Using game mechanics for field evaluation of prototype social applications: a novel methodology

(Article begins on next page)

26 November 2024

Using game mechanics for field evaluation of prototype social applications: a novel methodology

Amon Rapp[a], Federica Cena[a], Cristina Gena[a]*, Alessandro Marcengo[b] and Luca Console[a]

[a]Computer Science Department, University of Turin, Turin, Italy; [b]Telecom Italia – Research & Prototyping, Turin, Italy

This paper describes a novel methodology to evaluate a social media application in its formative phase of design. Taking advantage of the experiences developed in the Alternate Reality Games, we propose to insert game mechanics in the test setting of a formative evaluation of a prototypical social system. As a use case, we present the evaluation of WantEat, a prototypical social mobile application in the gastronomical domain. The evaluation highlighted how the gamification of a field trial can yield good results when evaluating social applications in prototypical status. From a methodological point of view, gamifying a field trial overcomes the cold start problem, caused by the absence of active communities, which can prevent the participation of users and therefore the collection of reliable data. Our experience showed that the gamification of a field evaluation is feasible and can likely increase the quantity of both browsing actions and social actions performed by users. Based on these results, we then are able to provide a set of guidelines to gamify the evaluation session of an interactive system.

Keywords: gamification; field studies; social applications

1. Introduction

In this paper, we aim at illustrating a new method for evaluating social applications, that is, interactive systems with social features that allow people to interact with each other and contribute to the system's contents. In order to evaluate a social application, an active community motivating users in interacting and performing social actions is required. The lack of it usually prevents generating sufficient user engagement for gathering reliable data on the application's social features. This may happen when a prototype is under evaluation: in this case, it is often impossible to have a meaningful social context in which participants can interact among each others and use the social features of the system (Persson, Blom, and Jung 2005).

In this situation, common evaluation methods, such as laboratory experiments and field studies, fail. On the one hand, laboratory tests are too artificial (Lew and Nguyen 2011) for creating a believable social experience for users to perform social actions. On the other hand, field evaluations, in order to be effective, have to last for a long time in order to allow the emergence of a spontaneous community of users: this could be so expensive in terms of time, efforts, and costs that it could be unrealisable for a prototype application (Rogers et al. 2007), by requiring a stable system 'deployed on a platform that would be widely spread among users in order to achieve a critical mass of users' (Persson, Blom, and Jung 2005). We will

describe in the following the evaluation of a social application prototype, WantEat (Console et al. 2013), which we unsuccessfully previously tried to evaluate both with a usability test and a field trial.

The goal of this work is to find a proper methodology for evaluating a social application at an early stage of deployment, through providing a meaningful social context for motivating users to use the social features of the application even in the absence of a real social community of users.

In order to reach this goal, we decided to exploit the opportunity offered by *gamification*, that is, the use of game design elements in non-game contexts (Deterding et al. 2011), which showed to be a valuable option in enhancing the user experience in applications and services (Hamari, Koivisto, and Sarsa 2014), in order to promote a playful experience that could encourage social participation and foster the emergence of a spontaneous social community in a field evaluation performed in a limited time span.

This paper will proceed as follows: after a discussion of the current evaluation methods used in testing social applications (Section 2), we illustrate WantEat, the social application we used to test our methodology, summarising the previous evaluations of this application in order to highlight the issues in testing social features in a prototype (Section 3). We, then, propose our solution for a gamified evaluation (Section 4), describing the evaluation

of WantEat and its results, and we conclude the paper by illustrating some guidelines for introducing game elements in interactive system evaluation (Section 5).

2. Background and theoretical problems in evaluating social features

In this section we provide a brief account of the fundamental principles presented in the paper. We first highlight how traditional evaluation techniques seem to be unsuitable for testing social applications (Section 2.1). Then, after introducing the term 'gamification' and its actual role in human–computer interaction (HCI), we describe how to introduce game mechanics in field evaluation methodologies (Section 2.2). For each discussed topic, we also cite the most relevant work in the literature.

2.1. Evaluating social applications

While the evaluation of interactive systems has a long tradition and a variety of available methodologies, for example, usability testing (Dumas and Redish 1994), experience sampling method (Consolvo and Walker 2003), and living laboratories (Intille et al. 2006), some problems emerge when what is to be evaluated is a social application.

According to Kim, Jeong, and Lee (2010), social applications are usually classified as *social networking applications*, that is, website or smartphone applications that allow users to stay connected with other people in online communities; and s*ocial media applications*, that is, website or smartphone applications that allow people to create, upload, post, tag, comment, and share user-created content (UCC).[1] The distinction between the two types is quickly fading, as social networking sites started to add features of social media sites, that is, the sharing of UCC, and social media sites are adding features of social networking sites, such as personal profiles and managing communities (Ellison and Boyd 2013).

Social applications enable a wide variety of social behaviours mediated by technology that are deeply related to people's everyday habits: they are meaningful only if they are connected with a social context made up of interpersonal ties, audiences, and pre-existing relationships, as the desire to communicate and share content with other people or friends is a primary driver of social applications (Ellison and Boyd 2013). For these reasons, traditional evaluation techniques, such as *experimental evaluations* and *field studies*, do not seem appropriate to test them, especially during a *formative evaluation phase*,[2] where the application is in a prototypical status and hence it does not already have an active community of users.

Experimental evaluations create an artificial context that does not usually allow users to be motivated to perform social actions in the normal way. They are widely adopted in the HCI field, but they were criticised as they lack realism in the four dimensions of appearance, content, task, and setting (Lew and Nguyen 2011). In other words, they could be missing 'ecological validity', which describes how closely the appearance, content, methods, and setting of the experiment approximate a real-life situation. This could threaten both their internal and external validity (Lew and Nguyen 2011).

In the *ubiquitous computing* field, experimental evaluations were moreover criticised, as the strong link between the ubiquitous systems and the physical contexts in which they are used makes it difficult to use the traditional usability evaluations carried out in laboratories (Abowd and Mynatt 2000). The situated nature of ubicomp systems emphasises the need of conducting evaluations that investigate how technologies are used by people in their everyday contexts (Rogers et al. 2007). This issue is even more pressing for *social applications* where the user actions take place in a social context. In fact, the experimental protocol can force participants to use certain functionalities, allowing the experimenter to gather data on all the aspects of the system that she considers relevant. However, it is not sufficient to analyse the social behaviour of users simply in term of cause–effect relationships between variables. Social actions require intrinsic motivations (Paulini, Maher, and Murty 2014) that can only be engendered if these actions can obtain a social recognition in a social context: for example, the use of a commenting feature can make sense only if there is already a comment from other users or if this comment could receive a response in a short time frame from someone else. Lampe and Johnston (2005), for example, show how users who received replies to their first comment in a social community took less than a third of the time than those who did not to post a second comment. These kinds of social rewards are not easily replicable in a laboratory setting, even using a Wizard of Oz methodology (Kelley 1983) to simulate the reaction of an active social network.

Field studies, on the other hand, take place in contexts that are similar to those of the everyday life and thus they seem to be more suitable to assess social applications: they observe and record what people do for their own situated purposes, making available data on users' real social habits (Rogers 2011). This is essential in evaluating the social actions triggered by a social application, as these actions are meaningful only if they are performed in a real social context. Field studies, recently, also became known as 'evaluations in the wild', to highlight the need to be conducted in situ, leaving users free to use an application without constraints and for their own situated purposes (Rogers 2011). Users are left free to interact with the application while their activities are logged, as in the evaluation of CenceMe, a mobile social network application (Miluzzo et al. 2008), and of iCITY DSA, a social adaptive website in the domain of cultural events (Gena et al. 2013). However, the lack of well-defined tasks leaves less control to the researchers, not allowing them to gather data on specific critical points of an application (Kjeldskov et al. 2004). So,

when the community of a social application is still inactive, its social features can be ignored by users, as they appear meaningless to them.

The evaluation session should last for a long time (entire months or even years) to allow the emergence of a spontaneous active community, as happens using the Living Lab methodology, where services are validated in 'collaborative multi-contextual empirical real-world environments' (Eriksson, Niitamo, and Kulkki 2005). For example, Hess and Ogonowski (2010) evaluate the build-up process of household networks, observing users in their home transformed in a Living Lab for three years. Alternatively, it is possible to extrinsically motivate users, creating artificial rewards (Brown, Reeves, and Sherwood 2011): in the Tiramisu Field Trial, for example, an evaluation in the wild of a crowd-sourcing computer system was conducted, paying users to reach the needed amount of participants to create a social effect (Zimmerman et al. 2011). So, running these kinds of studies is very cumbersome in terms of efforts, costs, and time (Rogers 2011), especially in relation to the short-lived nature of the deployed prototype (Korn and Bødker 2012).

We think that gamification can help to solve this issue, by inserting game mechanics in a field evaluation format. This idea leads us to the following research questions (RQs) motivating our work:

- RQ1: Can gamifying a field evaluation be a feasible and efficient solution in testing a prototype social application, being conducted without high costs and long-term studies?
- RQ2: Can the gamification of a field evaluation motivate participants in using the social features of a social application prototype even when a community of users is not already active?

2.2. HCI and gamification

In the field of HCI, the term gamification is well known and refers to the use of elements borrowed from video games domain with the aim to improve the user experience in non-recreational applications and services (Deterding et al. 2011). These elements can be ascribed to visual components employed in video game interfaces (e.g. the progress bar in LinkedIn), reward systems (e.g. the level climbing in Yahoo Answers), and features for recognising the users' status (e.g. the badges in GetGlue). Academic research provided successful examples of gamification: Barata et al. (2013) showed that introducing points, levels, badges, and leaderboards in an academic course can enhance student involvement and participation; Flatla et al. (2011) showed that it is possible to introduce the enjoyable experiences of games in calibration tasks, without drastically transforming the nature of these tasks; and Cechanowicz et al. (2013) highlighted that it is possible to obtain greater gamification effects combining different game elements.

However, it seems that HCI is not yet exploiting the overall possibilities that game mechanics offer (Laschke and Hassenzahl 2011). In particular, the possibility of using gamification practices within the process of designing interactive systems did not receive sufficient attention. While some methods such as role-playing, make-believe, and design games are methodologies that have been available to designers for a long time (e.g. Brandt and Messeter 2004; Seland 2006), the use of game mechanics during the evaluation stages of interactive systems has not been fully investigated.

Game elements have showed to be useful in enhancing real-life situations. In particular, the experiences from alternate reality games (ARGs) can suggest to us that a situation, which is initially experienced as difficult and burdensome, can become engaging by introducing some game elements. ARGs such as World Without Oil[3] and Lost Joules[4] exploit game mechanics in order to create new organisational practices, allow people to imagine different worlds, and change people's behaviours (Michael and Chen 2005). Often ARGs target a specific area of personal life and try to improve it by introducing mechanics borrowed from video game world, as in Chore Wars[5] and Quest to Learn.[6]

From these and other similar practices, we can see how games can encourage people to participate more actively, in a self-motivated and self-directed manner, making them intensely interested and genuinely enthusiastic (Rigby and Ryan 2011). For this to work, it is necessary to provide goals, interesting obstacles, challenges, and a well-designed continuous feedback system (McGonigal 2011).

Taking advantage of the experiences matured in the ARGs field, our idea is to insert game mechanics during the formative phase of evaluation of prototype social applications. This will allow us to create a format for evaluating prototype social applications in a limited time span and without high costs, performing a field study in a setting that is very close to the context of everyday use (RQ1). It will also allow us to motivate users in performing social actions, creating a meaningful context in which to perform social actions (RQ2).

3.   A mobile social media application: WantEat

We tested our methodology on WantEat (Console et al. 2013), a social application prototype in food domain. We decided to gamify the here described field evaluation, since we tried to evaluate WantEat social features by means of a usability test and of a field trial but without success. WantEat is an example of a social application which allows both social network activities and UCC activities: providing different types of social features that we were not able to test through traditional evaluation methodologies, it represents an optimal use case for testing our idea of applying gamification to a field evaluation.

WantEat is an intelligent mobile application in the gastronomy domain, which aims at putting together real and virtual worlds. By means of WantEat it is possible to make everyday objects smart and able to communicate with users. Objects can be not only gastronomic items, such as food products, market stalls, restaurants, shops, and recipes, but also geographic places and actors, such as cooks, producers, and shop owners. WantEat is based on the idea that smart objects could be gateways for enhancing the interaction between people and a territory and its cultural heritage.

WantEat introduces a novel and peculiar paradigm for supporting the user interaction with a social network of smart objects. Such interaction is made of two main phases: (i) getting in touch with the object and (ii) sharing information with the object and exploring its social network of 'friends' (both people and other objects connected with it).

Getting in touch. WantEat supports an interaction with everyday objects with no embedded electronics or tags. The contact between a user and an object can be created by taking a picture of the label of a product with the phone's camera (Figure 1(b)); geopositioning the user in a specific place, making contact with the objects related to that place; getting a recommendation; and searching or exploring bookmarks (Figure 1(a)).

Interacting with the object and its world. Once a contact with an object has been established, the user can interact with it and access its social network. The interaction model (conceived as a 'wheel') allows users to explore the network starting from a contacted object (Figure 1(c)). The object is in the centre of the wheel and the user can get in touch with it by simply touching it. The object tells the user about itself, providing both general knowledge and information synthesised from the interaction with people (including tags, comments, and ratings) (Figure 1(d)).

The object in focus is surrounded by a wheel that provides access to the social network of its friends (Figure 1(c)). Each friend belongs to one of four sectors; the partition into sectors depends on the object in the centre. In the example in Figure 1(c), the object in focus is a kind of cheese; the first sector 'Territorio' (Territory) contains the friends related to the territory, the production, and supply chain (e.g. producers, shops, and production places). The sector 'Persone' (People) contains people who are friends of the object in focus (e.g. people who bookmarked it or who wrote a comment on it); the sector 'Prodotti' (Products) contains other products that are connected with the object in focus (e.g. a wine that goes well with a cheese); and the sector 'Cucina' (Cuisine) contains entities related to cuisine, such as restaurants, recipes, and so on. The user can continue the exploration of the network by changing the object in focus (Figure 1(e)). This can be done by simply dragging an item towards the wheel miniature in one of the corners (Figure 1(f)). At this point the whole wheel is recomputed and displayed to the user.[7]

Thus, the user can perform *browsing actions,* which we called *'display the object and its network'* (1(c) and 'more info about the object' 1(d)); or she can tell something to the object (thus, perform some *social actions*): she can add tags, comments, and ratings or can bookmark the object (Figure 1(g)).

In the past, we conducted evaluations of WantEat but none of them allowed us to evaluate its social features. We will briefly describe them in order to point out the challenges we encountered in evaluating a prototypical application with social features.

During the deployment of the application, we carried out a *first evaluation in the form of* a *usability test*[8] on an early prototype. While we gathered useful insights on the usability of the system's interface (Marcengo et al. 2012), we could not assess the acceptability of its social feature because the experimental protocol was perceived as too artificial by the users for performing any meaningful social behaviour. Users could only imagine the consequences of their social actions since there was not a fully running community of users that could provide them with
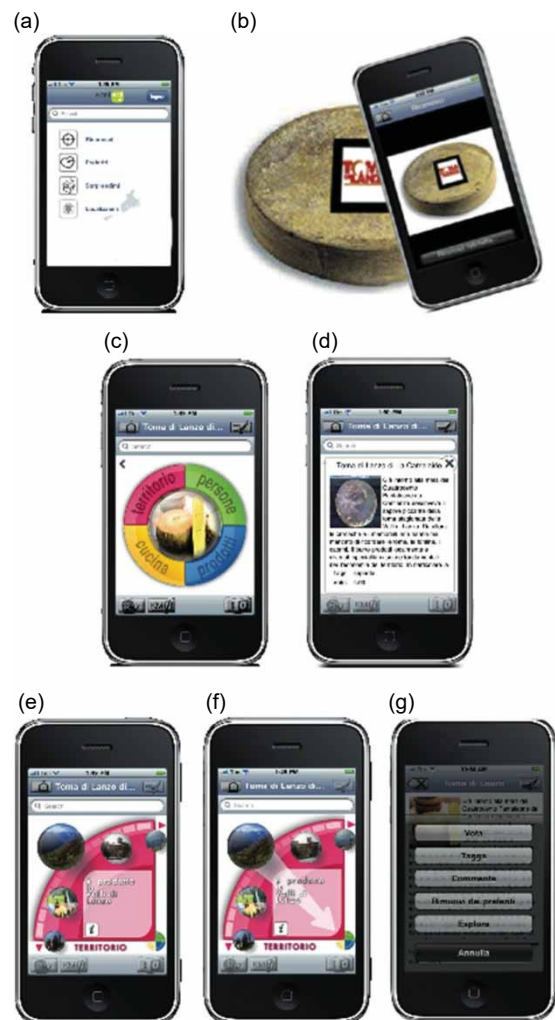

Figure 1. Example of the wheel on an iPhone.

real feedback. The tasks provided also forced participants to use the social features of WantEat in order to evaluate them, but these tasks were perceived as abstract and external, resulting in vague and superficial responses which were considered highly unreliable.

Later on, we carried on a *second formative evaluation* as a *field study*[9] on a high fidelity prototype in order to gather data about the usage of WantEat in a real context of use, because of the failure of the previous evaluation. We conducted the trial at Salone del Gusto 2010, the biggest food fair in Italy that attracts about 200,000 paying visitors, hosted every two years in Turin, Italy. However, due to the lack of an active community of users and formal tasks that forced successfully the usage of its social features, users mostly performed few social actions during the evaluation, as it will be detailed in the following. The 'cold start problem' (Salton and McGill 1984) prevented users from being really involved in the study.[10] Users did not participate because they did not perceive an immediate usefulness of their social actions, as they could not see to the effects of their social behaviours. The result was that we did not have sufficient data to perform a reliable analysis on the users' social actions. The social features of the application were perceived as useless or meaningless and often they were not even noticed by the users.

Therefore in order to solve these problems, we conducted a *gamified field evaluation*, as described in the next section, providing a comparison with this field evaluation.

4. A gamified field test

The novel occasion to test our approach was the international food fair Cheese.[11] This exhibition, held in the Town of Bra in Piedmont, Italy, attracts about 300,000 visitors. We aimed at stimulating the spontaneous usage of WantEat in all its aspects and in a context as close as possible to a real-use situation via game mechanics. To this aim, we set up a *gamified field evaluation*.

The goal was to offer a gamified test experience completely merged with the fair visiting experience that could allow us to gather more reliable data related to the use of the application's social features, compared with those gathered during the previous field trial.

4.1. Method

First of all, in order to motivate visitors, a viral event[12] was created based on an ad hoc developed format. During the four days of the fair, WantEat was installed on the users' iPhones with their permission. Testers could use it in whatever way they would like and for as long as they desired. The high degree of freedom given to users, however, was balanced by means of 'game missions' that substitute the formal tasks of the laboratory setting with engaging and playful objectives to accomplish. Each participant received a leaflet with game instructions and a map of the fair

with highlighting of the areas in which the application was working. Inside the exhibition, 10 items (cheeses) were selected as the focus of the game. They were located in different areas of the fair, and were marked as recognisable by the application. The main objective of the game was to recognise with the application (by taking a picture) at least five cheeses, 'taste' them, and perform some social actions using the social features of the application. Each social action (e.g. leaving a comment, tagging, and rating) allowed the player to earn 500 points: at the achievement of 6000 points, she was awarded a T-shirt with the application logo.

A live leaderboard at the installation base informed all participants of their current position, showing them the distance from the user with the highest score during the four-day fair.

The prize was an incentive for users to return to the installation point, where, simultaneously with the delivery of the T-shirt, they were interviewed. The semi-structured interview aimed at assessing the user experience with the WantEat app.

Besides the basic actions, participants could accomplish some special but more cumbersome missions, which allowed them to earn additional points (10,000 or 20,000 points). The special missions aimed at stimulating cooperation between participants and fostering the exploration of all the locations of the fair. Users were encouraged to exchange a special identification coin (which was provided with game instructions) using the application communication features (Mission 1), or to discover some secret places that could be recognised by the application (Mission 2). The accomplishment of these extra-missions was not required in order to obtain the T-shirt, since we wanted to encourage everyone to play, offering a relatively easy goal to reach (6000 points) and, at the same time, optional objectives that could motivate more serious players, that is, users who wanted to excel in the game, to win the collective challenge.

4.2. Sample

In the four-day event, 157 participants attended the trial. Participants were recruited in the same manner that we recruited users at Salone del Gusto 2010, by asking them whether they owned a smartphone and, in case of a positive response, whether they wanted to try a novel application for the Cheese fair.

They were not informed of the game format (and of the rewards available) in advance. The information about the goals of the evaluation and the opportunity to win a T-shirt were provided only after they accepted to participate in the trial and installed the application on their smartphone.

The sample was smaller than the one of Salone del Gusto as we had fewer resources (fewer interviewers and only 160 licences for installing the app on users' iPhones).

4.3. Results

We aimed at analysing user actions on the app and, in particular, social actions. To this aim, during the experiment we collected usage data and analysed them by means of descriptive statistics on log data and then we also performed correlations in order to explore the relationship between the variables of interest. Moreover, we performed interviews in order to reach a deep and qualitative understanding of the usage of the application and its features.

4.3.1. Analysis of user actions

In the four-day event, 157 people attended the trial and installed the system on their smartphones. Of the 157 users, 110 (70.06%) have actively participated in the evaluation. In total they have interacted with 102 objects, with an average of 19.4 objects per user (STD= 25.65). In particular, 72 users (64.63%) interacted with more than the required 5 objects, with an average of 28.08 objects per user and STD of 28.08, and about the 40% of subjects (namely 45) interacted with more than the 10 selected objects (the 10 cheeses recognisable by the system), with an average of 40.53 objects per user (STD= 29.13). Moreover, excluding the entry points scattered through the fair (the 10 cheeses that had to be photographed), around 40% of total active users (45 out of 110) browsed the other 92 objects in the app (other cheeses, the producers linked to the cheeses, etc.).

The users totally performed 2134 actions on the app: 1584 (74.23%) were browsing actions, 547 (25.63%) were social actions, and 3 (0.14%) were deletion actions (i.e. Delete Bookmark action). As we have noticed earlier, users explored several contents and thus made a lot of *browsing actions*, and in particular:

- 1109 (70.23%) actions were actions like 'display the object and its network' (Figure 1(c));
- 470 (29.76%) were like 'more info about the object' (Figure 1(d)).

These last actions were the entry point for performing social actions. In all, 102 out of 110 different users (93% of total active users) asked for more info about the object, and 70 users (63.64% of total active users) made some sort of social actions. We have defined these 70 users as *contributing* users, because they rated, commented, shared bookmarks, and tagged, bringing contributions to the application contents. In particular, regarding the distribution of the 547 social actions:

- 226 (41.33%) were ratings;
- 127 (23.2%) were comments;
- 120 (21.94%) were bookmarks;
- 74 (13.53%) were tags.

Looking in more detail at these contributing users, we obverse that users have made 7.81 social actions on average (STD =9.94), and these were mostly distributed on more than one object (28.89% of users made one social action per instance, 29.33% made two social actions per instance, and 41.78% made more than two social actions per instance). In all, 17 users (24.28%) made social actions greater than or equal to 12 (the minimum amount required to obtain the prize), and thus they won the prize.

In regard to the special missions, two users completed the first mission and one user completed the second mission.

However, if we consider the total number of social actions, we may notice that:

- 66 users (60% out of total 110 active users) have rated 226 times, with an average of 3.42 ratings per user (STD =3.24), and with 11 users (10% out of total) covering more than 50% of rating actions;
- 32 users (26.9% out of total active users) have added 127 comments, with an average of 3.96 comments per user (STD =3.15). In particular, 8 users (7.27% out of total) inserted 69 comments (53% out of total comments);
- 35 users (31.8% out of total active users) have added 120 objects to the shared bookmark list, with an average of 3.42 bookmarks per user (STD= 3.41), and 7 users (6.36% out of total) have made more than one half of add/share bookmark actions (54.5%);
- 19 users (17.3% out of total active users) have inserted in total 74 tags, with an average of 3.89 tags per user (STD =3.07). Among them 5 users (4.54% out of total) inserted almost one half of the tags (49.62%).

Looking at these results, we may notice that the game mechanics had a positive impact on users' contributions (63.64% of users made some sort of UCC activities). This is especially true if compared to the '90-9-1' Rule formulated by Nielsen,[13] which states that the majority of users (90%) just consume content produced by others and a small set (9%) provides small contributions every now and then, while a very small fraction of users (1%) accounts for most of the UCC. Nielsen's Rule is consistent with Preece et al.'s study (2004) which claims that in online communities, content consumers (*lurkers*) outnumber content producers (*posters*). Thus, we can claim that gamification seems to promote user participations and the production of UCC activities.

However, we still have to distinguish the *kind of contributions* the users have made: 29.9% of users wrote comments, a time-consuming action, since it requires a higher cognitive effort than tagging – about 4.85 (STD = 3.60) average word per comment per user wrt 1.15 (STD =0.50) average tags per user in our dataset – and the difficulties are increased by the fact that it is

performed with a smartphone. Most users (60%) simply rated items and 31.8% added items to the bookmark list, which are both soft contributing actions.[14] If we look at the data in more detail, we discover that few users (10% for ratings, 7.27% for comments, 6.36% for bookmarks, and 4.45% for tags) have made more than one half of the total social actions, thus somehow approaching on average the ratio between heavy and soft contributors and consumers. In our evaluation, according to Nielsen's expression, the average ratio is 36.36–56.36–7.27, so 36.36% of users (namely 40) have been simply consumers, 56.36% have been soft contributors (namely 62), and 7.27% have been heavy contributors (namely 8).

Another observation has to be provided. Rating and bookmarking actions were very frequent, and this is not surprising since they are less time-consuming than other social actions. They can be also defined as *micro-contributions*, which are less cumbersome UCC activities like rating and tagging. According to Frankowski et al. (2007), micro-contributions may motivate casual contributors, who can be very important to the community, and may provide a path for a casual contributor to become a heavy one. A more recent study (Wichmand and Jensen 2012) in the context of network games has shown that 'by developing and expanding our understanding of engagement to include micro forms of engagement and by giving the users the possibility to be recognized for their micro deeds, we are able to create more thriving and creative online communities'. Thus micro-contribution in user communities may stimulate lurkers to become more active and participative, and it is not surprising that micro-contributions, which are effort saving, are performed more than other more demanding actions.

As a final consideration, we highlight that even if tagging may be considered as a soft contribution, the subjects of our evaluation inserted very few tags (74 tags inserted by the 17.3% of users), and in general tagging was the least performed action. Surprisingly, users have made more comments than tags, and from the analysis of

tags it emerges that most tags look like subjective tags ('good', 'tasty', 'sweet', etc.) than organisational tags (tags that identify personal items) (for details, see Xu et al., 2006). Another point is that in the context of our application, the action of tagging is an end in itself, and this is in contrast with the usual user tagging that is required for achieving other goals (e.g. tags are requested before adding bookmarks, and before publishing pictures, videos, and documents).[15]

4.3.2. Correlational analysis

To obtain a broader view of the user social behaviour on the application, we have considered the co-occurrence of social actions per single user (spread on the overall interaction and not on the same object), and we have calculated the Pearson correlation coefficient since scores showed a normal distribution (see Table 1 for all the details). We considered just the correlations significant at the 0.01 level, and thus we have highlighted these correlations between users' actions:

- all users who commented very frequently rated ($r = 0.835$), frequently added bookmarks ($r = 0.732$), and also added tags ($r = 0.671$);
- all users who added bookmarks frequently rated ($r = 0.741$), commented ($r = 0.732$), and added tags ($r = 0.713$);
- all users who rated very frequently commented ($r = 0.835$), frequently added bookmarks ($r = 0.741$), and also added some tags ($r = 0.591$);
- all users who tagged frequently added bookmarks ($r = 0.713$), comments ($r = 0.671$), and rates ($r = 0.591$).

The reader may notice that all the users who commented often made other social actions, in particular rating and adding bookmarks, thus they have frequently made both heavy and soft contributions. All the users who added

Table 1. Correlations.

		Comments	Preferences	Rates	Tags
Comments	Pearson Correlation	1	0.732**	0.835**	0.671**
	Sig. (two-tailed)		0.000	0.000	0.000
	N	110	110	110	110
Preferences	Pearson correlation	0.732**	1.0	0.741**	0.713**
	Sig. (two-tailed)	0.000		0.000	0.000
	N	110	110	110	110
Rates	Pearson correlation	0.835**	0.741**	1.0	0.591**
	Sig. (two-tailed)	0.000	0.000		0.000
	N	110	110	110	110
Tags	Pearson correlation	0.671**	0.713**	0.591**	1.0
	Sig. (two-tailed)	0.000	0.000	0.000	
	N	110	110	110	110

**Correlation is significant at the 0.01 level (two-tailed).

bookmarks frequently made all the other social actions. All the users who rated often commented and shared bookmarks. All the users who tagged frequently added bookmarks, and this is not surprising since users are often required to tag when they add bookmarks on social media sites, and they less frequently performed the other social actions.

4.3.3. Comparison with prior field evaluation

As described earlier, in our study users seem to be motivated more than usual to participate and to perform social actions through the social features of the application. In this gamified evaluation, participants performed more social actions than those accomplished in the field trial at Salone del Gusto 2010, where the game mechanics were not present. During this field trial, 675 were active users (98.67%) out of 684 total participants,[16] while just 74 contributing users (10.96%) performed 179 (3,84%) social actions out of 4660 total actions.

The gamified evaluation with 70 (63.64%) contributing users on 110 active users allowed users to perform 547 social actions (25.63%) on 2134 total actions.

The distribution of the 179 social actions (performed at the Salone del Gusto) was:

- 84 (46.93%) ratings;
- 47 (26.26%) bookmarks;
- 33 (18.44%) tags;
- 15 (8.38%) comments;

thus confirming the tendency of users to prefer micro-contribution actions, even more than in the Cheese evaluation.

Applying Nielsen's rule at this past evaluation, we have found the following result: 89.04–10.67–1.63 corresponding to 601 consumers (89.04%), 63 soft contributors (9.33%), and 11 heavy contributors (1.63%), substantially in accordance with the proportions proposed by Nielsen.

Although the two trials were not conducted in the same event, the environmental conditions were similar: the trials took place in a big food fair attended by thousands of people, the WantEat application under evaluation was substantially the same, and users could use it freely for as long as they liked. Participants could recognise with a smartphone a limited set of items that were clearly highlighted as recognisable by the application in different marked food stands. Furthermore, participants were recruited in the same manner, by asking them whether they wanted to try a novel application, as in 'Cheese' evaluation they were not aware of the opportunity to win a T-shirt and that they are participating in a game before formally accepting to participate.

However, we can highlight some differences that could affect the comparison between the two datasets, by introducing some confounding variables: in the field evaluation

at Salone del Gusto 2010, users received an iPhone with the app pre-installed and could use it in a limited area of the fair (only in the area of the Piedmont food products, about 3000 square metres, in one of the four pavilions in which the fair took place), while at Cheese 2011 the app was installed directly on the users' phone and they could use it in different areas scattered through the fair. Furthermore, in the evaluation at Salone del Gusto 2010, the experimenters were closer to the participants than at Cheese 2011 evaluation. At Salone del Gusto, if participants wanted, the experimenters could accompany the participants in their first attempts in interacting with the application, while at Cheese 2011, the participants were left completely free to use the system after a short demo at the installation base (which was far away from the areas where the system worked).

Furthermore, the Salone del Gusto food fair was a paying sheltered event, while Cheese 2011 was a free open-air event, held in the streets of the Town of Bra.

Although the comparison between these two datasets does not assure the effectiveness of gamification in increasing user participation during a field trial, as the first evaluation does not represent a proper control group for the other one due to the differences we highlighted, our results represent a cue of the goodness of our format, which will require additional testing for proving its validity. However, as long as our main aim was to generate more social actions than those performed in the evaluation at Salone del Gusto 2010, these results could be interpreted as a success of the gamified field trial.

4.3.4. Qualitative interviews

We also performed 37 qualitative interviews in order to gather more insights about the use of the applications and the user engagement. We interviewed all the participants who returned to the installation base to withdraw the prize or to spontaneously report their own experience with the system.

Participants reported a high degree of engagement in using the application. It emerged that people wanted to expand the social features of the application not in the direction of connecting with new friends, but rather to express and share their views on food, to discover new products and recommend them to other people, or to build groups with others interested in the same products. Using the application, the participants also became aware of the potentialities of WantEat, expressing the desire of interacting directly with producers as a way to gain a personal relationship between consumers and producers. During the interviews the participants reported also the problems they encountered with the app during the study,[17] allowing us to fix them in the redesign of the prototype.

Participants highlighted how the presence of specific objectives motivated them in commenting and tagging: however, gaining the promised prize was not their final

aim. Users showed they were motivated by the structure of the game that allowed them to compete in order to stand out among the other participants. For example, a participant completed all the missions provided by the researchers only for 'finishing the game'. During the accomplishment of this challenge, she started to communicate with other users through the social features of the application also for sharing personal comments about the food she tasted. Another participant reported that he was motivated to leave detailed comment on cheese since he thought that they were useful for someone else. Another participant stated that she replied to several comments about a product posted by other users since she thought that they were wrong.

We also have to acknowledge how the presence of a well-integrated task structure with the fair experience motivates the users to participate correctly in the game. The comments posted and the other actions required by the game were always coherent: users did not use any mechanism for circumventing the rules (i.e. 'gaming the system'; posting empty, out of context, or nonsense comment; or repeating the same action, which would also earn points for the participant, was never performed). During the interviews, participants confirmed that they were not pushed only by gaining points, without considering the quality of their contributions: they wanted to post useful comments, tags, and ratings that could be useful for other people. These findings suggest that the goals and missions provided were not perceived as abstract and artificial, and instead motivated the users to participate.

In conclusion, the interviews confirmed that the game format of the evaluation was able to generate an engagement of participants, who performed social actions for accomplishing their own goals. From this perspective, the findings were more reliable than those from the previous evaluations, allowing us to assess the WantEat social features and to gather useful insights for its improvement.

4.4.   *Summary of the main findings*

The main conclusions that emerged from the results of our evaluation can be summarised as follows.

We showed how the gamification of a field study of a social application prototype can be conducted in a short time span and without high costs: in only four days the gamified format produced a large amount of data at a minimum expense by building up a meaningful social context in which participants could perform their social actions.

In fact, regarding the effectiveness of gamification in creating a user community from scratch in a field evaluation and, thus, motivating participants in using the social features of a given prototype social application, our research pointed some interesting cues of the goodness of the gamified format. Compared to the previous field trial we conducted at Salone del Gusto 2010, where the social actions were 3.84% out of total, the gamified evaluation

at Cheese 2011 produced 25.63% of social actions out of total actions. As emerged also from the interviews, users were deeply engaged in the use of the application in all of its aspects: the game mechanics favoured in users the willingness to use every WantEat functionality, allowing us to gather reliable results about the usage of the application in a real context of use.

So, we were able to answer all of the RQs presented in Section 2:

. RQ1: gamifying a field evaluation can be a feasible and efficient option when testing a prototype social application, by not requiring high costs or long-term studies to be conducted.
. RQ2: by means of gamification, a field evaluation of a prototype social application can motivate users in using its social features, by giving them a meaningful social context in which to perform their social actions.

Regarding the second RQ, we still have to emphasise that we are not able to assure that the insertion of game mechanics in a field evaluation format can increase users' participation in respect to a non-gamified field study, as our study lacked of a proper control group. However, the increase in the number of social actions performed by participants in the Cheese 2011 field study, compared with those performed during the Salone del Gusto 2010 field study, provides some good cues about the effectiveness of the solution. In addition, the data gathered through the qualitative interviews confirmed that participants were engaged by the game structure and motivated to perform the social actions required by the study.

We think that our study highlighted how this point is worthy of further investigation in future study.

5.   Discussion and guidelines

In this paper we wanted to show how it is possible to introduce gamification within the assessment methodologies of interactive systems. The Cheese field test highlighted how the gamification of a user test can yield good results when evaluating social application prototypes.

From a *methodological point of view*, gamifying a user evaluation helps with overcoming the cold start problem, caused by the absence of active communities, which can prevent the participation of users and therefore the collection of reliable data during prototype testing. Game mechanics can motivate users to participate, by quickly generating an active community from scratch, in which users can find a meaningful context to perform their social actions.

The role of game mechanics within user tests can reasonably be extended to all cases in which applications, interactive systems, or websites require motivated users' active participation. It is possible to think that the

gamification of an evaluation session may also be a good remedy to the artificiality of the laboratory setting. Even if the game situation does not recreate the daily experience of using the application, which often takes place in not-playful situations, it could push user participation and provide practical objectives that are similar to those experienced by people during their everyday life.

From our experience, we can provide a set of *guidelines* to gamify the evaluation session of an interactive system.

First, the gamification of a user test does not have to necessarily provide monetary or material incentives: unlike usability testing, where a gadget or a monetary compensation is almost always provided, in a gamified testing experience it is sufficient to provide the proper status reward to motivate users to participate. In our case, the T-shirt, according to the participants themselves, was a tangible reward but it was not the main reason for participating in the game: in the interviews, participants reported that the self-motivation to gain the objectives planned in the game format had a far greater influence.

Second, the gameplay should be well balanced with regard to the difficulty of achievable goals: on the one hand, too great difficulty would make the experience frustrating; on the other hand, if the completion is too easy to accomplish, it would make the experience boring. Providing incrementally complex achievable goals, structured by means of easier mandatory objectives and optional more challenging missions, provides the necessary motivation to all users to reach a minimum level of participation and, if necessary, to continue the game experience if they like it. For example, in our case the minimum of 6000 points to get the T-shirt prize was accessible by all (17 users, 15.45%, reached the quota), and special missions were directed only to the most motivated players (two users, 1.82%, completed the first mission and 1 user completed the second mission).

In addition, careful definition of the missions of the game should promote cooperation and sharing among the participants. Creating differentiated targets that leverage on the competition, but that are needed for their fulfilment of a social cooperation between users, is the right way to get active participation through the application under test (the first mission provided in our format included cooperation among users for its completion).

Still, the importance of the gaming stage in the design of an engaging gaming experience should not be underestimated: the articulation of sub-areas in which sub-objectives can be achieved and the balance between exploration and the control component on the surrounding spaces must never be missing. We, for example, designed to distribute the recognisable cheese in many sub-areas, some clearly marked on the map provided with the information leaflets, and others hidden and therefore to be discovered.

Finally, researchers should investigate, through an interview session, whether the application under evaluation produced some significant problems during the evaluation. The log analysis is not sufficient to produce unambiguous data. To make sure that gamification did not obscure a problem with the app under test, researchers should dig deep into the reasons that pushed participants in using the app (for example, users could have used a given application only because the gamification was fun, but not because the application was engaging). For this reason, a qualitative investigation, made of contextual interviews or focus groups, is necessary to clarify some aspects of the quantitative data gathered.

In conclusion, we still want to emphasise that the process of gamifying a field trial does not consist in adding a layer made up of leaderboards and points to a context that is already self-contained. To gamify an evaluation session means to make it engaging and fun by adding game mechanics carefully designed for a specific context. Game mechanics, that is, the 'rules of the game', the different actions and behaviours afforded to the player within a game context (Hunicke, LeBlanc, and Zubek 2004), need to be well balanced and provide specific goals that users can achieve on their own during the gaming experience. In addition, the expected reward systems should be adapted to the type of situation to gamify, so that the playful side merges with the 'serious' side of the experience, without being perceived as a superficial or merely external element. The design of challenges and game rules should therefore be carried out with an accurate understanding of the context of the experience to be gamified and a careful selection of the game mechanics suitable for that context. Understanding the context comes before the design of the gamified evaluation itself. Only in this way it is possible to create a deeply engaging experience that is necessary for effective gamification.

Conflict of interest disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. In this paper we focus on evaluating UCC activities in a social application. In the rest of the paper, we will use the term 'social features' to refer to these activities.
2. Formative evaluations are done during design phases for checking that a prototype of an application meets users' needs (Preece et al. 2002).
3. http://www.worldwithoutoil.org/
4. http://lostjoules.com/
5. http://www.chorewars.com/
6. http://q2l.org/
7. For more details on system architecture and the involved software components, see Console et al. (2013).
8. The test was conducted through five tasks that addressed users to navigate and search information in the prototype of the application. The evaluation has involved 12 users. The variables considered for the composition of the sample were 'Age' (18–35 and 36–60 years) and 'Experience and openness to technology' (Hard users, who daily used technological tools, such as videogame consoles, personal computers, applications on their smartphones, and Soft users, who did not daily use technological tools). From the intersection of these

two variables, four groups of three users each were formed. The test lasted one hour and was conducted in our laboratories.

9. The field study (Console et al. 2013; Marcengo et al. 2012) was conducted on a sample recruited by asking participants to try a novel application for the fair. On the basis of our previous usability test, we segmented the sample in four subgroups through the dimensions of age and attitude towards the new technologies of communication. The total sample has been of 684 users surveyed, divided into 228 Young/Hard users, 114 Old/Hard users, 175 Young/Soft users, and 167 Old/Soft users. Every user received the app pre-installed on an iPhone and could try it freely for as long as she liked in a limited area of the fair.

10. The 'cold start problem' is the lack of data at the beginning on an interaction with an interactive system, which prevents the system to effectively work (Salton and McGill 1984). In our case, it consisted in a lack of an active social community of real users.

11. http://www.cheese.slowfood.it/

12. A viral event is an event that exploits existing social networks by encouraging users to share information related to the event with their friends (Leskovec et al. 2007; Odén and Larsson 2011).

13. http://www.nngroup.com/articles/participation-inequality/

14. We define heavy and soft contributions according to the cognitive load required to the users (commenting is a more cumbersome activity than simply tagging).

15. We evaluated also some WantEat features (temporal navigation patterns, typology of user actions, and typology of tag), but this is beyond the scope of this paper.

16. Users at Salone del Gusto evaluation used our smartphones, and were sometimes accompanied by the experimenters during the trial. This probably justifies the higher percentage of active users with respect to Cheese.

17. The report of the data of the interviews regarding this point is beyond the scope of this paper.

References

Abowd, G. D., and E. D. Mynatt. 2000. "Charting Past, Present, and Future Research in Ubiquitous Computing." *ACM Transactions on Computer-Human Interaction* 7 (1): 29–58.

Barata, G., S. Gama, J. Jorge, and D. Gonçalves. 2013. "Improving Student Creativity with Gamification and Virtual Worlds." In *Proceedings of Gamification'13*, 95–98. New York: ACM.

Brandt, E., and J. Messeter. 2004. "Facilitating Collaboration through Design Games." In *Proceedings of the Eighth Conference on Participatory Design. Artful Integration: Interweaving Media, Materials and Practices – Volume 1 (PDC 04)*, Vol. 1, 121–131. New York: ACM.

Brown, B., S. Reeves, and S. Sherwood. 2011. "Into the Wild: Challenges and Opportunities for Field Trial Methods." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 1657–1666. New York: ACM.

Cechanowicz, J., C. Gutwin, B. Brownell, and L. Goodfellow. 2013. "Effects of Gamification on Participation and Data Quality in a Real-World Market Research Domain." In *Proceedings of Gamification 2013*, 58–65. New York: ACM.

Console, L., G. Biamino, F. Carmagnola, F. Cena, E. Chiabrando, R. Furnari, C. Gena, et al. (2013). "Interacting with Social Networks of Intelligent Things and People in the World of Gastronomy." *ACM Transaction on Interactive Intelligent Sytems (TIIS)* 3 (1), article 4. doi:10.1145/2448116.2448120.

Consolvo, S., and M. Walker. 2003. "Using the Experience Sampling Method to Evaluate Ubicomp Applications." *IEEE Pervasive Computing Mobile and Ubiquitous Systems* 2 (2): 24–31.

Deterding, S., D. Dixon, R. Khaled, and L. Nacke. 2011. "From Game Design Elements to Gamefulness: Defining 'Gamification'." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek '11)*, 9–15. New York: ACM.

Dumas, J. S., and J. C. Redish. 1994. *A Practical Guide to Usability Testing*. Norwood, NJ: Ablex.

Ellison, N. B., and D. Boyd. 2013. "Sociality through Social Network Sites." In *The Oxford Handbook of Internet Studies*, edited by W. H. Dutton, 151–172. Oxford: Oxford University Press.

Eriksson, M., V. P. Niitamo, and S. Kulkki. 2005. "State-of- the-art in Utilizing Living Labs Approach to User-centric ICT Innovation – A European Approach." White paper, December 15. http://www.vinnova.se/upload/dokument/Verksamhet/TITA/Stateoftheart_LivingLabs_Eriksson2005.pdf.

Flatla, D. R., C. Gutwin, L. E. Nacke, S. Bateman, and R. L. Mandryk. 2011. "Calibration Games: Making Calibration Tasks Enjoyable by Adding Motivating Game Elements". In *Proceedings of UIST'11*, 403–412. New York: ACM.

Frankowski, D., S. K. Lam, S. Sen, F. M. Harper, S. Yilek, M. Cassano, and J. Riedl. 2007. "Recommenders Everywhere: The WikiLens Community-Maintained Recommender System." International Symposium on Wikis 2007, Santa Barbara, CA, 47–60.

Gena, G., F. Cena, F. Vernero, and P. Grillo. 2013. "The Evaluation of a Social Adaptive Website for Cultural Events." *User Modeling and User-Adapted Interaction* 23 (2–3): 89–137.

Hamari, J., J. Koivisto, and H. Sarsa. 2014. "Does Gamification Work? – A Literature Review of Empirical Studies on Gamification." In *Proceedings of the 2014 47th Hawaii International Conference on System Sciences (HICSS '14)*, 3025–3034. Washington, DC: IEEE Computer Society.

Hess, J., and C. Ogonowski. 2010. "Steps Toward a Living Lab for Socialmedia Concept Evaluation and Continuous user-Involvement." In *Proceedings of the 8th International Interactive Conference on Interactive TV&Video (EuroITV '10)*, 171–174. New York: ACM.

Hunicke, R., M. LeBlanc, and R. Zubek. 2004. "MDA: A Formal Approach to Game Design and Game Research." In *Proceedings of the Challenges in Game AI Workshop, 19th National Conference on Artificial Intelligence (AAAI '04)*, 1–5. Menlo Park, CA: AAAI Press.

Intille, S., K. Larson, E. Tapia, J. Beaudin, P. Kaushik, J. Nawyn, and R. Rockinson. 2006. "Using a Live-in Laboratory for Ubiquitous Computing Research." In *PERVASIVE 2006*, edited by K. P. Fishkin, B. Schiele, P. Nixon, and A. Quigley, Vol. 3968, 349–365. Heidelberg: Springer, LNCS.

Kelley, J. F. 1983. "Natural Language and Computers: Six Empirical Steps for Writing an Easy-to-Use Computer Application." Unpublished doctoral diss., Johns Hopkins University.

Kim, W., O.-R. Jeong, and S.-W. Lee. 2010. "On Social Web Sites." *Information Systems* 35 (2): 215–236.

Kjeldskov, J., M. Skov, B. Als, and R. Høegh. 2004. "Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field." In *Mobile Human-Computer Interaction – MobileHCI*, edited by S. Brewster, M. D. Dunlop, Vol. 3160, 61–73, Heidelberg: Springer, LNCS.

Korn, M., and S. Bødker. 2012. "Looking Ahead: How Field Trials Can Work in Iterative and Exploratory Design of Ubicomp Systems." In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*, 21–30. New York: ACM.

Lampe, C., and E. Johnston. 2005. "Follow the (Slash) Dot: Effects of Feedback on New Members in an Online

Community." In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work (Sanibel Island, Florida, USA, November 06–09, 2005). GROUP '05*, 11–20. New York: ACM.

Laschke, M., and M. Hassenzahl. 2011. "Being a 'Mayor' or a 'Patron'? The Difference between Owning Badges and Telling Stories." In *Proceedings of Gamification: Using Game Design Elements in Non-gaming Contexts Workshop HOLD at the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, Vancouver, BC, 72–75.

Leskovec, J., L. A. Adamic, and B. A. Huberman. 2007. "The Dynamics of Viral Marketing." *ACM Transactions on the Web* 1 (1): 5.

Lew, L., and T. Nguyen. 2011. "Solomon Messing, and Sean Westwood. 2011. Of Course I Wouldn't Do that in Real Life: Advancing the Arguments for Increasing Realism in HCI Experiments." In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*, 419–428. New York: ACM.

Marcengo, A., A. Rapp, L. Console, and R. Simeoni. 2012. "Evaluating WantEat: A Social Network of People and Objects." In *Advances in Usability Evaluations. Part 2*, edited by F. Rebelo and M. Soares, 493–502. Boca Raton, FL: CRC Press.

McGonigal, J. 2011. *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. London: Penguin.

Michael, D., and S. Chen. 2005. *Serious Games: Games That Educate, Train, and Inform*. Boston, MA: Thompson.

Miluzzo, E., N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. 2008. "Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application." In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys '08)*, 337–350. New York: ACM.

Odén, N., and R. S. Larsson. 2011. "What Makes a Marketing Campaign a Viral Success? A Descriptive Model Exploring the Mechanisms of Viral Marketing." *Umea Universitet*, Faculty of Social Sciences, Department of Informatics. Accessed December 16, 2014. http://www.diva-portal.org/smash/get/diva2:433110/FULLTEXT01.pdf.

Paulini, M., M. L. Maher, and P. Murty. 2014. "Motivating Participation in Online Innovation Communities." *International Journal of Web-Based Communities* 10 (1), 94–114.

Persson, P., J. Blom, and Y. Jung. 2005. "DigiDress: A Field Trial of an Expressive Social Proximity Application." In *Proceedings of the 7th International Conference on Ubiquitous Computing (UbiComp'05)*, edited by Michael Beigl, Stephen Intille, Jun Rekimoto, and Hideyuki Tokuda, 195–212. Berlin: Springer-Verlag.

Preece, J., B. Nonnecke, and D. Andrews. 2004. "The Top Five Reasons for Lurking: Improving Community Experiences for Everyone." *Computers in Human Behavior* 20 (2): 201–223.

Preece, J., Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey. 2002. *Interaction Design: Beyond Human Computer Interaction*. New York: John Wiley & Sons.

Rigby, S., and R. M. Ryan. 2011. *Glued to Games: How Video Games Draw Us in and Hold Us Spellbound*. Santa Barbara, CA: Praeger.

Rogers, Y. 2011. "Interaction Design Gone Wild: Striving for Wild Theory." *Interactions* 18 (4): 58–62.

Rogers, Y., K. Connelly, L. Tedesco, W. R. Hazlewood, A. Kurtz, and R. E. Hall. 2007. "Why It's Worth the Hassle: The Value of In-Situ Studies When Designing Ubicomp." In *Proceedings of Ubicomp 2007*, Innsbruck: ACM.

Salton, G., and M. McGill. 1984. *Introduction to Modern Information Retrieval*. New York: McGraw-Hill Book Company.

Seland, G. 2006. "System Designer Assessments of Role Play as a Design Method: A Qualitative Study." In *Proceedings of the 4th Nordic Conference on Human-computer Interaction: Changing Roles (NordiCHI '06)*, 222–231. New York: ACM.

Wichmand, M., and S. S. Jensen. 2012. "Small is Beautiful – Lurkers Engaging Through Micro-Contributions." In *Proceedings of ECREA 2012, 4th European Communication Conference*, 24–27 October 2012, Istanbul.

Xu, Z., Y. Fu, J. Mao, and D. Su. 2006. "Towards the Semantic Web: Collaborative Tag Suggestions." In *Proceedings of the Collaborative Web Tagging Workshop at WWW 2006*, Edinburgh, Scotland. Accessed May 2008. http://www.semanticmetadata.net/hosted/taggingws-www2006-files/13.pdf (2006).

Zimmerman, J., A. Tomasic, C. Garrod, D. Yoo, C. Hiruncharoenvate, R. Aziz, N. R. Thiruvengadam, Y. Huang, and A. Steinfeld. 2011. "Field Trial of TIRAMISU: Crowd-Sourcing Bus Arrival Times to Spur Co-Design." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 1677–1686. New York: ACM.