Driving engagement and online social behavior of employees in an enterprise environment

Version of August 18, 2014

Laurentiu Catalin Stanculescu
Driving engagement and online social behavior of employees in an enterprise environment

THESIS submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER SCIENCE

by

Laurentiu Catalin Stanculescu
born in Bucharest, Romania
Driving engagement and online social behavior of employees in an enterprise environment

Author: Laurentiu Catalin Stanculescu
Student id: 4236963
Email: catalaur@gmail.com

Abstract

In the recent years, gamification, “the use of game design elements in non-game contexts”, has drawn the attention of an increasing number of scientists. Although several studies highlighted the benefits of gamification in several applications, its potential in the enterprise environment still needs to be understood.

In the enterprise context, companies strive to foster positive behaviour in employees in order to achieve important business needs. For example, it is important for enterprises to educate their employees about the company’s values, products and services. Also essential is to improve the work environment by facilitating social interaction between employees online. Lastly, companies also need to incentivise their employees to spread awareness of the enterprise’s vision and interests beyond the boundaries of the company.

This thesis contributes to the studies on enterprise gamification with a study performed at a large multinational enterprise. We designed and implemented a modular and extensible framework for studying gamification and instantiated it as a Q&A game combined with news sharing and social connections capabilities. We used the experimental tool to test the effectiveness of several game mechanics for promoting several types of behaviour.

The study involved N=206 IBM employees, organised in 4 experimental groups, and lasted between May 12th and July 11th, 2014. Results show that the implementation of game mechanics was beneficial for the goals of this thesis. First, the tool was received well by employees, who quickly became engaged. They answered quiz questions, invited colleagues to play the game and shared news. The level of engagement was dependent on the game mechanics assigned to an experimental group. Second, the tool has been proved successful for promoting knowledge acquisition. Third, the results related to social online behaviour were also encouraging, showing that several game mechanics can stimulate the desired response from employees.

Thesis Committee:

Chair: Prof. Dr. Ir. Geert-Jan Houben, Faculty EEMCS, TUDelft
University supervisor: Dr. Alessandro Bozzon, Faculty EEMCS, TUDelft
Company supervisor: Drs. Robert-Jan Sips, Center for Advanced Studies, IBM Netherlands
Committee Member: Dr. Martha Larson, Faculty EEMCS, TUDelft
“The larger the island of knowledge, the longer the shoreline of wonder.” — Ralph W. Sockman
Preface

I remember the summer of 2008, walking into the building of what later was going to be the place where I would spend the next four years of my life. While standing in line to submit my application for a Bachelor of Science at the University “Politehnica” of Bucharest, one professor said: “The next four years of your lives are not going to be easy, but none of the things that really matter in life are easy”. This phrase still resonates with me today, while I am about to earn a Master degree in Computer Science from Delft University of Technology. Along the way I have learned that one should not only look for, but also embrace the difficult things in life because those are the ones that bring the most personal satisfaction. Deciding to study abroad was one of the best decisions I took because it shaped my character into becoming the person I am today and taught me to always stand up straight.

Firstly, I would like to thank my parents Virgil and Elena-Gabriela who were always supportive of my decisions and gave me great advice when the right path to take was uncertain. I would not be here without their unconditional moral and financial support. Secondly, I am much obliged to Alessandro Bozzon for guiding me throughout the process of writing this thesis and teaching me how to be a critical and rigorous researcher. A great deal of appreciation goes to Robert-Jan Sips not only for encouraging and providing me with the resources to write this work at IBM, but also for providing instruction along the way. I am thankful to Martha Larson for giving me essential comments and feedback for improving this thesis. Finally, I am so grateful to my girlfriend, Catalina, for her love and support even from far away.

Laurentiu Catalin Stanculescu
Delft, the Netherlands
August 18, 2014
# Contents

## Preface

### Contents

<table>
<thead>
<tr>
<th>1 Introduction</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Problem Statement</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Legal and ethical aspects</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Methodology</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Original contribution</td>
<td>4</td>
</tr>
<tr>
<td>1.5 Thesis outline</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2 Background and Related Work</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 User engagement</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Gamification</td>
<td>9</td>
</tr>
<tr>
<td>2.2.1 Points</td>
<td>9</td>
</tr>
<tr>
<td>2.2.2 Leaderboards</td>
<td>10</td>
</tr>
<tr>
<td>2.2.3 Badges</td>
<td>10</td>
</tr>
<tr>
<td>2.3 Measuring user engagement</td>
<td>10</td>
</tr>
<tr>
<td>2.3.1 Subjective metrics</td>
<td>11</td>
</tr>
<tr>
<td>2.3.2 Objective metrics</td>
<td>11</td>
</tr>
<tr>
<td>2.4 Chapter conclusions</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3 The experimental tool: How much of an IBMer are you</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Overview</td>
<td>13</td>
</tr>
<tr>
<td>3.2 Requirements</td>
<td>14</td>
</tr>
<tr>
<td>3.3 The players</td>
<td>14</td>
</tr>
<tr>
<td>3.4 Player’s actions</td>
<td>15</td>
</tr>
<tr>
<td>3.5 Architecture</td>
<td>15</td>
</tr>
<tr>
<td>3.6 Data model</td>
<td>16</td>
</tr>
<tr>
<td>3.7 Modules</td>
<td>16</td>
</tr>
<tr>
<td>3.7.1 Authentication</td>
<td>16</td>
</tr>
</tbody>
</table>
## CONTENTS

3.7.2 Quiz ................................................................. 18  
3.7.3 Connections ...................................................... 19  
3.7.4 Invitation .......................................................... 20  
3.7.5 Trending topics ..................................................... 20  
3.7.6 Notifications ........................................................ 21  
3.8 Game mechanics ..................................................... 22  
3.8.1 The IBMer Score ................................................... 23  
3.8.2 Leaderboards ........................................................ 24  
3.8.3 Badges ............................................................... 24  
3.9 Implementation ....................................................... 25  
3.10 Building the questions database ................................. 27  
3.11 Chapter conclusions ............................................... 28  
4 Measuring user engagement and social behaviour ................. 29  
4.1 Experimental design ................................................ 29  
4.2 Experimental procedure .......................................... 31  
4.2.1 Recruiting .......................................................... 31  
4.2.2 Procedure .......................................................... 31  
4.3 Objective metrics ..................................................... 31  
4.3.1 The average session length .................................... 31  
4.3.2 The player retention curve ...................................... 31  
4.3.3 The user lifetime .................................................. 32  
4.3.4 The number of invites sent ..................................... 32  
4.3.5 The number of news shares ..................................... 32  
4.4 Subjective metrics .................................................... 32  
4.4.1 Perceived learning ............................................... 32  
4.4.2 Perceived user engagement ..................................... 33  
4.5 Chapter conclusions ............................................... 33  
5 Results and discussion ................................................. 35  
5.1 Overview .................................................................. 35  
5.2 Groups demographics .............................................. 35  
5.3 User engagement ..................................................... 37  
5.3.1 Which game mechanics are more effective for incentivising users to stay longer on the website? ......................... 37  
5.3.2 Which game mechanics are more effective for incentivising users to answer questions? ............................... 39  
5.3.3 Which game mechanics stimulate better perceived user engagement? ........................................... 41  
5.4 Social online behaviour ............................................. 42  
5.4.1 Which game mechanics are more effective for incentivising users to invite their colleagues to play the game? .............. 43  
5.4.2 Do IBMers have strong ties across departments? .......... 44  
5.4.3 Which game mechanics are more effective for incentivising users to share IBM related news on their professional network? 46
## CONTENTS

5.5 Perceived learning ................................................. 47
5.6 Chapter conclusions ............................................ 49

6 Conclusions and Future Work .............................. 51
   6.1 Conclusions ................................................... 51
   6.2 Future work .................................................. 52

Bibliography .......................................................... 53

A Terms of Service .................................................. 57

B The questionnaire ................................................ 59

C The user interface ................................................ 67
Chapter 1

Introduction

Over the last years, companies have taken advantage of the emerging topic of social media for a variety of business purposes. The social network sites (SNS), a particular type of social media, have been used to improve communication and data-sharing in enterprises, leverage social connections within and outside the enterprise and to maintain awareness of organisational dynamics [1, 2].

To take advantage of the opportunities offered by SNSs, companies need to engage their employees with such systems. However, inside the work environment, employees are often too busy with, or overwhelmed by the immediate tasks they have to perform as part of their job, which make it difficult to engage with the long-term vision of the company without proper motivation.

At the same time an increasing number of studies look at how successful different game mechanics are for engaging users with e-commerce systems, health/exercise systems or education/learning platforms[3]. Usage of game mechanics has been correlated with positive patterns in service use, improvement of social interaction and increased user activity and productivity [3]. However, some of these studies present shortcomings such as small sample sizes, absence of control groups, short experiment time frames [3]. It is important to address the pitfalls of the current studies in order to refine the research on gamification.

Enterprise gamification implementations focus on user engagement (42%), brand loyalty (22%) and brand awareness (15%) [4]. By increasing employee engagement, enterprises can achieve important business needs such as motivating and retaining employees, engaging internal communities, increasing adoption, learning and loyalty [4]. According to M2 Research¹, the market of enterprise gamification is expected to jump from $100 M in 2011 to over $2.8 B in 2016, therefore it is important to understand if gamification in an enterprise environment has the same outcomes as presented by studies done in non-enterprise contexts.

⁰Cover page picture from http://www.slideshare.net/gzicherm/maggie-buggie-v2
¹http://m2research.com/gamification.htm
1.1 Problem Statement

We address the problem of implementing gamification in an enterprise environment to support employee engagement and drive online social behaviour of employees in a community. Enterprises face both internal and external challenges. A source of internal challenge is the number of departments: enterprises usually have many business units which produce different types of products and services. It is unrealistic to assume the employees are acquainted with all of them. It is important for these companies to improve employee knowledge about company's history, products and services.

One external challenge companies have is promoting the vision to the more general public. Because of this, enterprises also find it important to motivate the social online behaviour of employees. In the context of the current work, the social online behaviour is defined as (1) establishing connections with other colleagues online and (2) promoting the company's products and services on social or professional networks.

This thesis focuses on applying game mechanics for incentivising people who work within a company network to learn about the company’s products and services, acquire knowledge of technology in general and get better acquainted with their colleagues. Employees are also incentivised to expand their online network and spread awareness of the enterprise’s vision and interests by sharing company news beyond the boundaries of the company. The knowledge acquisition aspect is beneficial for both the enterprise and the employees because the employee is better educated on topics such as the history, the products and the services of the company. The social aspect is also beneficial for both parties because the work environment is improved through building stronger professional connections and learning more about one’s colleagues. It is important to explicitly state the benefits because applying gamification should only be done for the right cause.

The work presented in this thesis has been carried out at IBM, a worldwide enterprise with nine business units and 430,000 employees. Previous research focuses on the impact of social networking at work [5], enriching the social graph using crowdsourcing [6] and the effects of removing gamification from the enterprise SNS [7]. To our best knowledge there are no studies for understanding how different game mechanics impact user engagement and the social online behaviour of employees in an enterprise.

We formulate the main research question of this work as:

**How can game elements be used to support user engagement and drive the online social behaviour of employees in an enterprise environment?**

To answer the main research question it is necessarily to understand what user engagement is, how to trigger and how to measure it. The research objective is to create a flexible and effective software platform for experimenting the impact of different game mechanics on user engagement and social online behaviour. For this, it is necessarily to:

- **Objective 1:** Have a modular implementation of different game mechanics.
- **Objective 2:** Have the means to transfer knowledge to employees.
- **Objective 3:** Have the means to support, access and analyse social connections.
- **Objective 4:** Have the means to support the distribution of information about IBM services and products outside the company.
Objective 5: Have the means to measure and understand what the effect of different
game mechanics on user engagement and social online behaviour is.

Objectives 1-4 closely relate to the design and implementation of the experimental tool
presented in Chapter 3, whereas Objective 5 links to Chapters 4 and 5.

1.2 Legal and ethical aspects

The current work bridges the gap between an employee’s professional and private informa-
tion by combining user data from social and professional networks with enterprise data.
Therefore, it is of a great importance to make sure this work accounts for the best practices
concerning transparency and privacy and is compliant with the Data Protection Act\(^2\).

Specifically, we incorporate these practices by being open with respect to what data the
employees agree to share when they use the system. The Terms of Service presented in
Appendix A can be accessed from the signup page and specify the purpose of the tool, what
kind of data is requested of the employees and how this data will be used. When signing
up, employees explicitly authorise the tool to access their personal information from LinkedIn
and/or Facebook and they have the option to deauthorize and opt out from the system at any
moment. Opting out will cause all the information about an employee associated with
this project to be deleted permanently.

Respecting user's privacy is not only required from a legal and an ethical point of view,
but is also a critical part of system design because trust is a prerequisite for engagement.

1.3 Methodology

We call our software platform “\textit{How much of an IBMer are you}”, a framework for the im-
plementation and deployment of game mechanics within the enterprise. The tool is made
available on the internal network which has the potential of targeting a large number of em-
ployees.

In order to meet Objective 1, we implement a pluggable way to serve different seam-
lessly integrated game mechanics. The mechanics we experiment with are: a general score
to summarise the user's performance in the game, feedback on answering quiz questions,
leaderboards and badges.

Objective 2 is achieved by a set of quizzes that test employees' knowledge on topics
such as facts about IBM, facts about technology in general and colleagues' skills. Feedback
is given to the player to support learning.

We satisfy Objective 3 by designing the tool as a hub for social connections. By using
information from social networks such as LinkedIn, Facebook and IBM Connections\(^3\) (the
enterprise social network) an employee can see all his connections in a single place. The
information from these networks is then combined with information from the game. There is
also an invitation system for supporting establishment of new connections.

\(^2\)http://ec.europa.eu/social/BlobServlet?docId=2507&langId=en
\(^3\)http://w3-connections.ibm.com
1.4 Original contribution

We accomplish Objective 4 by employing a news module that periodically fetches trending topics about IBM and displays them in the tool with the option to distribute them on the user’s social networks.

Finally, we meet Objective 5 by logging different parameters regarding the user’s interaction with the tool. Such parameters include information requests, number of actions of a type, total time spent on the website or number of invites sent. We combine objective metrics with a questionnaire to also assess users’ perception. To study the effect of game mechanics on user engagement and social online behaviour of employees, we divide our user base in user groups and deploy different game mechanics to different groups. At the end, we analyse the data and draw conclusions.

1.4 Original contribution

This thesis’s main contribution is the study of gamification in an enterprise context as a means of supporting user engagement and driving the social online behaviour of employees.

One contribution as part of the current work is a modular framework for studying gamification, which separates the game mechanics logic from the actions logic. Different game mechanics can be plugged into the system and integrated with the overall experience in a way transparent to the user. Designing a framework for studying gamification in an enterprise environment is not trivial as it requires a tremendous amount of work to properly design and implement the user experience, while taking into account strict security requirements.

A second contribution of this thesis is the analysis of the impact leaderboards, badges and the combination of the two on user engagement and the social online behaviour of employees in an enterprise. The analysis is carried out on both objective and subjective metrics.

A third contribution is the database of quiz questions. The number and quality of the questions is essential for engagement and therefore building such a database is a prerequisite. The resulting database contains questions to support a large number of games and is also specific enough to target our reference group.

1.5 Thesis outline

The thesis is organised as follows. We first describe background and related work in Chapter 2 where we define user engagement and the metrics for measuring it. We present, in the same chapter, gamification as a means of achieving engagement. Chapter 3 focuses on the design and implementation of the experimental tool “How much of an IBMer are you”. Chapter 4 focuses on describing the experimental design and the metrics used in the experiments. We summarise the results of the various experiments we conducted in Chapter 5. Finally, we formulate conclusions and state proposals for future work in Chapter 6.
Chapter 2

Background and Related Work

This chapter discusses two related topics, of great importance for this thesis: user engagement and gamification. First, we define user engagement from literature and present the different attributes. Second, we introduce gamification as a means for triggering user engagement. Third, we look at subjective and objective metrics for measuring user engagement.

2.1 User engagement

User engagement is, according to [8], “a category of user experience characterised by attributes of challenge, positive affect, endurability, aesthetic and sensory appeal, attention, feedback, variety/novelty, interactivity, and perceived user control”. In [9], user engagement is defined as “the emotional, cognitive and behavioural connection that exists, at any point in time and possibly over time, between a user and a resource” and is characterised by focused attention, positive affect, aesthetics, endurability, novelty, reputation, trust and expectation and user context, incentives and benefits. These two definitions agree on the multi-faceted aspect of user engagement and identify common attributes such as: positive affect, endurability, aesthetics, focused attention and novelty. The perceived user control attribute in [8] is contained in the reputation, trust and expectation attribute from [9] and the attribute of challenge in the first definition is incorporated in user context in the second definition.

The effect of a single attribute on user engagement is dependent on the application. Trust might be a prerequisite for engagement with a health service, but aesthetics might count more for using a movie service. When designing applications and user interfaces context is of great importance. The current work uses the broader definition of user engagement as outlined by Attfield, Simon, et al [9]. The remaining of this section describes the attributes of user engagement. Figure 2.1 illustrates the attributes of user engagement and emphasises the role of gamification as a way of triggering user engagement.

Focused attention

*Focused attention* is one characteristic of engagement and it involves excluding other things not related to the particular experience and a distorted perception of time during interaction.
2.1 User engagement

Figure 2.1: An overview of the attributes of user engagement with an emphasis on gamification as a means for triggering and supporting engagement.

The more engaged someone is with an activity, the more likely they will underestimate the time passed during that activity.

**Positive affect**

Another attribute of user engagement is positive affect. Being engaged in an experience means to manifesting positive emotions during the interaction. A user who was initially affectively hooked will show a desire for exploration and active discovery. This desire will contribute to emotional involvement and customer loyalty.

**Aesthetics**

*Aesthetics* is an important factor for engagement and refers to how visually appealing is an interface to the user. It closely relates to positive affect and it is known to stimulate curiosity. It is often related to design principles such as screen resolution and graphics.
Endurability

*Endurability* refers to the likelihood of remembering and the willingness to repeat an experience. Users can be encouraged to return if they are conveniently rewarded either with incentives or/and with interesting content. Endurability also relates to how likely is a user to be willing to recommend an experience to others.

Novelty

*Novelty* is another characteristic of user engagement. A novel experience is unfamiliar, unexpected or surprising and it appeals to one’s sense of curiosity. Novelty can be created though freshness of content, variety or innovation. However, to properly support engagement it is important to preserve a balance between novelty and familiarity.

Trust

*Trust* is a necessary condition for engagement. It can be seen as a insurance offered by technology and it is influenced by the reputation of the accessed service.

User’s context

User’s context is a combination of user’s motivation, incentives and benefits that have an effect on the overall experience. Motivation is a precondition for action and thus is important when studying user engagement. To be motivated means to be moved to do something. The Self-Determination Theory is based on “the hypothesis that there is a set of universal psychological needs that must be satisfied for effective functioning and psychological health” [10]. The SDT makes a clear distinction between intrinsic and extrinsic motivation based on the reasons for taking a particular action. According to the theory, being intrinsically motivated means performing an activity because is interesting and enjoyable, while being extrinsically motivated means to perform an activity for an outcome different from the activity itself. We classify the methods reported in the literature for engaging users with technology based on the distinction between the two types of motivation.

Extrinsic motivation refers to performing an activity to achieve a separate goal from the activity itself. In [11] the authors classify extrinsic information into: external regulation, introjection, identification and integration. The *external regulation* refers to behaviours motivated by an external reward or punishment. The *introjected regulation* involves the ego and is characterised by a person's behaviour being motivated by the need to enhance or maintain self-esteem and the feeling of worth. Another type of extrinsic motivation is *regulation through identification* which is a self-determined motivation due to the person identifying the personal importance of a behaviour and thus accepting the self-imposed regulations. The last, and most autonomous form of extrinsic motivation is *integrated regulation*, when the regulations are integrated to the self. This form of motivation is autonomous and unconflicted, similar to intrinsic motivation, with the important distinction that the behaviour is still motivated by an outcome separate from the activity itself.
An intrinsically motivated behaviour is performed for fun or by the challenge of performing an activity rather than by external rewards or pressures [11]. Intrinsic motivation can be defined in terms of the task being interesting or in terms of the satisfaction someone gains from engaging with the task. Intrinsic motivation closely links to the core needs defined by Deci and Ryan [12] as autonomy, relatedness and competence. Autonomy refers to the actions “for which people feel a full sense of choice and endorsement of an activity” [10], therefore it is important for people to have full control over their actions and decisions in order to be intrinsically motivated. Relatedness is defined as “the need to to feel belongingness and connectedness with others” [12]. Competence is a sense of efficacy when people feel they have the skills and competence to perform a task. Competence can be triggered by rewards and proper feedback.

The incentives of performing an action or showing a certain type of behaviour can be monetary or non-monetary. Monetary incentives are used in human computer interaction to motivate users to perform a desired action in return for financial benefits. By employing monetary incentives, extrinsic motivation is increased. However, because monetary incentives have been show to decrease intrinsic motivation [13], many aspects of the application have to be taken into account when deciding to use such incentives.

We present the most frequent monetary incentives used in literature [14]. Quota systems and discretionary bonuses are a form of monetary incentives where an agent is rewarded for achieving a certain performance level for the first time. Close to achieving the performance level, the motivation in significantly increased, but the effort level quickly drops after the evaluation.

Deferred compensation is another example of a monetary incentive and it is used to motivate people to perform complex and long-lasting tasks. It is similar to quota systems in that performance is evaluated at different points in time but with the subtle difference that compensation takes into account three points in time: $t_0$, $t_1$, and $t_2$. “At $t_0$ the agent is promised a reward for successfully passing a deferred compensation at $t_2$. The evaluation takes into account the period of time $[t_1,t_2]$.”

Relative evaluation is used as a monetary incentive in groups where members are evaluated with respect to the other members when the reward is money. In team-based compensation, an entire team is rewarded for showing a certain behaviour and it is used for cooperative-tasks or when evaluating a single individual is not possible.

Non-monetary incentives are the ones that do not involve financial compensation such as psychological and gamification incentives. Psychological incentives are the ones that stimulate competition and personal satisfaction and appeal usually to the intrinsic motivation of the user. The incentives triggered by gamification appeal to both the intrinsic and extrinsic motivation.

This section defined user engagement and described it’s different attributes. Having defined and explained the different attributes of engagement, we now turn to gamification, which represents a means of triggering and supporting user engagement.
2.2 Gamification

Gamification is defined as “the use of game design elements in non-game contexts” [15]. In this definition, “game” refers to the presence of rule systems and the competition between the actors to achieve a set of discrete goals. The “elements” are those found in games and play a significant role in gameplay. The designer of a gamified system borrows design elements from games but does not implement a full ‘game proper’ [15]. Finally, the “non-game contexts” are concerned with use cases which do not have to do with entertaining, but rather with engagement and improvement of the user experience in general.

Gamification represents a means of supporting user engagement and positive patterns in service use, but these effects depend heavily on the context in which is implemented as well as on the users it targets [3]. Deploying gamified systems promises “access to more ecologically valid user data on the different kinds of experiences [...] that arise from the interaction with these systems” [15]. Increased engagement results from the intrinsically motivating, gamefull experiences. In gamified systems, the developer of that system creates goals or challenges with specific winning condition. Users are rewarded when they achieve one of the goals.

We present the most common game elements used in gamification, identified by a literature review on peer-reviewed empirical studies on gamification [3].

2.2.1 Points

Points represent an important concept in gamification. Points indicate to the users that their actions are valued. Gamified systems use one of the five points systems [16], namely: experience points, redeemable points, skill points, karma points and reputation points.

Experience points are earned for taking different actions in the system and are a way for the designer to align his behavioural objectives with the player. Although experience points can only increase, some designs specify an expiration time for these points to create goal loops.

Redeemable points, as the name suggests, can be traded up for things. These points are sometimes referred to as coins and are usually used in virtual economy systems. Because they are traded out, there are legal constrains to consider when implementing them.

Skill points are assigned to specific activities within the game and usually reward alternate tasks and subgoals. They represent a way to steer a player towards goals which are not the main focus of the designer but are still important.

Karma points are points that do not bring any value to the user except for when the user shares them. These points create a behavioural path for altruism and personal reward.

Finally, the reputation points make up the most complex system. Such points are used in systems where trust between one or more parties is required but cannot be guaranteed. Reputation points are used as a proxy for trust and the complexity lies in the facts that some users will try to game the system.
2.2.2 Leaderboards

Leaderboards represent a performance comparison game element. They are employed in both single and multiplayer contexts and are used to map progress and move the players to take action, appealing to their sense of competition. In a single context leaderboards are used to compare a user’s latest score to his previous scores, while in a multiplayer context they are used to compare the player’s performance with other players in the game.

In a multiplayer context, a leaderboard can compare the performance of a player with the performance of all the other players in the game. Such a leaderboard is called an absolute leaderboard and can act as a demotivator for those users who rank low, since there is the risk that they will perceive too much of an effort to reach one of the top leaderboard positions.

The relative leaderboards try to overcome this by only comparing a player with his friends or showing the player in the centre of the leaderboards with a few other players above and below him. The public character of a leaderboard presents a unique challenge in terms of privacy. When sensitive or private information is displayed designing leaderboards calls for extra awareness [16].

2.2.3 Badges

Badges represent at a first glance a visual representation of user’s achievements “marking the completion of goals and the steady progress of play within the system”[17]. However, while used as a way of engaging and motivating users, they serve five psychological functions: goal setting, instruction, reputation, status/affirmation and group identification [18].

The goal setting function is met by challenging the users to achieve the goals that are set for them. However, to be an effective motivator, these goals should be in comfortable reach. Badges also provide instruction because they exemplify the types of activities and interactions that are desired by the system’s designer. Since badges show user’s achievements they are a valuable encapsulation of user’s interests, skills sets and past experience on which reputation is built.

Status/affirmation is accomplished by reminding the user of his past accomplishments and communicate those accomplishments to others without bragging. Finally, badges promote group identification through the perception of similarity between the individual and the group.

Badges have been proven to drive user behaviour [17] on StackOverflow, motivate check-ins on FourSquare [19] and seem to encourage desired study practices in TRAKLA2 [20].

This section defined gamification as a means of triggering and supporting user engagement and presented the most common game elements in the literature. Next section is concerned with metrics for user engagement since it is important to be able to recognise when user engagement has been achieved.

2.3 Measuring user engagement

Having defined user engagement, we now have to understand how to measure it. User engagement can be measured by subjective or/and objective metrics.
2.3 Measuring user engagement

### 2.3.1 Subjective metrics

The subjective metrics involve surveys to capture the details about that specific interaction and the user. Such an example is [21] where the authors constructed a multidimensional scale to measure user engagement and assessed its reliability and validity. After identifying six attributes of user engagement (perceived usability, aesthetics, focused attention, felt involvement, novelty, and endurability), the authors run a principal component analysis on a list of more than 400 items which refer to one of the attributes of engagement to identify the most parsimonious set of items.

### 2.3.2 Objective metrics

The objective metrics overcome drawbacks of subjective metrics such as: halo effect, leniency or severity, central tendency and restriction of range and interrater reliability or agreement [22]. The halo effect represents: (a) “a tendency to attend to a global impression of each ratee rather than to carefully distinguish among the levels of different performance dimensions”; (b) “a rater’s inability or unwillingness to distinguish among the dimensions of a ratee’s job behaviour”; (c) “a tendency to place a given ratee at the same level on different dimensions” [22]. The leniency or severity refers to: (a) “a tendency to assign a higher or lower rating to an individual than is warranted by that ratee’s behaviour”; (b) “a response set attributed to easy or hard-nosed raters whose rating are consistently higher or lower than is warranted, given some external criterion of known true performance level”; (c) a shift in mean ratings from the midpoint of the rating scale in the favourable or unfavourable direction”; (d) “a rating level effect” [22]. The central tendency and restriction of range is a third rating error which is defined as a “rater’s unwillingness to go out on the proverbial limb in either the favourable or unfavourable direction” [22]. Interrater reliability or agreement refers “to the extent to which two or more raters independently provide similar ratings on given aspects of the same individuals’ behaviours” [22].

The subjective perception of time is measured by asking the user to estimate how much time has passed during the activity. Subjective perception of time is an objective measure since it is assessed against the time.

Follow-on task performance is another way of measuring engagement. This measure is concerned with how well somebody performs on a different task immediately after the user used the system for which we want to measure engagement. When playing games, the more engaged the user is, the more it takes them to complete an unrelated puzzle task [9].

Sensors are another source of objective measurements. Such sensors are eye tracking sensors, mouse pressure sensors or biosensors.

The player retention curve is a useful indicator of the proportion of the players who revisited the game and the frequency of their visits. It has been used to assess the enjoyability and popularity of human computations games [23].

The user lifetime represents the number of correct and incorrect questions answered by a user. It has been used to understand the effects of various engagement incentives in a gamified crowdsourcing system [24].
The online behaviour of a user will give an indication of the depth of engagement with a site. Examples of online behaviour metrics are the number of comments per post, number of subscriptions to feeds or search engine patterns [9].

2.4 Chapter conclusions

This chapter focussed on user engagement definition and metrics. The chapter also stresses the importance of gamification as a medium to trigger and support user engagement. Some key attributes of engagement will come back in Chapter 3 when we describe design decisions in our experimental tool. The most common game elements used in gamified systems presented in this chapter will also be further described in Chapter 3. The metrics used for measuring engagement will come along in Chapter 5 when we describe the experimental design.
Chapter 3

The experimental tool: How much of an IBMer are you

One of the main contributions of this thesis is the design and implementation of the game “How much of an IBMer are you”. The tool plays a central role in this thesis because it supports the research of the effects of game mechanics on user engagement and the social online behaviour of employees. We first give an overview of the tool with an emphasis on the architecture, present the different modules, present the players and describe the game mechanics we implemented and the motivation for doing so. Lastly, we describe the process of building the quiz questions database.

Also essential to this thesis is the experimental design in Chapter 4 and the analysis of the information collected when the game was played by the employees in Chapter 5.

3.1 Overview

The tool “How much of an IBMer are you” is an online website which can be accessed by all IBM employees behind the company’s firewall. The tool provides a quiz game with questions from three categories: IBM facts, World Wide Technology and You & Your Network. The users score points by answering questions and can receive badges for reaching different milestones.

Another important feature is the social hub which allows an employee to see his connections from social networks (LinkedIn, Facebook and IBM Connections) in a single location. The list of connections is combined with information from the game such as score, earned badges or current position on the leaderboard. To support the formation of new connections we build an invitation system. A user can invite his connections to play the game and receive points and/or badges for doing so.

The tool also provides employees with trending topics to keep them up to date with what the company is doing. The users are rewarded for distributing these news articles on their social networks.

The modular design allows to easily plug other game mechanics in, making this an essential tool for studying gamification. The actions to which we want to apply gamification can also be plugged in and out according to the requirements.
3.2 Requirements

The tool expands beyond the IBM use case. Being extensible, it can easily accommodate other types of actions users are incentivised to take. The plug-and-play nature of the game mechanics modules makes it easy to expand the research to other gamification elements which are not extensively explored in literature. Having such a tool can lay a strong foundation to study the effect on game mechanics in a variety of applications and contexts, while providing the right framing for doing so.

3.2 Requirements

Designing a framework for studying gamification in an enterprise environment is not trivial. There is a tremendous amount of work needed to properly design and implement the user experience. User engagement, triggered through the use of gamification, imposes some important requirements:

1. It requires an appropriately designed user experience balanced between familiarity and novelty.
2. The user interface has to be visually appealing, clean and simple because in an enterprise context time is crucial.
3. The incentive mechanisms have to be fully integrated with the overall experience.
4. There has to be transparency in the way the user’s data is used and ensure an employee can opt out from the tool at any point.

A poorly designed tool would not be compatible with studying gamification and therefore putting so much emphasis on the overall user experience is a prerequisite.

The scope and context of the application strongly demands to stress test the tool. This means we analysed performance metrics under high usage with many concurrent connections and real traffic to understand if there are bottlenecks. Based on the results we tuned the tool to improve it’s performance.

The complexity of the tool called for extra attention for beta-testing. We had a few iterations between development and testing to ensure the quality of the user’s experience.

An additional requirement for studying the effect of multiple game mechanics is to have the means to dynamically enable modules. This is needed to support multi versions of the tool based on logged in user. To implement this requirement we designed the tool around a Routing controller which will be presented in Section 3.5.

3.3 The players

The players of our game are the IBM employees. Worldwide, IBM employs around 430.000 people in 170 countries. They are spread across nine business units and have a large variation of skills and interests. Having potentially access to such a large number of employees makes IBM an ideal place for studying gamification in an enterprise context.
3.4 Player’s actions

The set of actions any player can take in the game are the following: answer quiz questions, link social networks, see their connections, their score and invite those colleagues not in the game, view trending topics and share these news to social or professional networks. Depending on the experimental group, a user’s set of actions might contain viewing leaderboards and/or badges.

3.5 Architecture

The framework we designed and implemented for supporting the work in this thesis has a modular design where user’s actions and game mechanics are separated. One of the key components is the Routing controller which servers the functions of asserting user’s permissions to load a resource and loads the resources accordingly. This module is linked to the Authentication module, the actions modules Quiz and Trending topics and to the game mechanics modules Leaderboards and Badges.

The Authentication Module, as the name suggests, handles user’s authentication with social providers such as LinkedIn and Facebook as well as the corporate IBM Connections. Another important module is the Database Controller which implements the methods for connecting and writing to and reading from the database.

The modules that deal with user’s actions or game mechanics are part of our instantiation of the framework and will be explained later in this chapter.
3.6 Data model

Having presented the architecture of the tool, we now concentrate on describing the data model which served as a blueprint for the implementation of the database model. The most important entities are: user, question, social questions, invitations, peers, earned badges and ibmer score.

The user entity stores information about the different users in the system such as name, job role and picture. It represents a central entity because many of the other entities implement a foreign key to it. The question and social questions entities store the questions used in the quiz together with the answer options and correct answer. Combined with the answer log and social answers log they allow our algorithms to select, when possible, unique questions for every user. The peers stores all users’ connections from the different social networks. For each peer the name, job role, profile picture url and source network are saved. The invitations entity stores information about which invites has been sent to whom and implements a foreign key to the peers table. For each invitation we store the inviter’s email address, the invitee’s email address, the time when the invitation was sent and whether the user accepted the invitation or not. The earned badges summarise user’s achievements. For each achievement we store the user id, badge id and the time at which the achievement was earned. The ibmer score stores the user’s daily IBMer score. The logical representation of the database schema is represented in Figure 3.2.

3.7 Modules

We split the functionality into modules for maintainability and extensibility reasons. Every module is composed of a controller to implement the logic and a view to present the information to the user.

3.7.1 Authentication

The Authentication module is in charge of handling registration and authentication. The module uses the HybridOAuth¹ library to implement social authentication and the logic for communicating with the different social providers.

The registration is a two-step process. The user has to first authenticate with LinkedIn and then provide the IBM email address. The choice for LinkedIn was motivated by the following reasons. LinkedIn does not only provide a way of identifying a user, but also a list of connections which can be used to build the social network in the game. We recognise the choice for the default authentication mechanism might limit the number of players since not every IBM employee has a LinkedIn account. To further investigate this, we used the LinkedIn search engine to look for the number of users that report to work for IBM. We find 386,000 results (90%) of the employees have a LinkedIn account. Authenticating with LinkedIn is effortless for the user because there is no need of filling in a user name and a password every time and is secure because no login information is stored in the local database.

¹http://hybridauth.sourceforge.net
The experimental tool: How much of an IBMer are you

Figure 3.2: The logical representation of the database schema.m
3.7 Modules

The experimental tool: How much of an IBMer are you

Figure 3.3: a) The initial login page. When a user clicks on Connect with LinkedIn he will be redirected to LinkedIn to complete his login credentials if he is not already authenticated with LinkedIn. b) The second step of the signup process is illustrated. In this example, the email address has already been identified by the system and the user has to confirm it.

The second step in the registration process is to provide the IBM email address. To make this process as easy as possible, we use the Faces API\(^2\), a platform for quickly find IBMers, to automatically find the email address of the player based on the name. When several possible emails are identified, the player has to confirm the correct one. The work email is required for the invitation system which will be described later in this section, but also serves the purpose of promoting trust.

Identification information is stored in the database. Figure 3.3 represents the view associated with the authentication process.

3.7.2 Quiz

At the core of this tool is module for answering quiz questions. The goals of this module are to retrieve questions from the database, process the answers given by a player, provide feedback back to the player on the correctness of the answers and to log the answers in the database. The quiz module closely relates to Objective 2 of this work: “have the means to transfer knowledge to employees”. The user can choose between three categories:

- **IBM Facts** questions are related IBM history, technology and facts.
- **World Wide Technology** questions touch upon knowledge about information technology and technology in general.
- **You & Your Network** questions revolve around user’s connections headlines and skills.

An example question in IBM facts is: “In which two consecutive years, IBM researchers won the Nobel Prize in Physics?”. An example question in World Wide Technology is: “At any given time, how many GPS satellites are visible in the sky from any location on Earth?”. An example of a social question is: “Which of your peers is Research Intern at IBM?".

Players interact with the module by initiating game sessions, each featuring 10 rounds. Each round contains a question to answer. Players can abandon the game at any point. Before a question is served, the player will have to chose the category. On the question page the question category, text and answers are displayed together with a timer indicating how much time is left for answering the question. The user has 20 seconds to choose an answer,

\(^2\)http://faces.tap.ibm.com
The experimental tool: How much of an IBMer are you

3.7 Modules

Figure 3.4: a) On the choose category page the user selects the category of the next question. b) The feedback a player receives when answering a question in IBM Facts. In this example the user has chosen the wrong answer and has been highlighted in red. The correct answer is shown in green.

After which the system will show the correct answer. The time constraint introduced by the timer puts time pressure on the players which in turn makes the questions more challenging. The timer also lowers the chances of someone cheating by looking up the questions’ answers. At any point during a game, the player is informed about the number of questions answered and the number of remaining questions. After the 10th question, a summary page with the new score, percentage of correct questions and the average answering time is shown.

There are three mechanisms for selecting a question for the current user: random not seen, random wrong question and random seen.

Random not seen makes sure a user receives a distinct question as long as the pool of available questions has not been exhausted.

Random wrong answer will select, when random not seen questions are not available anymore, one of the previous questions for which the user gave a wrong answer.

Random seen will select a random question without looking at user’s question history and only when the other two mechanisms fail.

Figure 3.4 shows the view associated with this module. In 3.4 it is illustrated a) the view for selecting the category for the next question and b) the feedback a player receives after answering a question in IBM facts.

The system provides feedback every time a user answers a quiz question specifying if the answer provided was correct or incorrect. The correct answer is always mentioned. Being told the correct answer is more effective for learning than only giving information regarding the correctness of a response. In our tool, a correct answer is indicated by highlighting the corresponding text in green, while an incorrect answer is indicated by highlighting the answer in red.

3.7.3 Connections

The tool How much of an IBMer are you aims at bringing all player’s connection to a single place and enrich the information available from social networks with information from the game. The module links back to Objective 3 of this work: “have the means to support, access and analyse social connections”.
3.7 Modules

The experimental tool: How much of an IBMer are you

Figure 3.5: The Connections view displays all users’ connections on LinkedIn and Facebook who work for IBM. In this example the user decided to search for Robert. The system found two connections with that name. The connection from LinkedIn is already in the game and has an IBMer score of 242.

The module Connections serves the purpose of identifying IBMers connected to the current user on social networks. Currently, only LinkedIn and Facebook networks have been implemented, but the tool can be easily extended to accommodate other social networks. We consider an IBMer a connection who has reported on their social network working for IBM in the present. For each peer we retrieve from the social network the id, name, profile picture url, headline and source network which we store in the database. The connections are then displayed to the user grouped by source network. For each connection we display a picture, the job role at IBM, name, the score in the game if available and whether an invitation to play the game has been sent to that user. To quickly find a connection we implement a name filter which dynamically updates the list of connections. Figure 3.5 shows an example.

3.7.4 Invitation

The Invitation module allows a player to invite his colleagues to play the game. Every IBMer who is a connection to the current player can be invited. The email address of the invited user is automatically detected using the Faces API, an internal IBM tool. When an invitation is sent, the system will send an email which appears to originate from the inviter and is addressed to the invitee. In the email we display the profile picture, the IBMer score of the inviter, a personal message if provided and a list of other players in the game with their corresponding score. The invitee can accept to play the game by following the link in the email. This link is personalised to allow us to track when a user has accepted the invitation. Figure 3.6 shows an example.

3.7.5 Trending topics

The Trending topics module is responsible for fetching IBM related news and displaying them to the user. The motivation behind this module is to create an environment where employees can stay up to date with the trending topics within IBM and promote at the same time
The experimental tool: How much of an IBMer are you

3.7 Modules

The company’s products and services externally. The module represents a way to achieve Objective 4 of this thesis: “have the means to support the distribution of information about IBM services and products outside the company”.

The news are fetched from an IBM RSS feed every two days and saved in a local database. The two days interval was chosen because the RSS feed we used is rarely refreshed daily. When displayed, only the most recent 6 news articles from the database are shown to keep the content fresh. Every article has two buttons for sharing the information on social networks while also displaying the number of times an article has been shared on different social networks. Figure 3.7 shows the trending topics view.

The number of shared news items is checked with a cronjob\(^3\). Only the users who visited the Trending topics page are considered, to avoid potentially reaching the API limitations. We retrieve the news feed from LinkedIn and Facebook and look for IBM related links. We recognise this approach is more resource intensive than incrementing the shared news count when the share button is clicked on the website by a user. The later approach could give false information if the user shares an article and then deletes it while looking at the news feed will always give the correct number of shares.

### 3.7.6 Notifications

The Notifications module serves the purpose of notifying a user about different events. This module is of utmost importance because it closes the engagement loop, by showing visual feedback to users on their performance in the game, achievements and successfully completed actions.

The notifications are pulled out from a queue and displayed on the website. They get dismissed by moving the mouse cursor over them or clicking the close button. The notifications are triggered by the different modules when an event takes place. Examples of events

---

\(^3\)A cronjob is a a time-based job scheduler in Unix-like computer operating systems
3.8 Game mechanics

The experimental tool: How much of an IBMer are you

Figure 3.7: The trending topics page displays information related to IBM. Every news article has a title, a summary and buttons for distributing the article on user’s social networks. In the example, the first news article has been shared 177 times on LinkedIn and 5 times on Facebook. The share count represents the total number of shares on the social network and is not a measure of the popularity of the article in the tool.

Figure 3.8: The figure shows how notifications are displayed. In this example, upon visiting the website the information that the user earned a new badge and that a connection accepted the invitation are displayed.

that issue a notification are: the IBMer score increases or decreases, an invitation has been accepted, the number of news shared has increased or decreased or a new badge has been rewarded. Figure 3.8 shows an example.

3.8 Game mechanics

The tool How much of an IBMer are you provides a framework for implementing and deploying different game mechanics.
3.8.1 The IBMer Score

The score used in the tool is called The IBMer Score. Exploring the existing literature on game mechanics, it become apparent to us that there are no references for building a scoring system. The score itself brings a lot of power to many games and for this reason most developers keep it as a secret. Because of this, we developed our own scoring scheme.

The IBMer Score summarises the players’ actions in the game with an integer greater than zero and represents an incentive for players to perform balanced actions in the tool. It is displayed next to the profile picture of the player, similar to other services and is visible to both the player and the player’s connections in the game. The actions that influence the score are: question answering, inviting colleagues to play the game, sharing news and connecting multiple social networks. To motivate the players to return to play we introduce a discount factor. Equation (3.1) illustrates how the score is computed at time $t$ without discounting. The term $qa_{score}(t)$ (3.3) represents the score corresponding to answering questions, $ui_{score}(t)$ (3.4) the score for inviting colleagues to play and $ns_{score}(t)$ (3.5) is the score for sharing news. $mIBM(t)$ (3.6) represents an initial score the user receives based on the number of connected networks. Equation (3.2) shows how the final score is computed with the discount factor incorporated. Although we put a lot of effort into designing the score, we didn’t carry any simulations to optimise the formula, as the optimisation is beyond the scope of this thesis.

\begin{equation}
IBM_1(t) = qa_{score}(t) + ui_{score}(t) + ns_{score}(t) + mIBM(t) \tag{3.1}
\end{equation}

\begin{equation}
IBM(t) = \begin{cases} 
\text{avg}(\omega \ast IBM(t-1) + IBM_1(t)) + mIBM & \text{if } IBM_1(t) \leq IBM(t-1) \\
\text{avg}(IBM(t01) + \omega \ast IBM_1(t)) + mIBM & \text{if } IBM_1(t) > IBM(t-1)
\end{cases} \tag{3.2}
\end{equation}

\begin{equation}
qa_{score}(t) = \begin{cases} 
\alpha \ast \frac{ca}{mq} & \text{if } \alpha \ast \frac{ca}{mq} \leq 0.9 \\
0.9 + \log^2(\alpha \ast \frac{ca}{mq}) & \text{if } \alpha \ast \frac{ca}{mq} > 0.9
\end{cases} \tag{3.3}
\end{equation}

\begin{equation}
ui_{score}(t) = \begin{cases} 
\beta \ast \frac{mi}{ui} & \text{if } \beta \ast \frac{mi}{ui} \leq 0.9 \\
0.9 + \log^2(\beta \ast \frac{mi}{ui}) & \text{if } \beta \ast \frac{mi}{ui} > 0.9
\end{cases} \tag{3.4}
\end{equation}

\begin{equation}
ns_{score}(t) = \begin{cases} 
\gamma \ast \frac{ms}{us} & \text{if } \gamma \ast \frac{ms}{us} \leq 0.9 \\
0.9 + \log^2(\gamma \ast \frac{ms}{us}) & \text{if } \gamma \ast \frac{ms}{us} > 0.9
\end{cases} \tag{3.5}
\end{equation}

\begin{equation}
mIBM(t) = \eta \ast \frac{cn}{ln} \tag{3.6}
\end{equation}

\footnote{http://www.klout.com}
### 3.8 Game mechanics

The experimental tool: How much of an IBMer are you

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>question answering weight</td>
<td>0.4</td>
</tr>
<tr>
<td>$\beta$</td>
<td>invite peers weight</td>
<td>0.2</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>share news weight</td>
<td>0.2</td>
</tr>
<tr>
<td>$\eta$</td>
<td>connected networks weight</td>
<td>0.2</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>discounted variable for number of consecutives days played</td>
<td>1-1.1</td>
</tr>
<tr>
<td>$ca$</td>
<td>correct answers given by this user</td>
<td>-</td>
</tr>
<tr>
<td>$aq$</td>
<td>number of attempted questions by this user</td>
<td>-</td>
</tr>
<tr>
<td>$mq$</td>
<td>mean number of attempted questions / user</td>
<td>10</td>
</tr>
<tr>
<td>$u$</td>
<td>users invited by the user</td>
<td>-</td>
</tr>
<tr>
<td>$m$</td>
<td>mean number of invites / user</td>
<td>5</td>
</tr>
<tr>
<td>$cn$</td>
<td>number of connected networks</td>
<td>2</td>
</tr>
<tr>
<td>$tn$</td>
<td>total number of implemented networks</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3.1: The list of parameters used for computing the IBMer score.

The individual terms in equations (3.3), (3.4), (3.5) are computed using a log function when the value is greater than 0.9. This is to limit the effect an action has on the score. If a player takes many actions of the same type, for example, answering only questions, the score will increase with more questions played but the increase will become smaller and smaller with every new question answered. A complete description of all the variables used in computing the IBMer score and their meaning can be found in table 3.1. When the value is not provided, it means it is computed dynamically.

#### 3.8.2 Leaderboards

We also implemented and experimented with leaderboards. Leaderboards display the ranking of a user compared to other users. It has been suggested that leaderboards can decrease motivation when they include all the users. This is because low ranked users perceive the effort of reaching the top too much and not worth it. To overcome this, we implemented relative leaderboards where a user’s rank is compared to its peers only. The peers represent those employees the player is connected to on social networks. The players are ranked based on their IBMer score. Our hypothesis was that the users will strive to reach one of the top leaderboards positions by changing their behaviour. An example of a leaderboards is shown in Figure 3.9.

#### 3.8.3 Badges

We also considered badges as a way of incentivising user’s activity. Badges are rewarded to a user for performing one or multiple actions of a given type and they serve as a summary of a user’s accomplishments. Our hypothesis is that badges can increase participation and influence user behaviour on a website. We reward users for multiple actions:
Providing their email address: When signing up, people receive the Proud to be an IBMer badge for successfully providing their email address.

Linking social networks: Players also receive the Social as it can get badge for linking their Facebook account.

Answering questions: When answering questions, players earn multiple badges for first correct answer in a category, or multiple (consecutive) answers.

Inviting peers: Players can also receive badges when one or a number of colleagues accept the invitation to play the game.

Sharing news: When sharing news, players are also rewarded. Badges are earned for sharing the first, five or twenty five news articles on LinkedIn or Facebook.

The three main categories of actions for which we reward badges are: social actions, participation and answers. In the tool, we show the list of available badges, the list of earned badges and the progress towards earning a new badges. We do so because knowing how many actions of a type need to be taken to achieve a new badge act as an effective motivator. To provide additional incentive, we also inform the user how many people have earned a specific badge. An example is shown in Figure 3.10.

More information on the badges used in the game How much of an IBMer are you can be found in Table 3.2.

### 3.9 Implementation

*How much of an IBMer are you* is a web application. The backend side was written in plain PHP and makes usage of HybridAuth, a library for implementing the social sign on. The code is modular to make the application easy to maintain and extend. We designed a routing controller which is in charge of deciding which modules needs to be loaded and whether a resource should be made available to the user. To give an example, if a user requests to play the game, the routing controller will first check if the user is authenticated. If the condition is met, the controller will load the necessary module, in this case the Quiz module. Both the...
3.9 Implementation

The experimental tool: How much of an IBMer are you

Figure 3.10: The Badge view shows a summary of user’s achievements. In this example, the user earned 4 of the 8 possible badges from the Social category. For the badges that were not earned yet a progress bar is displayed.

<table>
<thead>
<tr>
<th>Type</th>
<th>Action</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social actions</td>
<td>Link your IBM email</td>
<td>Proud to be an IBMer</td>
</tr>
<tr>
<td>Social actions</td>
<td>Link your Facebook account</td>
<td>Social as it can get</td>
</tr>
<tr>
<td>Social actions</td>
<td>Invite your first friend</td>
<td>Getting friendly</td>
</tr>
<tr>
<td>Social actions</td>
<td>Invite 5 friends</td>
<td>Networker</td>
</tr>
<tr>
<td>Social actions</td>
<td>Invite 20 friends</td>
<td>Social butterfly</td>
</tr>
<tr>
<td>Social actions</td>
<td>Share your first news item</td>
<td>Announcer</td>
</tr>
<tr>
<td>Social actions</td>
<td>Share 5 news items</td>
<td>Booster</td>
</tr>
<tr>
<td>Social actions</td>
<td>Share 25 news items</td>
<td>Publicist</td>
</tr>
<tr>
<td>Participation</td>
<td>Accepting the invitation link or signing up.</td>
<td>Welcome aboard</td>
</tr>
<tr>
<td>Participation</td>
<td>Login 3 days in a row</td>
<td>Enthusiast</td>
</tr>
<tr>
<td>Participation</td>
<td>Login 7 days in a row</td>
<td>Fanatic</td>
</tr>
<tr>
<td>Participation</td>
<td>Login 30 days in a row</td>
<td>Legendary</td>
</tr>
<tr>
<td>Game</td>
<td>Complete your first game</td>
<td>Just starting</td>
</tr>
<tr>
<td>Game</td>
<td>Complete 10 games</td>
<td>Try to stop me now</td>
</tr>
<tr>
<td>Game</td>
<td>Complete 50 games</td>
<td>Marathon runner</td>
</tr>
<tr>
<td>Answers</td>
<td>First correct answer</td>
<td>Good answer</td>
</tr>
<tr>
<td>Answers</td>
<td>10 correct answers in a row</td>
<td>Nice Streak</td>
</tr>
<tr>
<td>Answers</td>
<td>30 correct answers in a row</td>
<td>Awesome Streak</td>
</tr>
<tr>
<td>Answers - category</td>
<td>30 correct answers in IBM Facts</td>
<td>Thinker</td>
</tr>
<tr>
<td>Answers - category</td>
<td>200 correct answers in IBM Fact</td>
<td>IBM Guru</td>
</tr>
<tr>
<td>Answers - category</td>
<td>30 correct answers in WW Technology</td>
<td>Proud techy</td>
</tr>
<tr>
<td>Answers - category</td>
<td>200 correct answers in WW Technology</td>
<td>Tech guru</td>
</tr>
<tr>
<td>Answers - category</td>
<td>30 correct answers in You &amp; Your peers</td>
<td>Know thy friend</td>
</tr>
<tr>
<td>Answers - category</td>
<td>200 correct answers in You &amp; Your peers</td>
<td>Acquaintance champion</td>
</tr>
</tbody>
</table>

Table 3.2: The list of badges used in the game and the conditions for rewarding them.
game mechanics as well as the actions to be gamified are implemented in separate modules
and can be enabled, disabled or extended when needed. The user interface is generated
dynamically and uses technologies such as jQuery\(^5\) for improving the user interaction, the
D3 javascript visualisation library\(^6\) for displaying graphs, Toastr\(^7\) for displaying notifications
and Intro.js\(^8\) for building the introductory tutorials.

To store the data, we use a relational database system called PostgreSQL, chosen for
familiarity reasons. The entire systems runs on a Linux Virtual Machine in the IBM’s Cloud.
To perform periodic maintenance tasks we run a number of scripts using the cronjob software
available on Linux. Such tasks include: retrieving news from a RSS feed, checking the
number of news shared by a user, updating the IBMer score daily, building social questions
or resending invitations. We usually perform these tasks asynchronously from the main
application, during the night, to optimise resource usage.

### 3.10 Building the questions database

As mentioned before, players can take quizzes where they can choose questions from three

To build the questions from the first two categories we designed a public accessible web
interface. Figure 3.11 shows the interface for adding questions. We took the crowdsourcing
approach because formulating good quality, clear and unambiguous questions is very time
consuming. Having a number of users contributing with a small number of questions was
a better approach. The users received instruction on how to formulate questions and were
provided with example questions in each category. The IBM Facts questions were focusing
on the history of IBM, facts about IBM and IBM technologies, while the World Wide Technolo-
gies questions were concerning knowledge about information technology and technology in
general, historical facts and trending topics about technology.

We avoided specifying sources of information for building these questions because we
wanted to allow the users to be creative and submit diversified questions. To maximise
the number of distinct questions submitted, we implemented an algorithm that checks for
duplicate questions while the user is typing based on text similarity and correct answer.

The You &and Your Network questions are related to a user’s social connections. The
questions we ask are related to headlines and skills. Once a user signs up for the game,
a batch job will crawl information about his peers to build the questions. We designed the
algorithm for creating the questions such that there is only one correct answer. Although not
obvious, when asking about skills this is an important thing to have in mind.

\(^5\)http://jquery.com
\(^6\)http://d3js.com
\(^7\)http://github.com/CodeSeven/toastr
\(^8\)http://usablica.github.io/intro.js
3.11 Chapter conclusions

This chapter focussed on the design and implementation of the game “How much of an IBMer are you”. Next chapter, Chapter 4 explains the methodology used for measuring user engagement and the social online behaviour of employees and discusses how the tool is used in the experiments.

Figure 3.11: The interface for adding questions for the game. Users can see an explanation of the categories and a brief tutorial on how to formulate questions. They can also see random questions from the database on the Examples page.
In this chapter we explain the methodology used for measuring user engagement and the social online behaviour of employees in an enterprise context. The experimental design is a between subjects post-test only control group design.

First, we motivate the choices we made for the experimental design with a discussion about internal validity. Second, we describe the experimental procedure. Third, we present the objective and subjective metrics used in this thesis to answer the main research question of this thesis: “How can game elements be used to support user engagement and drive the online social behaviour of employees in an enterprise environment?”.

4.1 Experimental design

The participants were 206 IBM employees (N = 206) who signed up the tool “How much of an IBMer are you” in the period in the period of May 12th, 2014 to July 11th, 2014. In a between-subject post-test only control group experimental design, they were assigned to one of the four treatment groups:

Group 0 received the default game mechanics: an IBMer score and feedback for answering quiz questions.

Group 1 received, in addition to Group 0, leaderboards.

Group 2 received, in addition to Group 0, badges.

Group 3 received, in addition to Group 0, both leaderboards and badges.

The choice for these particular game mechanics is motivated by their popularity. Feedback is required to support learning. The score represents a global performance indicator in the game and closely relates to leaderboards. Badges are popular in a number of applications and studies have showed them effective for driving behaviour, as explained in Chapter 2.

After the participants interacted with the tool for a period of time, we measured the variables of interest objectively by looking at interaction logs and subjectively by means of a questionnaire.
4.1 Experimental design

The design has a few characteristics that deserve mention. First, it is a between subjects design. Since the subjects are only presented with a treatment, we limit the chance that the participants suffer boredom or be susceptible to the learning effect. One disadvantage of this design is that it requires a large number of participants. We overcome this requirement by deploying the tool on the IBM’s internal network with a large potential number of employees. Another disadvantage is the presence of variability in the individual characteristics of the participants. Because different participants follow different treatments it might happen that these individual differences introduce variance in the outcome. We minimise the effect which individual characteristics of the participants might have on our measurements by randomly assigning participants to treatment groups.

Second, the design is a control group post-test only. We measure how the treatment groups influence the dependent variables after the participants have used the tool. The control group is Group 0. This design accounts for standard threats to internal validity such as history, selection, instrumentation and testing. History is a threat when “an observed effect might be due to an event which takes place between the pre-test and post-test, when this effect is not the treatment of research interest” [25]. In our experiments such a history effect is the possibility that the participants discuss the game among themselves. These discussions might influence subject’s perceptions of and attitudes towards the experimental tool. This effect cannot be ruled out but we can assume that all the experimental groups will be affected in the same way and the existence of a control group minimises the history threat.

The testing thread happen when “an effect might be due to the number of times particular responses are measured” [25]. In the present context this threat is dealt with as we only measure the outcome of the treatment once.

Instrumentation “is a threat when an effect might be due to a change in the measuring instrument between pre-test and post-test and not to the treatment’s differential impact at each time interval” [25]. In the current experiments we control for this thread by using the same methodology to measure the dependent variables for all the treatment groups.

The independent variables for the experiment are the presence of leaderboards and presence of badges which are coded into the treatment group number. The dependent variables are user engagement and social online behaviour measured according to the metrics described in Section 4.3.

<table>
<thead>
<tr>
<th></th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Feedback</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Leaderboards</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Badges</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 4.1: The treatment groups used in the experiments.
4.2 Experimental procedure

4.2.1 Recruiting
We placed one banner and a few hundred flyers at IBM Headquarters in Amsterdam, Netherlands, inviting employees to play the online game “How much of an IBMer are you”. Other participants signed up to the game after receiving an email in their corporate email box from someone already in the game. During the signup process the employees were randomly assigned to one of the four treatments groups.

4.2.2 Procedure
The participants interacted with the experimental tool for a number of days, depending on the signup date, between May 12th and July 11th 2014. After this interval, we performed an objective quantitative analysis by examining web logs and usage patterns to see which groups were more engaged with and more socially active online in the tool. Following the objective analysis, we carried out a subjective analysis by means of a questionnaire. We sent an invitation to all employees who played the game to fill in a questionnaire. The interested reader can refer to the questionnaire in Appendix B. The questions in this questionnaire are meant to better understand the data but also to answer additional questions dealing with user’s perception.

4.3 Objective metrics
One objective of the current work is to measure engagement and social online behaviour of employees after being subjected to one of the four treatments. In the current context we measure user engagement by the average session length, player retention curve and the player lifetime. We measure social behaviour by the number of invites a player sent in the game and the number of news shared.

4.3.1 The average session length
The average session length is defined by the total time a player spent in the tool divided by the number of sessions. For each player we compute the total time spent in the game and the number of sessions from the requests log. Two requests are part of the same session if the time difference is less than 15 minutes. The longer the value, the more engaged a user is. The average session length falls into the online behaviour metrics and is a good indicative of user engagement, while normalising for the sign up time.

4.3.2 The player retention curve
The player retention curve is defined by the proportion of players who revisited the game and the frequency of their visits [23]. Because the player retention curve has been used to assess the enjoyability of different games with a purpose, we can not only compare the four treatments but also relate to other games from literature. The steepness of the slope...
indicates the level of engagement. A steep slope implies that many people played the game only a few times while a flatter slope indicate that only a few people abandoned the game after playing just a few times, while many people played a large number of games. Within each treatment group, we compute the number of games played by all players and display a log-log plot of the data.

4.3.3 The user lifetime

The user life time represents the number of correct and incorrect quiz questions answered by a user. We use this metric to understand how different game mechanics influence the number of correct and incorrect answers submitted by a user.

4.3.4 The number of invites sent

The number of invites sent refers to the invites a player sent from the tool to another IBMer connected on this player’s social networks. This metric closely relates to the first part of our definition of social online behaviour which is “establishing connections with other colleagues online”. The more invites a player sends the more active socially is in the game.

4.3.5 The number of news shares

The number of news shares refer to how many news articles a player shared on his/hers social networks from the tool. This metric relates to the second part of our definition of social online behaviour which is “promoting the company’s products and services on social or professional networks”.

4.4 Subjective metrics

One objective of the current work is to transfer knowledge to employees by having them answer quizzes on topics such as facts about IBM, facts about technology in general and colleagues’ skills. We assess learning subjectively by means of a questionnaire.

Another objective of this thesis is to quantify engagement of employees. We do this objectively as described in Section 4.3, but we also evaluate engagement subjectively to understand users’ perception.

4.4.1 Perceived learning

We assess learning in the three categories (IBM Facts, World Wide Technology and You & Your network) by asking employees to select on a 5-point Likert scale their level of agreement with the following questions:

Q 1: I have learned about IBM by answering questions in IBM Facts.
Q 2: I have learned about Technology by answering questions in World Wide Technology.
Q 3: I have learned about my colleagues by answering questions in You & Your Network.

The answer choices were: Strongly disagree, Disagree, Neither agree nor disagree, Agree and Strongly agree.
4.4.2 Perceived user engagement

To subjectively measure user engagement, we use the multidimensional scale constructed by O'Brien et al [21]. The authors identify six attributes of engagement: perceived usability, aesthetics, focused attention, felt involvement, novelty, and endurability. A Principal Component Analysis was carried out on the set of items referring to one of the six categories to identify the most parsimonious set of items. Based on the analysis we choose the top three attributes of engagement based on the percentage of variance explained. We end up with focused attention, perceived usability and aesthetics. We use the same PCA analysis to select one statement that represents best each of the three attributes.

Q 4: I found this website confusing to use. (perceived usability)
Q 5: The website was aesthetically appealing. (aesthetics)
Q 6: I forgot about my immediate surroundings while playing on the website. (focused attention)

The employees that agreed to fill in the questionnaire were asked to select on a 5-point Likert scale their level of agreement with these questions. As with the previous questions, the answer choices were: Strongly disagree, Disagree, Neither agree nor disagree, Agree and Strongly agree.

4.5 Chapter conclusions

This chapter focused on the methodology used for measuring user engagement and the social online behaviour of employees in an enterprise context. We presented the experimental design with an emphasis on the threats to internal validity. We then described the experimental procedure, including the process of recruiting participants. Finally, we defined the metrics used in the experiments.

Both the objective and the subjective metrics will come back in Chapter 5 when we present the results of the various experiments. The treatment groups defined here will also be used in the coming chapters.
Chapter 5

Results and discussion

In this chapter we present the various experiments we performed in order to answer the main research questions of this work. The insights gathered will help us understand user engagement and social online behaviour in an enterprise context.

First, we present high level usage statistics of the tool we deployed internally at IBM. Second, we analyse the employees’ assignment to experimental groups and provide information on gender, role, country and management for each group. Third, we describe the analysis we performed and discuss the results.

5.1 Overview

The game How much of an IBMer are you has been played 778 times in the period of May 12th, 2014 to July 11th, 2014. The number of games per user ranged between 0 and 103. The 206 employees spent in the game a total number of 66 hours, while individual play time was between 1 minute and 8.3 hours. The players answered a total of 2061 questions in IBM Facts, 1720 questions in World Wide Technology and 3995 social questions, invited 553 colleagues to play and shared company related news 299 times.

5.2 Groups demographics

Table 5.1 shows the gender, country, department and the number of managers/non-managers within the experimental groups. By looking at the gender we notice that, 77.67% of the employees in our experiments are male and about 22.33% are female. This ratio is not unexpected since worldwide, the International Business Machines employs 26% female and 74% male.

The employees who signed up for the game are spread across nine countries and three continents. A large percentage of 78.15% works in The Netherlands, 8.73% in Belgium, 7.76% in Romania, 1.45% in the United Kingdom, 1.45% in the United States of America and less than 1% in Argentina, France, Peru and Germany. Because we advertised the game internally in the IBM’s HQ in the Netherlands we expected to have a large number of users from the Netherlands. User registrations coming from other countries were mostly due
### 5.2 Groups demographics

#### Results and discussion

<table>
<thead>
<tr>
<th>Number of employees</th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>41</td>
<td>41</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td>female</td>
<td>7</td>
<td>12</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Belgium</td>
<td>2</td>
<td>11</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>France</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Germany</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Netherlands</td>
<td>41</td>
<td>39</td>
<td>39</td>
<td>42</td>
</tr>
<tr>
<td>Peru</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Romania</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>United States</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>undefined</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Management</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>manager</td>
<td>13</td>
<td>9</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>non-manager</td>
<td>35</td>
<td>44</td>
<td>45</td>
<td>39</td>
</tr>
<tr>
<td><strong>Department</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Business Services</td>
<td>19</td>
<td>17</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>IBM Sales &amp; Distribution</td>
<td>12</td>
<td>14</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>IBM Integrated Operations</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Global Technology Services</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>IBM Software Group</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>IBM Systems &amp; Technology</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>IBM Integrated Supply Chain</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>undefined</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5.1: Employee demographics across the four treatment groups.
to the invitation mechanism in the game, since no advertisement campaign was employed outside.

The distribution of managers and non-managers is as follows: managers make 20.87% of the total number of players while non-managers represent 79.12%.

The employees work in the Global Business Services department (30.58%), IBM Sales & Distribution (31.55%), IBM Integrated Operations (17.47%), Global Technology Services (7.76%), IBM Software Group (8.25%), IBM Systems & Technology (0.48%) and IBM Integrated Supply Chain (1.94%). We were not able to retrieve the department for 4 employees.

The random assignment yielded acceptable results taking into account the relatively small sample size. Although the number of female employees is not equally distributed across the four groups, the differences are not large. The male employees, on the other hand, were better distributed across the groups. The number of managers and non-managers are also well distributed across the groups. All the groups have assigned employees from different countries and different departments.

### 5.3 User engagement

This section is concerned with answering one of the main research questions of this thesis RQ 1: Which game mechanic is more effective for driving user engagement? by analysing the average session length, the user lifetime and the player retention curve across the treatment groups.

We split the research question into three sub research questions:

**RQ 1.1:** Which game mechanics are more effective for incentivising users to stay longer on the website?

**RQ 1.2:** Which game mechanics are more effective for incentivising users to answer questions?

**RQ 1.3:** Which game mechanics stimulate better perceived user engagement?

#### 5.3.1 Which game mechanics are more effective for incentivising users to stay longer on the website?

To answer RQ 1.1 we use the average session length defined by the total time in minutes spent on the website by a user divided by the number of sessions of the user. We look at the average session length in the four treatment groups and use the Mann-Whitney-Wilcoxon non-parametric test [26] to test for statistical significant difference across the groups. The
5.3 User engagement

<table>
<thead>
<tr>
<th></th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>0.0049</td>
<td>0.0134</td>
<td>7.08e-05</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>0.0049</td>
<td>0.8125</td>
<td>0.0981</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.0134</td>
<td>0.8125</td>
<td>0.0837</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>7.08e-05</td>
<td>0.0981</td>
<td>0.0837</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.3: Mann-Whitney-Wilcoxon pair-wise significance test for AvgSessionLength variable.

Mann-Whitney-Wilcoxon test was preferred to the t-test because it does not make any assumptions about the distribution of the data.

Our hypothesis is that badges and leaderboards will improve the average session length, while combining them will further improve the results. The average session length statistics are summarised in Table 5.2. When users are provided with either leaderboards (Group 1) or badges (Group 2) we see an improvement in the average session length. When we combine leaderboards and badges (Group 3) the effect is even more pronounced resulting in even longer sessions.

We now want to further investigate if the difference between means is statistically significant. For this we employ the two samples Mann-Whitney-Wilcoxon to test the null hypothesis that the distributions of the two groups are identical.

The p-values for the statistical test are summarised in the Table 5.3. The p-values show how likely the observed data would be if the null hypothesis was true. We have strong evidence to reject the null hypothesis that the two samples come from the same distribution when comparing Group 0 with Groups 1,2 and 3. This means that the players tend to stay on the website longer when they see leaderboards or receive badges. The effect is even more pronounced when the two game mechanics are combined. When used alone, there is no preference towards either leaderboards or badges.

Another way of looking at how motivated players are to spend time on the website is to analyse the player retention curve defined by the proportion of players who revisited the tool and the frequency of their visits. The number of visits are translated to the number of different sessions.

In Figure 5.1 we compare the slopes of the lines between the four treatment groups. When comparing our groups, we can see that leaderboards and badges produce a flatter line compared to the control group which suggests in turn higher level of engagement. The slope of the two lines are almost the same. As suggested by previous analysis, we see that the combination of leaderboards and badges produce the best results (the flattest curve). The interested reader can refer to 5.2 for a comparison between our game and other games with a purpose in literature.
Results and discussion

5.3 User engagement

5.3.2 Which game mechanics are more effective for incentivising users to answer questions?

To answer RQ 1.2 we use the number of total and correct questions in the three categories: IBM Facts (TotalIBMQuestions and CorrectIBMQuestions), World Wide Technology (TotalWWTQuestions and CorrectWWTQuestions) and You & Your Network (TotalSocialQuestions and CorrectISocialQuestions). The authors in [24] study the effect of incentive mechanics on the number of questions a user answers on their platform by means of a Poisson regression. Because of the similarities with the work presented in [24], we take the same approach with the difference that we use a negative binomial regression [27] to account for over-dispersed data.

Our hypothesis is that badges and leaderboards will improve the number of questions answered. We also believe than combined they give the best results. Table 5.5 shows the coefficients and the significance levels computed for each mechanism by the negative binomial regression model. The intercepts corresponding to the control group are omitted from the table. Giving users both leaderboards and badges has the strongest impact on increasing the number of both correct and total number of questions. Only leaderboards or only badges does not increase the number of total answered questions nor the number of correct questions.

When looking at the number of total and correct World Wide Technology questions in Table 5.6, we notice that only giving badges or only giving leaderboards has a significant effect on both increasing overall participation and the number of correct answers. Badges are slightly better across both variables. When combined, they give the strongest effect increasing the number of answered questions by 370% ($e^{1.31}$) and the number of correct
5.3 User engagement

Results and discussion

<table>
<thead>
<tr>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM Facts</td>
<td>7.1458</td>
<td>7</td>
<td>11.0555</td>
</tr>
<tr>
<td>World Wide Technology</td>
<td>3.6458</td>
<td>7.6603</td>
<td>8.4629</td>
</tr>
<tr>
<td>You &amp; Your Network</td>
<td>8.9375</td>
<td>18.9811</td>
<td>12.4074</td>
</tr>
</tbody>
</table>

Table 5.4: Average number of questions per user in the three question categories across the treatment groups.

<table>
<thead>
<tr>
<th>Variable: TotalIBMQuestions</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaderboards</td>
<td>-0.0206</td>
<td>0.9410</td>
</tr>
<tr>
<td>Badges</td>
<td>0.4466</td>
<td>0.1067</td>
</tr>
<tr>
<td>Leaderboards+Badges</td>
<td>0.7281</td>
<td>0.009 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: CorrectIBMQuestions</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaderboards</td>
<td>0.0567</td>
<td>0.9410</td>
</tr>
<tr>
<td>Badges</td>
<td>0.5125</td>
<td>0.1067</td>
</tr>
<tr>
<td>Leaderboards+Badges</td>
<td>0.8461</td>
<td>0.0093 **</td>
</tr>
</tbody>
</table>

Table 5.5: The effect of different game-mechanisms in incentivising users to answer more IBM Questions (TotalIBMQuestions) and more correct IBM Questions (CorrectIBMQuestions). The coefficients were computed by a negative binomial regression model (***: 0.1% significance, **: 1% significance, *: 5% significance).

<table>
<thead>
<tr>
<th>Variable: TotalWWTQuestions</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaderboards</td>
<td>0.7425</td>
<td>0.0325 *</td>
</tr>
<tr>
<td>Badges</td>
<td>0.8542</td>
<td>0.0303 *</td>
</tr>
<tr>
<td>Leaderboards+Badges</td>
<td>1.3121</td>
<td>0.0006 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: CorrectWWTQuestions</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaderboards</td>
<td>0.9305</td>
<td>0.021 *</td>
</tr>
<tr>
<td>Badges</td>
<td>0.9431</td>
<td>0.307 *</td>
</tr>
<tr>
<td>Leaderboards+Badges</td>
<td>1.5055</td>
<td>1.20e-05 ***</td>
</tr>
</tbody>
</table>

Table 5.6: The effect of different game-mechanisms in incentivising users to answer more World Wide Technology Questions (TotalWWTQuestions) and more correct World Wide Questions (CorrectWWTQuestions). The coefficients were computed by a negative binomial regression model (***: 0.1% significance, **: 1% significance, *: 5% significance).

questions by 448% \(e^{1.50}\) compared to the control group.

Table 5.7 displays the results of fitting our regression model on the number of total and correct social questions. Leaderboards show an improvement in participation for both variables, while badges do not have such an effect. The social questions ask users about their peer's job role and skills and are considered easy to answer questions. Therefore when presented with leaderboards, the users will try to improve their score and thus the position in the leaderboard by answering questions that do not take too much effort. As with the previous results, the combination of leaderboards and badges is the best for incentivising user
5.3 User engagement

### Table 5.7: The effect of different game-mechanisms in incentivising users to answer more Social Questions (TotalSocialQuestions) and more correct Social Questions (CorrectSocialQuestions). The coefficients were computed by a negative binomial regression model (***: 0.1% significance, **: 1% significance, *: 5% significance).

<table>
<thead>
<tr>
<th>Variable: TotalSocialQuestions</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaderboards</td>
<td>0.7532</td>
<td>0.030 *</td>
</tr>
<tr>
<td>Badges</td>
<td>0.3437</td>
<td>0.323</td>
</tr>
<tr>
<td>Leaderboards+Badges</td>
<td>1.4394</td>
<td>4.14e-05 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: CorrectSocialQuestions</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaderboards</td>
<td>0.8024</td>
<td>0.021 *</td>
</tr>
<tr>
<td>Badges</td>
<td>0.3560</td>
<td>0.307</td>
</tr>
<tr>
<td>Leaderboards+Badges</td>
<td>1.5400</td>
<td>1.20e-05 ***</td>
</tr>
</tbody>
</table>

The results suggest that leaderboards and badges always increase user participation across all six variables. However, the effect of only badges or only leaderboards is dependent on the type of questions users answer. Used alone, leaderboards are effective only for incentivising users to answer World Wide Technology and Social questions. When used alone, badges are only effective for incentivising users to answer questions in the World Wide Technology category. For a quick overview of the number of average questions per user, the interested reader can refer to Table 5.4.

### 5.3.3 Which game mechanics stimulate better perceived user engagement?

To answer RQ 1.3 we use partially the questionnaire developed by O’Brien et al [21]. We specifically look at the focused attention, perceived usability and aesthetics attributes of user engagement as these three attributes were identified by the authors as the most important attributes after principal component analysis. We analyse the data resulting from N=41 respondents who answered the questions related to user engagement. The employees who answered the questions from the questionnaire were recruited among the users who signed up for the game. We invited all the players, by means of an email, to answer the questionnaire related to their experience with the game “How much of an IBMer are you”. We also contacted a subset of users on the corporate instant messaging service to kindly ask them to fill in the questionnaire. To construct the questionnaire and to manage the answers, LimeSurvey\(^1\) was used. The 41 respondents were split into the four treatment groups as follows: there were 6 participants in Group 0, 13 participants in Group 1, 10 participants in Group 2 and 12 participants in Group 3.

Our hypothesis is that employees who receive either badges or leaderboards or the combination of the two will perceive themselves better engaged with the tool than the control group. To process the answers, we recode the Likert scale on an ordinal scale with 1 (Strongly disagree) to 5 (Strongly agree) and compute the means and standard deviations

---

\(^1\)https://www.limesurvey.org/en/
5.4 Social online behaviour

Results and discussion

<table>
<thead>
<tr>
<th>Variable: Aesthetics</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>3.3333</td>
<td>0.8164</td>
</tr>
<tr>
<td>Group 1</td>
<td>4</td>
<td>0.7071</td>
</tr>
<tr>
<td>Group 2</td>
<td>3.8</td>
<td>0.4216</td>
</tr>
<tr>
<td>Group 3</td>
<td>4.3</td>
<td>0.6215</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: Focused Attention</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>3.5</td>
<td>0.5477</td>
</tr>
<tr>
<td>Group 1</td>
<td>3</td>
<td>0.7071</td>
</tr>
<tr>
<td>Group 2</td>
<td>3.1</td>
<td>0.9944</td>
</tr>
<tr>
<td>Group 3</td>
<td>3.1</td>
<td>0.9962</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: Perceived usability</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>2</td>
<td>0.6324</td>
</tr>
<tr>
<td>Group 1</td>
<td>2</td>
<td>0.8164</td>
</tr>
<tr>
<td>Group 2</td>
<td>1.6</td>
<td>0.6992</td>
</tr>
<tr>
<td>Group 3</td>
<td>1.7</td>
<td>0.4923</td>
</tr>
</tbody>
</table>

Table 5.8: First order statistics of the perceived user engagement across the four treatment groups.

across our treatment groups. We test for statistically significant differences using the Mann-Whitney-Wilcoxon non-parametric test. The Mann-Whitney-Wilcoxon test was preferred here because it makes the assumption that the data is at least ordinal data which always holds for a Likert-scale.

The results are summarised in Tables 5.8 and 5.9. Looking at the resulting data, there is no statistical significant difference between groups for the focussed attention and perceived usability attributes. When analysing responses concerning aesthetics, we notice that users in Group 3 perceived the tool to be more visually appealing than users in the other groups. This is an interesting fact because the tool looks the same regardless of the group, but it seams that receiving leaderboards and badges compared to only leaderboards or only badges or none of the two strongly influences user’s perception.

In the current context, the data shows that the presence of badges and leaderboards influences the user’s perception about aesthetics. The other two attributes are not influenced. However, it is important to remain critical of the results, because we assess user’s perception and the subjective data is susceptible to multiple drawback as suggested by [22].

5.4 Social online behaviour

This section is concerned with answering one of the main research questions of this thesis RQ 2: Which game mechanic is more effective for promoting online social behaviour of employees? by analysing the number of invitations sent and the number of news shared across the treatment groups.

We split the research question into three questions:
Results and discussion

5.4 Social online behaviour

<table>
<thead>
<tr>
<th>Variable: Aesthetics</th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>0.1244</td>
<td>0.2053</td>
<td>0.0293</td>
<td>0.0093 *</td>
</tr>
<tr>
<td>Group 1</td>
<td>0.1244</td>
<td>0.4849</td>
<td>0.3778</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.2053</td>
<td>0.4849</td>
<td>0.0711</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>0.0293 *</td>
<td>0.3778</td>
<td>0.0711</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: FocussedAttention</th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>0.1605</td>
<td>0.5139</td>
<td>0.3224</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>0.1605</td>
<td>0.5924</td>
<td>0.9538</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.5139</td>
<td>0.5924</td>
<td>0.8088</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>0.9538</td>
<td>0.7565</td>
<td>0.8088</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: PerceivedUsability</th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>0.8784</td>
<td>0.2562</td>
<td>0.2821</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>0.8784</td>
<td>0.2298</td>
<td>0.3194</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.2562</td>
<td>0.2298</td>
<td>0.6795</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>0.2821</td>
<td>0.3194</td>
<td>0.6795</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.9: Mann-Whitney-Wilcoxon pair-wise significance test for perceived user engagement across the four treatment groups.

<table>
<thead>
<tr>
<th>Variable: InvSent</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaderboards</td>
<td>1.5013</td>
<td>0.0187 *</td>
</tr>
<tr>
<td>Badges</td>
<td>0.8515</td>
<td>0.1936</td>
</tr>
<tr>
<td>Leaderboards+Badges</td>
<td>1.6756</td>
<td>0.0090 ***</td>
</tr>
<tr>
<td>AcceptedInv</td>
<td>1.1787</td>
<td>3.33e-16 ***</td>
</tr>
</tbody>
</table>

Table 5.10: The effect of different game-mechanisms in incentivising users to invite colleagues (InvSent). The coefficients were computed by a negative binomial regression model (***: 0.1% significance, **: 1% significance, *: 5% significance).

RQ 2.1: Which game mechanics are more effective for incentivising users to invite their colleagues to play the game?

RQ 2.2: Do IBMers have strong ties across departments?

RQ 2.3: Which game mechanics are more effective for incentivising users to share IBM related news on their social/professional networks?

5.4.1 Which game mechanics are more effective for incentivising users to invite their colleagues to play the game?

To answer RQ 2.1 we use the number of invitations sent to play the game by players to their colleagues. Our hypothesis is that badges and leaderboards will improve the number of invitations sent, while combining the two game elements will result in an even stronger effect.
5.4 Social online behaviour

Results and discussion

<table>
<thead>
<tr>
<th>Variable: InvAccepted</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>InvPM</td>
<td>0.1737</td>
<td>2e-16 ***</td>
</tr>
<tr>
<td>InvNPM</td>
<td>0.1993</td>
<td>1.3e-07</td>
</tr>
</tbody>
</table>

Table 5.11: The effect of including a personal message (InvPM) and not including a personal message (InvNPM) on the numbers of invitations accepted (InvAccepted). The coefficients were computed by a negative binomial regression model (***: 0.1% significance, **: 1% significance, *: 5% significance).

Before starting the analysis we suspect that a large number of players did not send any invitations in the game. Looking at the data, we observe that this is the case, with only 21.07% of the players having sent at least one invitation. We want to use a regression model to fit our data and we believe a negative binomial regression is better for fitting this type of data, but the excessive number of zeros might be better explained by a zero-inflated Poisson regression[28]. We fit the two regressions and test how well they approximate our data by means of a Vuong closeness test[29]. The null hypothesis is that the two models are equally close to the actual model, against the alternative that one model is closer. Although the result suggests that the zero-inflated Poisson regression model is superior, we do not have enough evidence to reject the null hypothesis (p-value: 0.327). Therefore, we chose to use the negative binomial regression for consistency with the other results.

Table 5.10 shows the coefficients and significance levels of fitting a negative binomial regression model on the variable InvSent while accounting for incentive types and number of accepted invitations. We incorporate the number of accepted invitations (AcceptedInv) because players only receive points and earn badges when invitations are accepted. The intercept corresponds to the control group and has been omitted from the table. By looking at the significance levels we discover that leaderboards have a positive effect on the number of invitations sent. Giving users badges does not result in the same effect. When combined, leaderboards and badges produce the largest increase in the number of invitations sent.

Although not central to answering the sub research question, we want to investigate whether invitations which include a personal message are more likely to be accepted by the invitee. For this we run again a negative binomial regression model on the number of invitations accepted (InvAccepted) when players include a personal message (InvPM) or they do not include a message (InvNPM). Table 5.11 summarises the regression coefficients and significance levels. The regression coefficients suggest that, contrary to one would expect, including a personal message does not help in increasing the acceptance rate.

5.4.2 Do IBMers have strong ties across departments?

We answer RQ 2.2 representing the invitation exchanged by employees by a graph. We use the number of invitations sent in the game as a proxy for strong ties because we believe employees will send invitations to those colleagues they have a strong connection with. A strong tie refers to friendship and familial relationships and is characterised by “the amount of time and emotions invested in a relationship”[30]. Our hypothesis is that employees have stronger relationships with colleagues from the same department, than colleagues in other departments.
Results and discussion

5.4 Social online behaviour

Figure 5.3: The graph shows the invitations sent in the game. The different colours represent departments and the node size represent the number of invitations sent by the player.

The nodes in the graph represent players (inviter or invitee) and the edges represent invitations sent between these players. A node is color and size coded. The color shows the department and the size represent the number of invitations sent by the player. The more invitations sent, the larger the node. We represent the graph data in the Graph Exchange XML Format\(^2\) and import it with Gephi\(^3\) for visualisation.

During the data preparation step we retrieve the inviter name, invitee name and the time of the invitation. Furthermore, we automatically determine the department of the person invited using the Faces API\(^4\).

Once imported, we further process the graph, partitioning the nodes by department and ranking them by the number of edges going out. We then run the Fruchterman Reingold force-directed drawing algorithm\(^5\) to display the final graph.

Looking at the resulting visualisation in Figure 5.3 we notice that employees send invitations to both colleagues in the same department as well as other departments. From the

---

\(^2\)http://www.gexf.net
\(^3\)http://gephi.github.io/
\(^4\)http://faces.tap.ibm.com
\(^5\)http://en.wikipedia.org/wiki/Force-based_algorithms
5.4 Social online behaviour

Results and discussion

Table 5.12: The effect of different game-mechanisms in incentivising users to share more news from the game (NewsShared) or post IBM related content (NewsPostedTotal) on LinkedIn. The coefficients were computed by zero-inflated Poisson regression model (**: 0.1% significance, ***: 1% significance, *: 5% significance).

<table>
<thead>
<tr>
<th>Variable: NewsShared</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badges</td>
<td>-0.3066</td>
<td>0.7867</td>
</tr>
<tr>
<td>Leaderboards</td>
<td>0.6108</td>
<td>0.6167</td>
</tr>
<tr>
<td>Badges+Leaderboards</td>
<td>0.1874</td>
<td>0.8701</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: NewsPostedTotal</th>
<th>Coefficients</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badges</td>
<td>0.4277</td>
<td>0.4173</td>
</tr>
<tr>
<td>Leaderboards</td>
<td>0.2168</td>
<td>0.6735</td>
</tr>
<tr>
<td>Badges+Leaderboards</td>
<td>0.5232</td>
<td>0.3440</td>
</tr>
</tbody>
</table>

total number of invitations sent in the game, 48.42% were inter-department, while 51.57% were intra-department. The employees in IBM Global Technologies and IBM Sales and Distributions sent out the most invitations to colleagues outside their departments, while the employees in Global Technology Services and Global Business Services sent out the most invitations to colleagues inside their departments. The data suggests IBMers have strong ties in the department they work in, but also in other departments. Another important observation is that the department of the user sending the invitations plays a role in the spread of invitations.

5.4.3 Which game mechanics are more effective for incentivising users to share IBM related news on their professional network?

To answer RQ 2.3 we use the number of news shared from the game by the employees on their professional network. Our hypothesis is that badges and leaderboards will act as an effective incentive to improve the number of news shared.

Looking at the data, there are only 11 employees who shared news articles on their LinkedIn professional network, making up only 5.33% of the total number of employees who played the game. Since this variable (NewsShared) has mainly values of zero, we run a zero-inflated Poisson regression model. Table 5.12 summarises the regression coefficients and significance levels. The results show that there is no statistically significant difference between the control group and the groups that received either badges, either leaderboards or the combination of the two.

Although we cannot confirm the hypothesis, it might be that the content of the news was not appealing enough to worth redistribution even with an additional incentive. We want to further investigate this alternative hypothesis by extending our analysis to include all IBM related news shared on the user’s LinkedIn network (NewsPostedTotal). We include in this analysis all the posts a user makes on his/hers social networks that include the word “IBM”. This makes sense because employees also receive points for posting IBM related content which did not originate from the tool. There are 44 IBMers who shared IBM related content on their LinkedIn network after signing up to the game, representing 21.35% of our sample
5.5 Perceived learning

Results and discussion

<table>
<thead>
<tr>
<th>Variable: ShareNewsFromTheWebsite</th>
<th>NO (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>20</td>
</tr>
<tr>
<td>Group 1</td>
<td>36.36</td>
</tr>
<tr>
<td>Group 2</td>
<td>66.66</td>
</tr>
<tr>
<td>Group 3</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 5.13: Percentage of employees who did not know that sharing news from the website brings them points in the game.

size. It is important to note that through the LinkedIn API\(^6\) we can only access user’s content subsequent to the moment when the user authorised the application. We could have avoided this limitation of the API by imposing a waiting time between the moment an employee signed up for the game and the moment (s)he was permitted access to the game, but that would not have been practical. Analysing the coefficients of the resulting zero-inflated Poisson model on the variable NewsPostedTotal, we see that there is no significant statistical difference between the treatment groups. The alternative hypothesis is therefore not supported.

We also want to investigate the alternative hypothesis that the employees did not know they can earn points by sharing IBM related content on their professional networks. We assess whether the alternative hypothesis is true by means of a questionnaire. We ask the participants to rate their level of agreement with the following question: I could earn points in the game by sharing news from the tool on LinkedIn or Facebook. In total, 37 employees answered the question related to earning points by sharing news. There were 5 participants in Group 0, 11 participants in Group 1, 9 participants in Group 2 and 12 participants in Group 3. Table 5.13 summarises the results. In Group 0 and Group 3 20% and 25% percent of the users weren’t aware they could score points by sharing news from the tool. In Group 1 and 2 the percentage is even larger, totalling 36.36% and 66.66%. These values show that it is important to make sure it is clear for the players which actions rewards them. The results presented in this section show that there is more work to be done in order to understand the effect of the current game mechanics on this type of social behaviour. Although we couldn’t prove the effectiveness of leaderboards and/or badges on incentivising employees to share news, these game mechanics were proved successful for promoting other types of social behaviour.

5.5 Perceived learning

This section is concerned with answering one of the secondary research questions of this thesis RQ 3: Which game mechanic is more effective for supporting learning? To answer RQ 3 we analyse the data resulting from N=41 respondents who answered the questions related to learning from the questionnaire as in Subsection 5.3.3. There were 6 participants in Group 0, 13 participants in Group 1, 10 participants in Group 2 and 12 participants in Group 3.

\(^6\)http://developers.linkedin.com
### Table 5.14: First order statistics of the effect of learning in the three categories (IBM Facts, World Wide Technology, and You & Your network) as perceived by the respondents to the questionnaire.

<table>
<thead>
<tr>
<th>Variable: LearnAboutIBM</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>4</td>
<td>0.6324</td>
</tr>
<tr>
<td>Group 1</td>
<td>4</td>
<td>0.5773</td>
</tr>
<tr>
<td>Group 2</td>
<td>3.7</td>
<td>0.8232</td>
</tr>
<tr>
<td>Group 3</td>
<td>4.2</td>
<td>0.3892</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: LearnAboutWWTechnology</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Group 1</td>
<td>4</td>
<td>0.8164</td>
</tr>
<tr>
<td>Group 2</td>
<td>3.4</td>
<td>0.8432</td>
</tr>
<tr>
<td>Group 3</td>
<td>4</td>
<td>0.4264</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable: LearnAboutColleagues</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>3.3</td>
<td>1.211</td>
</tr>
<tr>
<td>Group 1</td>
<td>3.4</td>
<td>0.8697</td>
</tr>
<tr>
<td>Group 2</td>
<td>3.6</td>
<td>0.8432</td>
</tr>
<tr>
<td>Group 3</td>
<td>3.7</td>
<td>0.4923</td>
</tr>
</tbody>
</table>

Our hypothesis is that the employees who received either badges or leaderboards or the combination of the two will perceive they learned more than the employees in the control group. We recode the answers on the Likert scale on an ordinal scale from 1 (Strongly disagree) to 5 (Strongly agree) and compute the means and standard deviations across the four treatment groups. We use Mann-Whitney-Wilcoxon non-parametric test [26] to test for statistical significant differences between groups. The results are summarised in Tables 5.14 and 5.15. Looking at perceived learning in IBM Facts and You & Your Network, there is no statistical difference between the four groups. The World Wide Technology category, however, shows a statistical difference between Group 2 and Group 3 (p-value: 0.0393). The employees receiving badges weren’t sure whether they learned about technology, while the employees receiving both badges and leaderboards thought they learned.

The results show that, although not statistical significant, the employees felt that they have learned. The effects of learning are more pronounced in the IBM Facts and the World Wide Technology categories and less in the Social category. The subjective data indicates that the tool is effective for the goal of promoting learning, partly because the feedback the players receive when answering questions and less because of the leaderboards and/or badges. It is important to be critical towards the results because of the small sample size and the subjective nature of the data which comes with multiple drawback as suggested by [22].
We measured user engagement objectively by the average session length, the number of total and correct questions answered and the player retention curve and subjectively by means of a questionnaire. Deploying leaderboards or badges has resulted in an increased average session length. When combined, the two game mechanics further improve this metric. When looking at the total and correct number of questions answered, leaderboards or badges show increased participation only for some question types. Again, when used together, leaderboards and badges give the best results, showing a positive effect on all the question types. Looking at the player retention curve, we notice that both leaderboards and badges increase the number of user sessions and combined show the strongest effect.

Analysing the respondents’ answers to the questionnaire, we notice that, from the three attributes of engagement, only the users’ perception of aesthetics is influenced by the game mechanics. The other two attributes, focused attention and perceived usability are not influenced.

The online social behaviour has been measured according to our definition (1) establishing connections with other colleagues online and (2) promoting the company’s products and services on social or professional networks. We used the number of invitations sent for the first part of the definition and the number of news shared on LinkedIn for the second part of the definition.

The results of the analysis performed on the number of invitations sent show that leaderboards have a positive effect on the number of invitations sent, while the combination of

<table>
<thead>
<tr>
<th>Group</th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>1</td>
<td>0.5312</td>
<td>0.5613</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td></td>
<td>0.3994</td>
<td>0.4442</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.5312</td>
<td>0.3994</td>
<td>0.1193</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>0.5613</td>
<td>0.4442</td>
<td>0.1193</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>0.7766</td>
<td>0.07758</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>0.7766</td>
<td>0.0723m</td>
<td>0.7565</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.07758</td>
<td>0.0723</td>
<td>0.0393 *</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>1</td>
<td>0.7565</td>
<td>0.0393 *</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Group 0</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>0.9627</td>
<td>0.6883</td>
<td>0.5666</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>0.9627</td>
<td>0.4226</td>
<td>0.1982</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.6883</td>
<td>0.4226</td>
<td>0.9098</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>0.5666</td>
<td>0.1982</td>
<td>0.9098</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.15: Mann-Whitney-Wilcoxon pair-wise significance test for perceived learning across groups.

5.6 Chapter conclusions
leaderboards and badges strengthens this effect. We discover that, contrary to one would expect, including a personal message with the invitation, does not help improve the acceptance rate. As part of the analysis, we investigate the invitations graph and conclude that IBMers are well connected intra and inter-department.

Looking at the number of news shared on LinkedIn, we notice that the differences between the experimental groups are not statistical significant. To try to understand why this is the case, we formulate two alternative hypothesis. The analysis shows the importance of more work for understanding the effect of current game mechanics on this type of social behaviour.

Lastly, we look at perceived learning by analysing respondents’ answers to a questionnaire. The results show that leaderboards or badges or the combination of the two does not impact learning neither negatively nor positively, exception being made by the World Wide Technology category. In the World Wide Technology category, the group receiving both badges and leaderboards has perceived to have learned more than the group receiving only leaderboards.
Chapter 6

Conclusions and Future Work

6.1 Conclusions

The main objective of this thesis was to research how can game elements be used to support user engagement and drive the online social behaviour of employees in an enterprise environment. To support our research, we created a flexible and effective software platform which (1) provides a modular implementation of several game mechanics, (2) has the means to transfer knowledge to employees, (3) has the means to support, access and analyse social connections, (4) has the means to support the distribution of information about services and products outside the company and finally (5) offers the means to measure our variables of interest.

We deployed our platform at IBM where we conducted various experiments which lasted from May 12th to July 11th, 2014 and involved 206 employees. In a between-subject post-test only control group experimental design, we randomly assigned participants to one of the four treatment groups. For each group, we analysed both subjective and objective data related to user engagement, knowledge acquisition and social online behaviour.

The experiments provided in Chapter 5 resulted in useful insights, showing that the implementation of game mechanics was beneficial for the goals of this thesis. First, the tool was received well by employees, which quickly became engaged. They answered quiz questions, invited colleagues to play the game and shared news. The level of engagement was dependent on the game mechanics a group received. Second, the tool has been proved successful for promoting knowledge acquisition. Third, the results related to social online behaviour are also encouraging, showing that the implementation of various game mechanics are effective for encouraging some types of social behaviours.

Building the platform for studying gamification was not a trivial task, as it required a great deal of work to properly design and implement the user experience. A poorly designed tool would not have been compatible with studying gamification. To support knowledge acquisition, we also invested a considerable amount of time into building the questions database. We were rigorous with defining the experimental design and accounting for various threats to the internal validity of our experiments. Lastly, the application of proper statistical tests guaranteed a sound mathematical analysis.
6.2 Future work

We identify several directions for future work.

The promising results we obtained when we analysed social behaviour of employees suggest further exploration. For instance, the news module should be further developed to include fresh content tailored to the user’s interests, rather than only providing general news. Another direction to explore is to better promote the tutorials that explain the functionality of the tool. This will improve the understanding of the tool for the user and potentially make the game mechanics more effective.

As part of this work, we developed a questionnaire to assess users’ perception of the tool “How much of an IBMer are you”. The questionnaire’s goals were to: (1) better understand if the users were aware of the actions that brought them points; (2) understand users’ motivation for using the different modules in the tool; (3) assess perceived user engagement. We should further process the answers to this questionnaire to include the items not included in the analysis we performed. We believe that we can refine the objective analysis by using data from the questionnaire.

The questions database can be further expanded to keep people engaged with the tool. Although the size and content of the database we build was appropriate for the goals of the current work, we believe that a larger number and more diverse questions will prove effective for supporting long time engagement.

Our work explicitly focused on leaderboards, badges and the combination of the two. However, the research community would benefit from expanding these game mechanics to include the ones that are less popular. Such examples are levels and progress.

In systems potentially accessing personal data, like the one described in this work, there is a need to protect the private data from attacks, de-anonymization and exploitation beyond the explicitly allowed use. We believe that an important direction for future work is to explore the privacy by design approach to system design. There is also a huge potential in the analysis of opt out tools and the presence of more granular controls to manage what information of an employee is shared with his/hers colleagues. Having more means to protect one’s privacy can prove beneficial for engagement and mass adoption of such a tool.

Finally, we recognise the importance of conducting similar studies in other companies to see if the conclusions drawn by the current work are generalizable beyond the IBM use case.


Appendix A

Terms of Service

In this appendix we give an overview of the Terms of Service used in the experimental tool “How much of an IBMer are you”.

About the project
How much of an IBMer are you is a project developed at the IBM Center for Advanced Studies Benelux in collaboration with the Delft University of Technology. The aim of this project is to explore how games can affect employee engagement, learning and social behaviour. How can we facilitate learning among employees by building an enjoyable experience? How can we influence the social activeness of the employees? How can we better assess the knowledge and skills of employees? Our goal is to answer these questions through a scientific approach.

What data you provide us with?
When you click on Connect with LinkedIn button you provide permission to our application to access your basic profile information, including your headline, location and job description summary. The LinkedIn information is used for the proper functioning of the game. For example we use your name and location to identify you on IBM Connections, and to identify the IBMers part of your network in the game. Upon successfully logging in with LinkedIn, we ask you to confirm your IBM email. We will use your email only for two purposes: 1) to notify you about updates in the game, and 2) to support the game’s invitations system.

For a better social gaming experience, we also encourage you to connect your Facebook account. When you authorise our application with Facebook, you provide us access to information about basic profiles information, including the list of your friends. Again, we use this information to display the IBMers on your network in the game. We also look at your news feed to see when you shared IBM related news.

How we use your data?
Your data will be safely stored and used only and exclusively for research purposes, and will be accessible exclusively by the members of the research team. Your game activity will be anonymised and used only in aggregated form. At any time during the game, you will be provided with the option of opting-out of the system.
Appendix B

The questionnaire

In this appendix we illustrate the questionnaire which has been sent out to employees who used the experimental tool “How much of an IBMer are you?”.

General questions about the tool.

1. I have watched the general tutorial on how to interact with the tool.
   * Please choose only one of the following:
   ○ Yes
   ○ No

2. Please indicate your level of agreement with the following statements. *
   Please choose the appropriate response for each item:
   
<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I forgot about my immediate surroundings while playing on the website.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>This website was aesthetically appealing.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I found this website confusing to use.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
General questions about answering quiz questions in the tool.

[] I could earn points in the game by answering quiz questions. *
Please choose only one of the following:
- Yes
- No

[] I have watched the tutorial on the website on how to play the game. *
Please choose only one of the following:
- Yes
- No

[] Please indicate your level of agreement with the following statements. *
Please choose the appropriate response for each item:

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I felt motivated to answer quiz questions.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I have learned about IBM by answering questions in IBM Facts.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I have learned about Technology by answering questions in World Wide Technology.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I have learned about my colleagues by answering questions in You &amp; Your Network.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
[A] Please indicate what was your motivation for playing the game.

Please write your answer here:
General questions about the News module in the tool.

[ ] I could earn points in the game by sharing news from the tool on Linkedin or Facebook. *

Please choose only one of the following:

- Yes
- No

[ ] I could earn points in the game by posting IBM related content which does not originate from the tool on Linkedin or Facebook. *

Please choose only one of the following:

- Yes
- No

[ ] I have watched the tutorial on the website about how to use the News module. *

Please choose only one of the following:

- Yes
- No
Please indicate your level of agreement with the following statements.*

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I felt motivated to share news from the website.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The news in the tool were Interesting for me.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I shared IBM related information on my Facebook network before using the tool.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I shared IBM related information on my LinkedIn network before using the tool.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I shared news stories on my Facebook profile from the game.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I shared news stories on my LinkedIn profile from the game.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I share more IBM related links on Facebook after using the tool than I used to share before.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I share more IBM related links on LinkedIn after using the tool than I used to share before.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Please indicate what was your motivation for sharing news.

Please write your answer here:
General questions about the Invitation module in the tool.

[ ] I could earn points on the website by inviting my colleagues to play. *

Please choose only one of the following:

- Yes
- No

[ ] I have watched the tutorial on the website about how to use the Invitation module. *

Please choose only one of the following:

- Yes
- No

[ ] Please indicate your level of agreement with the following statements. *

Please choose the appropriate response for each item:

I felt motivated to invite my colleagues to play.

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

[ ] Please indicate what was your motivation for inviting friends.

Please write your answer here:
Feedback.

[ ] How should we improve the game to keep you engaged?

Please write your answer here:

[ ] May we approach you for participating in other research projects? *

Please choose only one of the following:

- Yes
- No
Appendix C

The user interface

In this appendix we illustrate the user interface of the experimental tool “How much of an IBMer are you?".

Figure C.1: The initial login page.
The user interface

Figure C.2: Second step of the signup process.

Figure C.3: The Choose Category view.
The user interface

Figure C.4: An example of answering a quiz question.

Figure C.5: The Invitation view.
The user interface

Figure C.6: The Trending Topics view.

Figure C.7: An example of displaying notifications to the user.
The user interface

Figure C.8: The Leaderboards view.

Figure C.9: The Badges view.