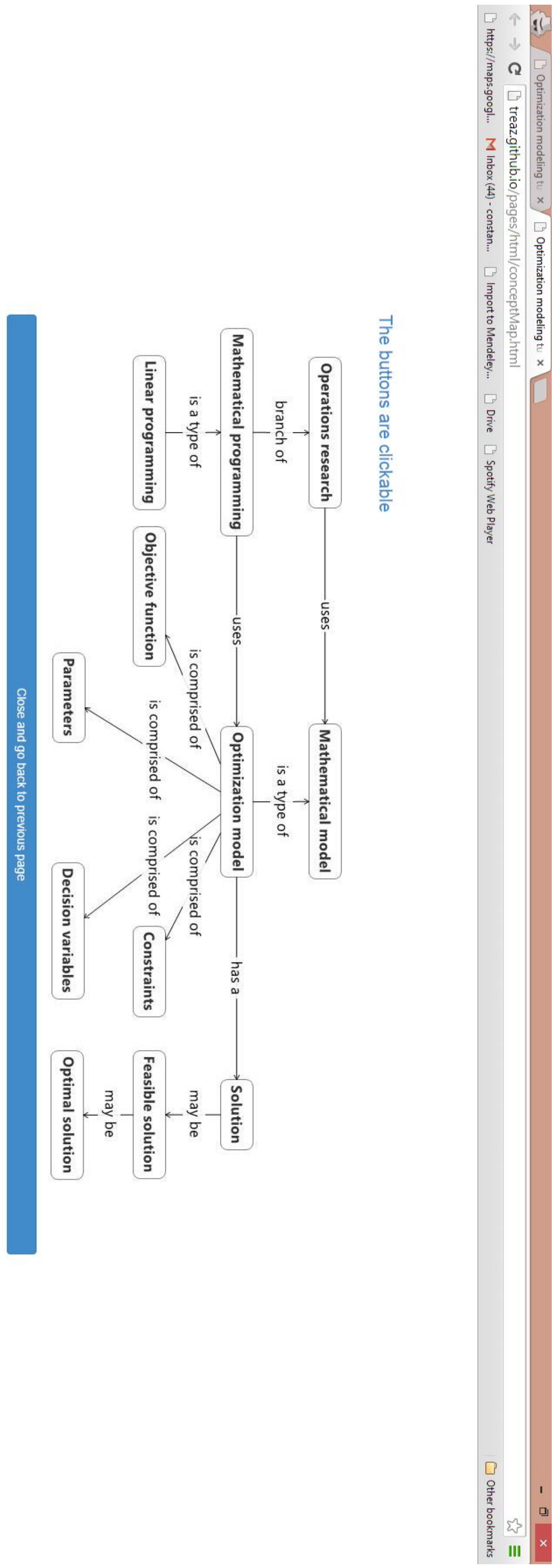


Figure 28 Concept map page

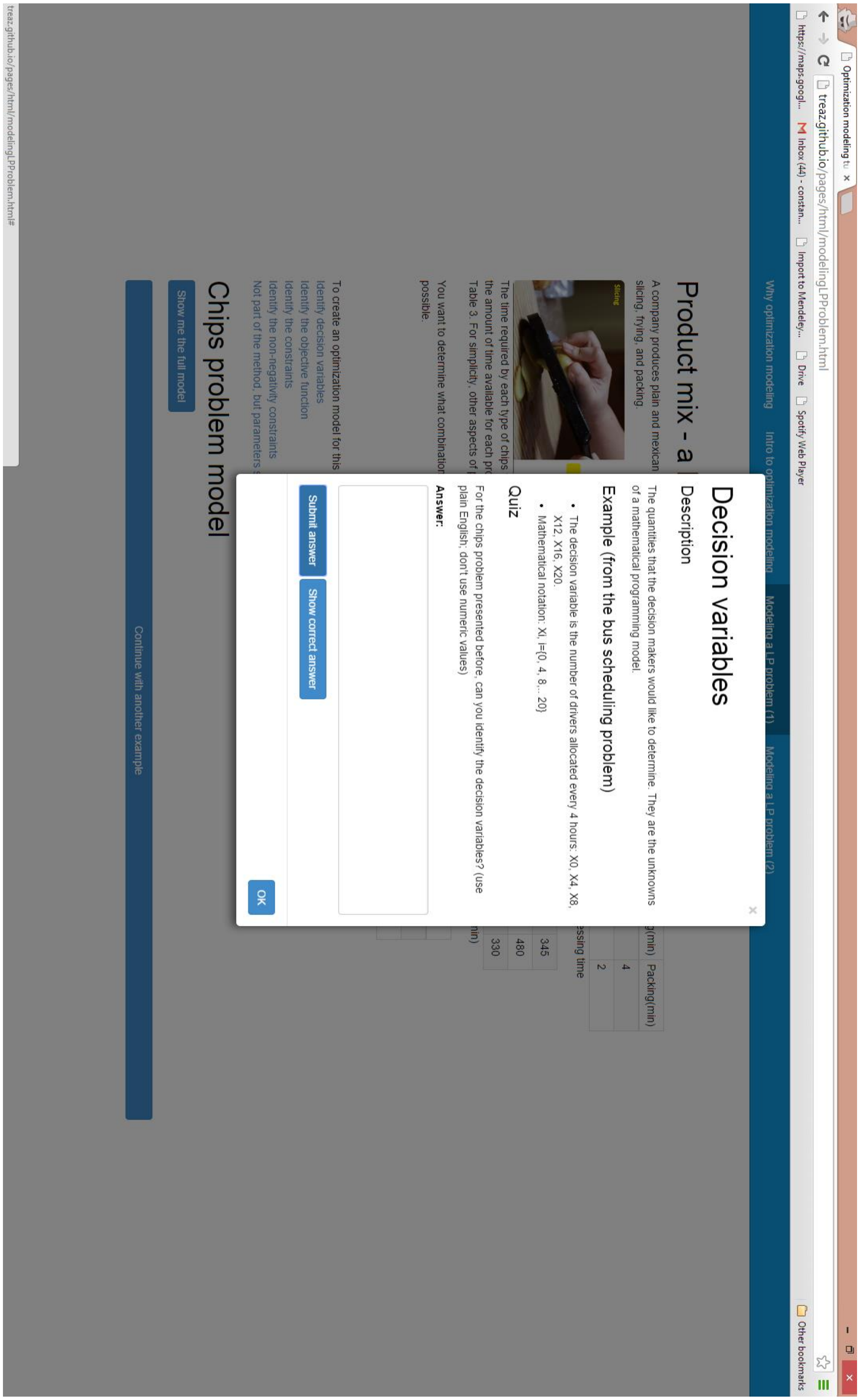


Figure 29 An OR concept quiz page

7.8 Screenshots of the gamified OR digital interactive tutorial

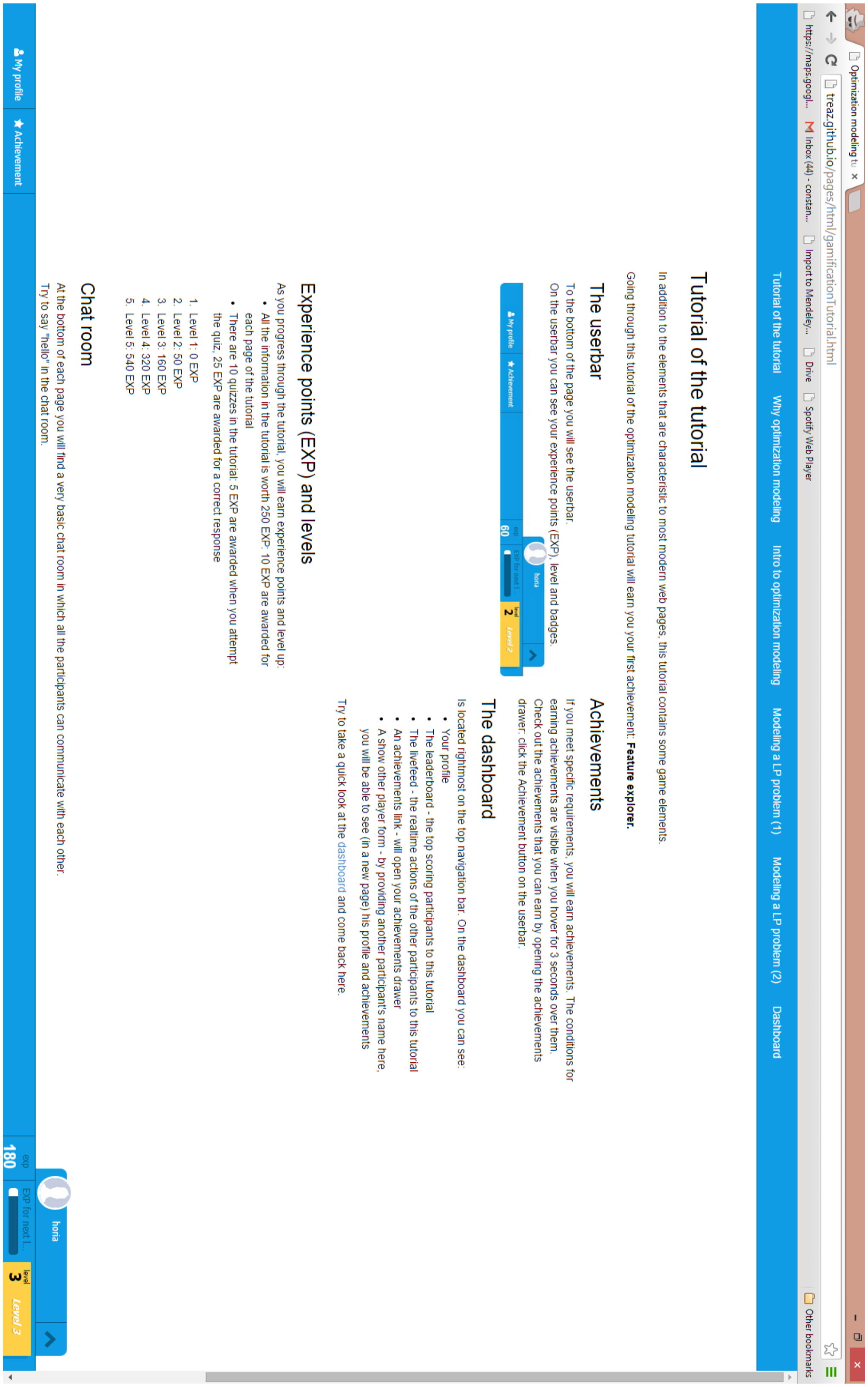


Figure 30 The tutorial of the tutorial page - explains the basic rules of the gamified application

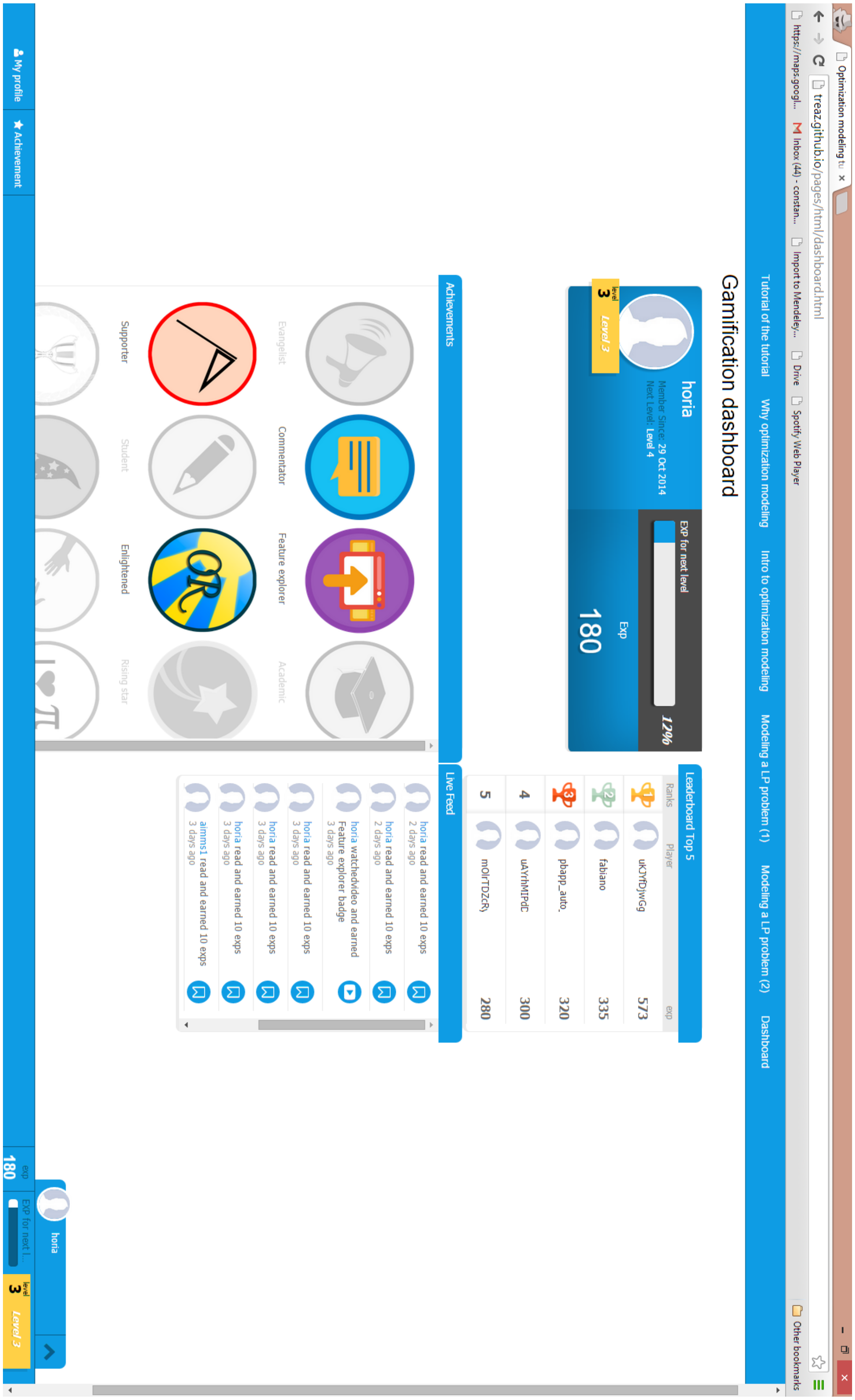


Figure 31 The gamification dashboard page

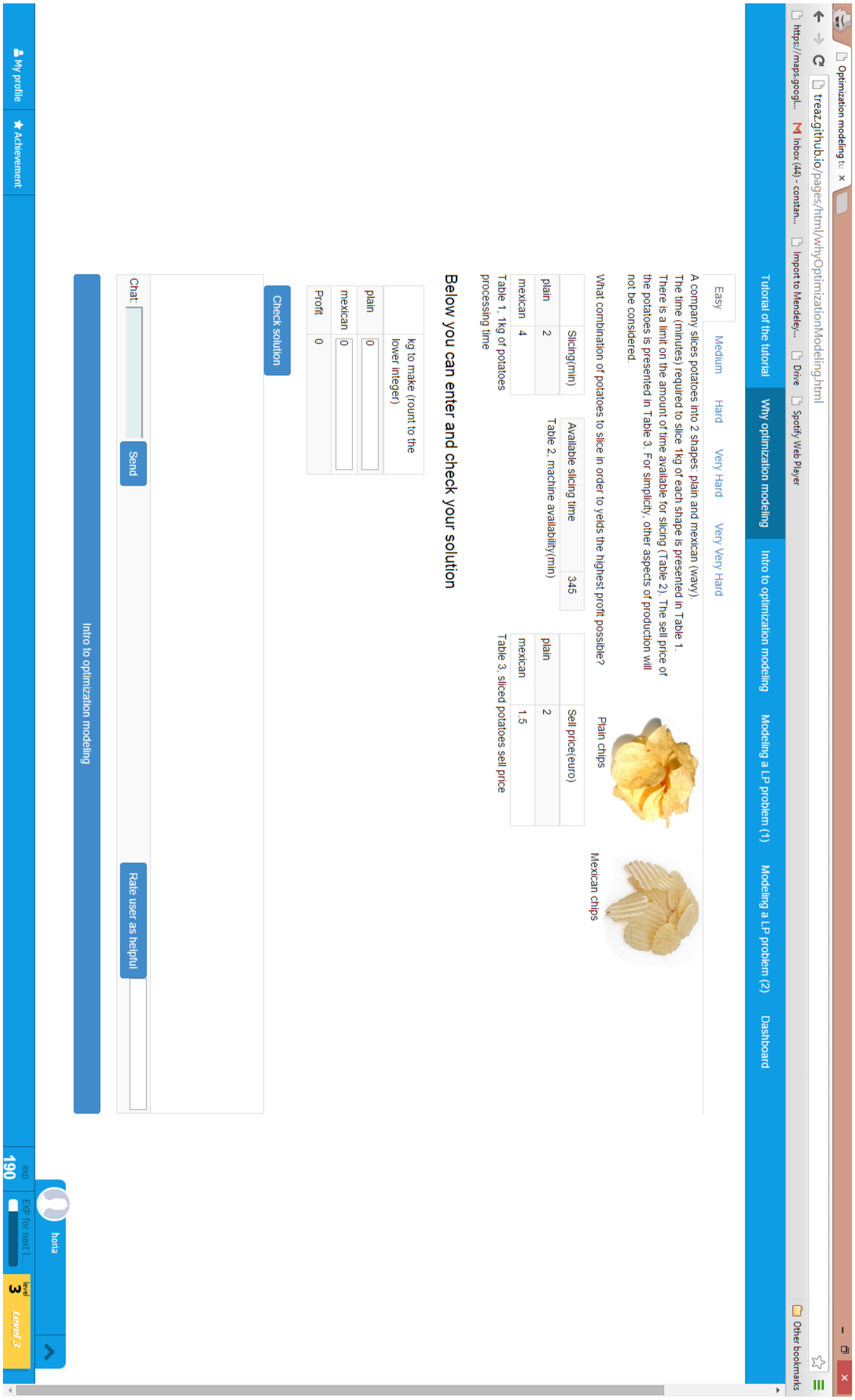


Figure 32 A perspective of the chatroom

7.9 Email template sent to participants to online tutorial

Hello participant,

To make it easier for you to participate, the experiment is now delivered online:

1. a pre questionnaire ~5 minutes
2. a tutorial (try to spend at least 50 minutes here)
3. a post questionnaire ~16 minutes

Before you get started, a few notes:

- Participation in the experiment is anonymous (hence the general usename).
- If a section feels too hard, move on. You might come back to it later, you might not.
- Please do not cheat.

Your credentials:
usename: user48
password: CFDn

The link to the pre questionnaire is:
https://treaz.typeform.com/to/jAnajC