Pervasive Games Field Trials: Recruitment of Eligible Participants through Preliminary Game Phases

Vlasios Kasapakis, Damianos Gavalas and Nikos Bubaris

Department of Cultural Technology and Communication
University of the Aegean
Mytilene, Greece

{v.kasapakis, dgavalas@aegean.gr}, nbubaris@ct.aegean.gr

Abstract. This paper investigates a novel methodological approach in performing user evaluation trials for pervasive games. The evaluation process of a pervasive role playing game, *Barbarossa*, is used as a case study. *Barbarossa* involves a preliminary and a main execution phase. The former is freely available to anyone and may be played anytime/anywhere requiring no organizational and orchestration investments from the investigators team. The latter defines three inter-dependent player roles acted by players who need to collaborate in a treasure hunting game to achieve a common game goal. The eligibility of players for participating in the main game phase derives from those ranked relatively high in the preparatory phase. Drawing on concepts of cultural theory, we design the preparatory phase as an affective environment out of which the potential evaluators will emerge. The main hypothesis investigated is that the execution of such cost-effective preparatory phases may serve as a means for recruiting highly qualified subjects for user trials on pervasive game research prototypes, thereby increasing the reliability and quality of evaluation results. This hypothesis has been validated through the user evaluation trials performed on both the *Barbarossa* game phases.

Keywords: Pervasive games; evaluation; evaluator subject; field trial; Android.

1. Introduction

The evaluation of experimental systems and applications in relatively unconstrained settings outside of the laboratory, wherein groups of 'real' users are assigned 'real' tasks, represents a growing trend in computer and cognitive sciences. Field trials provide a rich source of information for HCI researchers since the users experience the systems in their 'natural' operational context. Hence, beyond investigating technical flaws and malfunctions, field trials offer insights on user understandings, perceptions, practices and the eventual uses of evaluated systems [14]. Field trials are even more important in testing ubiquitous systems, whose operation largely depends on sensing (and adapting to) the environmental and social context; these aspects are difficult to 'simulate', hence, they limit the reach of classical laboratory studies [1, 4, 25].

Field trials for evaluating pervasive computing systems engage considerable human resources (e.g. for trial organization activities, system programming to log detailed usage data, trial monitoring, carrying out user interviews, authoring user experience reports, etc), involve large groups of evaluators (to ensure the statistical validity of results) and commonly run over prolonged periods of time[6, 11, 24]. Namely, these trials are too expensive, labor-intensive and time consuming to fail. Recent studies have discussed methodological challenges in running user trials, documenting several key issues that influence the trials' outcome [6]. For instance, adjustment of users' behavior to fit the expectations of the investigators; potential social relationships among the participants; the way that trials are run (e.g. questions asked by investigators). Oddly, the HCI literature lacks reports on the impact of the evaluators themselves (i.e. the means of recruitment and the composition of the participant groups) on the overall quality of results derived from field trials (validity, reliability and transferability, among others).

Field trials are particularly relevant to pervasive games. Those refer to an exciting new form of computer games which extends the gaming experience out to the physical realm, building upon a combination of hybrid interfaces, mobile device equipment, wireless networking, positioning systems and context-sensing technologies [22]. This emerging gaming mindset is rather challenging not only for developers who explore the boundaries of available technologies, but also for HCI researchers and cognitive scientists who investigate methods to achieve high quality interactive experience for users. Along this line, evaluation trials are crucial in assessing the usability, playability and quality of experience aspects of pervasive game prototypes. As a result, field trials deliberately oriented towards capturing those factors (rather than technical feasibility aspects) become increasingly commonplace in pervasive games literature [8, 23, 29].

The focal objective of evaluation trials is to receive unbiased feedback from neutral, 'external' subjects about the usability and experience perceived throughout the game sessions. The recording of neutral views expressed by the evaluators may indicate technical flaws or even suggest essential script, usability or technical improvements. Sadly, trials participants do not equally contribute to providing useful feedback. A small atypical subset of users (referred to as 'lead participants') commonly offer particular insight into the behavior under investigation [6]. A reasonable interpretation of this imbalance in feedback contribution could be that not all evaluators are equally interested in and engaged with the tested system. Likewise, not all evaluators are prospective users of the system when this becomes fabricated and launched to the market. This assumption implies that it would be more meaningful to consider recruiting field trial participants from the pool of the system's potential target group. Such qualified evaluators would presumably be more self-motivated, committed in exploring the offered features of the system, keen to report faults and offer constructive feedback, thereby enhancing the reliability of evaluation findings and maximizing the usefulness of field trials.

We argue that the above suggested principle to narrow the pool of potential field trial participants among those provably sharing interest on the system particularly holds in the case of games. This argument rests on the specificities of gamers as a cultural group. Gamers, be it classical computer game, pervasive game or even non-digital game enthusiasts, are like-minded individuals establishing cohesive

communities (gamer groups) each resembling a tribe, i.e. possessing distinctive ethnographic elements [7]. This is substantiated by the findings of a recent study which revealed strong correlation of evaluators' attitude towards pervasive elements in games depending on their demographic profile or whether they have been gamers or non-gamers [28]. It should be meaningful then, to recruit pervasive game evaluators among individuals sharing this unique cultural identity. Dedicated gamers will be expected to engage with game scenarios and actively participate in the evaluation process, besides being more experienced, therefore eligible to assess the assets and weaknesses of the game. Such participant qualities should be even more critical when considering pervasive game scenarios which involve several interdependent game roles and require relatively long-term commitment to complete. In such cases, potential player dropouts may severely impede the whole evaluation process, frustrate affected co-players and require additional investments. Participants with strong interest on pervasive games are more likely to behave responsibly and not obstruct the playing of the game, hence, the execution of field trials. Nevertheless, safeguards are needed to ensure the objectivity of participants in expressing their views about the game, namely the reliability of trials results.

Herein, we propose a novel methodological approach in performing user evaluation trials for pervasive games. The main hypothesis investigated is that a competitive preparatory game phase, acting as a first layer of participation, would help in designating a group of players with shared common interest into the game itself. Intuitively, such a 'qualification round' would populate a pool of game-aware players, eligible to act as potential participants in the main game phase (in effect, the main investigated 'end product'). The main game phase could serve as a more exclusive layer of participation and, practically, comprise the 'official' field trials of the game. To maximize cost-effectiveness, the preparatory phase should be able to run without any serious organizational and orchestration requirements, i.e. engage minimal labor and/or financial investments. Furthermore, this phase should be open to anybody interested and address the widest possible audience so as to ensure the demographic, cultural and geographic diversity of participants, in addition to facilitating reach of players socially unconnected with the investigators (hence, presumably neutral towards the game). In this manner, the preparatory phase could be utilized as an effective participant recruitment tool, reducing dropouts and ensuring smoother execution of pervasive games evaluation trials; besides, it may offer insights over a large, representative sample of the game's potential target group, thereby highlighting future marketing opportunities. Furthermore, we investigate the case where dedicated game 'tribe' members invite individuals (socially related with them) to join the game. The hypothesis here is that such recruitments of inexperienced 'guest' players, recommended by provably committed players, may also serve as an effective tool to attract qualified evaluators. We argue that the above described methodological approach is applicable and transferable to other ubiquitous systems possessing similar characteristics, i.e., those that involve inter-dependent roles and require long-term commitment.

The above detailed hypotheses have been tested in the evaluation of *Barbarossa*, a pervasive role-playing game. Herewith, we document experiences from running a trial of both the preparatory and the main phases of our game; yet, the key objective of the trials has not been to assess our system but rather to validate our research

hypotheses. Along this line we opted to implement a freely available preparatory game mode in the Android application market (Google Play) to serve as the first layer of participation, reaching a large number of potential participants and also enabling players worldwide to participate into the game, thereby ensuring openness and neutrality in the recruitment process. Finally, we developed a main game phase, which served as more exclusive layer of participation; the participants of the main game phase have been selected amongst the most committed and interested players based on their performance in the preparatory phase of *Barbarossa.*

Playing a pervasive game is a multifaceted activity that situates game experience within an everyday socio-cultural environment. In this sense, cultural studies may help to better understand issues that concern the perceptual practices of the gamers. On the other hand, the abstract conceptual and methodological tools that cultural studies deploy seem discouraging for pervasive game designers who seek to employ applied principles and variables in their work [27]. At this crossroad, finding ways to connect such theories and concepts with pervasive game design models is a crucial and a challenging endeavor.

The remainder of this paper is structured as follows: Section 2 provides a conceptual framework for designing field trials by drawing on contemporary cultural theory. Section 3 presents the *Barbarossa* game scenario with Section 4 providing technical implementation details. Section 5 presents the methodology and results of the user evaluation trials. Section 6 presents related work and Section 7 concludes our work.

2. Linking field trials methods with cultural theory

Along with ubiquitous computing and game studies, cultural studies is a constructive contributor to the theory and practice of pervasive games [27]. Studies in this field mostly associate the embodied and spatial perceptions of the gamer, who moves in both digital and physical space, with broader phenomena such as urbanism and global mobility [12, 13, 15]. In this paper, we attempt to extend the contribution of cultural studies to the development of pervasive game theory and design by drawing on two central conceptualizations of contemporary cultural theory: (a) the performative production of subjects; (b) relational ontologies. Within the scope of this paper these conceptualizations are chosen because (a) they underpin and extend contemporary questionings in the design of field trial methods in ubiquitous computing, such as the dichotomy of natural/laboratorial setting, the degree of users' participation, the inclusion of contingencies, etc [6] and (b) they instigate a conceptual framework for deploying recruitment practices for the profile of gamers that a pervasive game, such as *Barbarossa*, requires (namely, long-term commitment, interdependency of actions, etc).

The performative production of subjects. The formation of identities is not a mere product of activating predefined properties of general categories. Rather, identity is performed through the actions of a subject in a particular setting. Performative actions, such as playing a game, are dynamic processes that foster contingent outcomes as they may strengthen or change existing ideas of identity formation and its related practices [5]. As far as the methods of recruiting gamers for evaluating a

pervasive game are concerned, the approach of the performative production of subjects will not focus on the formation of a representative sample, from the start, in accordance with the best suitable predefined features of general categories, such as age, sex, gamer's prior level of experience, etc. The performative approach will rather prioritise the formation of a setting out of which the desired attributes of the evaluator will emerge in practice. This setting is neither laboratorial nor "natural". This setting is being formed by the performative actions of playing the game itself. In other words, the group of evaluators is not pre-organised through the application of some external to the specific game-experience criteria. The group of the evaluators is self-organised as they respond to the environment of the game itself.

Relational ontologies. The approach of the performative production of the subjects contests the idea that there is a fixed set of variables that determine and can fully explain actions. It favors relational ontologies over essentialist ontologies. Being synthetic rather than analytic, this approach prioritizes interaction by foregrounding the mutual acknowledgement of the participating agents (both human and technological) not as distinct entities but as a coming-together situation [17]. In this sense, the design of evaluation methods is neither system-centric nor user-centric. The issues that concern usability and experience design are not reduced to some inherent properties of the system or the user but they are considered as emergent properties of their relation. Following this line of thought, the design process of evaluating *Barbarossa* put emphasis on the creation of the preparatory phase as an affective environment that could function as glue for the gamer-system interaction. This affective environment was designed so as to meet and foster crucial preconditions for playing a pervasive game, such as the sense of engagement, inter-dependency and long-term collaboration between the gamers themselves and between the gamers and the system. In doing so, typical methods for evaluation, such as questionnaires, interviews, log data, etc, are not dismissed; they are still useful tools for articulating results of interaction. However, these typical methods follow the relational process described above instead of preceding it.

The performative production of subjects indicates that evaluators are not chosen but they are becoming within the field trials. By actively performing the identity of a gamer in the affective environment of the preparatory phase of *Barbarossa*, some participants may develop to evaluators while some others may not. Both cannot be defined in advance. Becoming an evaluator is an event within a series of interactions between the gamer and the system. For an event to happen, a successful execution of interactions is needed. However, this is not sufficient. Affective responses, arousal of interest and desire for further participation are the catalysts. Therefore, an eventful approach for recruiting gamers for evaluation should favor the above motivations.

3. Game Scenario

The intuitions behind *Barbarossa* game scenario have been to design:

- an open and portable preliminary game phase scenario (i.e. to be played anytime, anywhere, by anyone), which will enable the designation of a narrow game-oriented group of players;

- a complex main phase scenario, thematically connected to that of the invitational phase, which compels the collaboration among interdependent roles and requires long-term player commitment.

Along this line, *Barbarossa* has been designed as a two-phase trans-reality role playing game. The first game phase is available from the Google Play app store under the title *The Conqueror*[1]. The overall scenario has been largely inspired by the medieval story of the pirate brothers Barbarossa, set in the Aegean Sea.

Preliminary game phase scenario

In the first phase scenario the Barbarossa pirate brothers Aruj, Khzir and Ilyas following a battle against the Knights of St. John outside the castle of Mytilene and assisted by some Traitors within the castle walls conquered the city. The Pirates then settled in the city center placing the Traitors around it. In the meantime, a group of Liberators plunder the areas around the city borders. In this phase the players act as Knights of the St. John who try to liberate the conquered city by eliminating their enemies. Acting so, the players use a mobile (Android) application which utilizes Google Maps and a turn-based role-playing game, which allows them to complete and create quests located into the surrounding area of Mytilene. Upon completing quests the players gain experience points that indicate their commitment and attribution to the game.

Main game phase scenario

The players ranked higher (in experience points) in the first phase are invited to participate in the second game phase called *The Interplay*; in the latter, Mytilene is liberated and the Knights rush into the castle to capture Aruj, Ilyas and Khzir. The three Pirates, being aware that the Knights are after them, hide all their treasures into a treasure chest somewhere in the crypts of the city castle. Aruj and Ilyas lock the chest with one combination lock each, while Khzir takes it and ridding his horse leaves the crypts to hide it. In a while, the Knights arrive at the crypts and manage to catch Aruj as a prisoner. The commander of the Knights charges two Guards to change shifts into moving Aruj around the city constantly so no one would know his exact location to try to set him free. The two Guards, being passionate gamblers, start to gamble theirs shifts using two dices. Finally, Ilyas manages to flee the crypts trying to find an ally Traitor to guide him out of the city while a Knight chases after him.

In order to complete the second game phase, the three players need to cooperate to unlock the treasure chest hid by Khzir somewhere in the city. Note that the *The Conqueror* application detects the distance of the players from the Mytilene center (located on Lesvos island, Greece) while playing, grouping players into two separate rankings categories, the *Insiders* who act in Mytilene (≤ 3.5 km from the city center) and the *Outlanders* who play away from the city, anywhere in the world. One of the *Outlanders* and one of the *Insiders* as well as a guest (invited into the game by the *Insider* player) are invited into the second phase based on their total experience points, collected in the first phase.

---

[1] The Android application files along with the full reports of Google Play about the traffic of The Conqueror and ManHunt (which is presented below) applications can be found at the official site of Barbarossa, www.BarbarossaRPG.com, in the APKs & Reports section.

The *Insider* acts the *Treasure Hunter* who pursuits the "*Treasure Hunt*" scenario in which she has to locate a real chest hidden somewhere in the city of Mytilene, locked with 2 combination locks (see Figure 1). The invited guest of the *Insider* player acts the *Pirate* who pursuits the "*Set Me Free!*" scenario trying to free Aruj from his guards so as to obtain the combination used by Aruj to lock the chest. Finally, the *Outlander* player pursues the "*ManHunt*" scenario acting as the *Knight* who chases Ilyas to kill him before Ilyas reaches his ally Traitor (who will guide him out of the city); by doing so, the *Outlander* obtains the second combination, the one that Ilyas used to lock the treasure chest.

Figure 1. *Barbarossa* Phase II Treasure Chest.

By completing their assigned missions (based on separate, yet, supplementary scenarios) the three players have to locate the locked chest and open it using the two obtained lock combinations. The complete game flow including the roles, scenarios and goals of Barbarossa is presented in Figure 2 below.
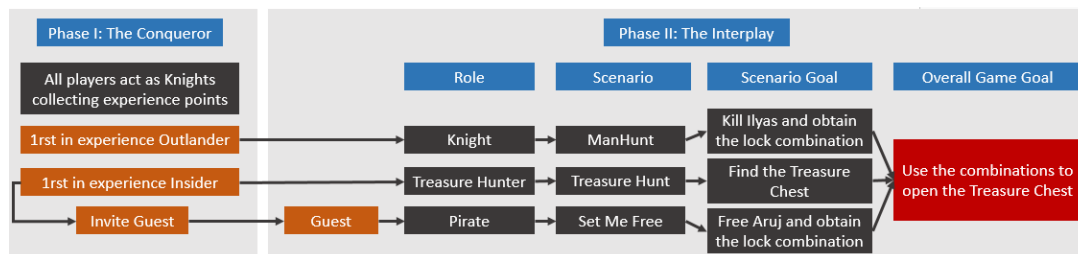

Figure 2. *Barbarossa* game flow.

4.   Game Implementation

4.1.   Phase I: The Conqueror

In the first phase of *Barbarossa* the players are expected to create and complete quests around the city of Mytilene in order to earn experience points, achieve high ranking in the game 'leaderboards' and secure invitation into the second phase. To participate, the players first create a personal game account. Upon logging into the game the players take control of a virtual *Knight* who possesses four (4) basic attributes (see Figure 3a). *Experience* is collected by completing quests and affects the damage that the *Knight* causes to her enemies. *Experience* is also the main ranking criteria among players. *Health Points* represent the damage the *Knight* can tolerate until she dies. *Potions* are used to recharge *Health Points* that the players loses when struck by her enemies. *Ratings* are the number of votes a player receives

from other players based on the quality of the quests she created. Finally, the current time is taken into account, on the basis that the players cause more damage to their enemies during nighttime (21:00 – 06:00), as they supposedly surprise their enemies while sleeping.

Quest Creation

Barbarossa suggests a user-generated content model wherein quests are created and consumed by the players. To create a quest, a player firstly assigns an enemy at a physical location (by tapping on a Google Map interface) within a given area around the city of Mytilene, which is divided into three sections denoted by concentric circles (see Figure 3b). To align with the game scenario, an enemy placed within the red circle would be a *Pirate* as pirates lived at the city center. An enemy placed in between the red and green circles is denoted as *Traitor* as those were placed around the city center by the pirates. Finally, enemies placed between the green and brown circles are *Liberators* who plunder the areas at the borders of the city.

Thereafter, the players are expected to fill in a title and a storyline about the quest. Players are advised to try connecting the storyline with the physical location where they placed the enemy. In their assistance, a tool called "Address" yields information about the enemy's location (by performing reverse geocoding and obtaining the address of the physical location, the tool performs a location-based Google search).

When the enemy position is set and the quest title and storyline are complete the player uploads the quest paying a *Health Potion* as a fee. All other players can then undertake the quest. A player completing a quest is requested to rate it. The votes of those players are transferred as *Ratings* to the creator of the quest and can be converted to *Health Potions* which are crucial in the fights against the enemies as they may be utilized to regenerate *Health Points*.

Quest Completion

To pursue a quest the player first obtains a list of the available quests around the city of Mytilene (see Figure 3c). Upon selecting one, its physical location is presented on the map. The players can save up to 3 quests into their local quest log and, thereon, they can pursue them to eliminate the enemies.

Upon starting a quest the game engine displays the player's location on the map via a *Knight* marker along with the corresponding enemy marker (e.g. a pirate icon if the enemy is a pirate). The title and storyline of the quest are also presented (see Figure 3d). Upon tapping the enemy marker, the fight against the enemy starts.

The position of the player is used, firstly, to determine whether the player is an *Outlander* or an *Insider*. If the player is an *Outlander* then the experience points that she gains by killing the enemies and completing the quests are the default. If, on the other hand, the player is an *Insider* the experience points earned upon completing a quest will be inversely proportional to her distance from the enemy. Tapping on the enemy marker, the game engine informs the player on the experience points she will earn upon completing the quest (see Figure 3e).

When the player starts the battle, a 2D turn-based fighting game launches. The player may utilize a variety of weapons and special attack patterns to eliminate the

enemy and complete the quest (see Figure 3f). Upon wining the fight, the player is requested to rate the quest using a 5 star rating scale (see Figure 3f) based on its quality (i.e., the title, storyline and relevance with the enemy's physical location). When the quest rating is uploaded, the experience of the player is updated, while she also receives a *Health Potion*.

(a)
(b)
(c)
(d)
(e)
(f)
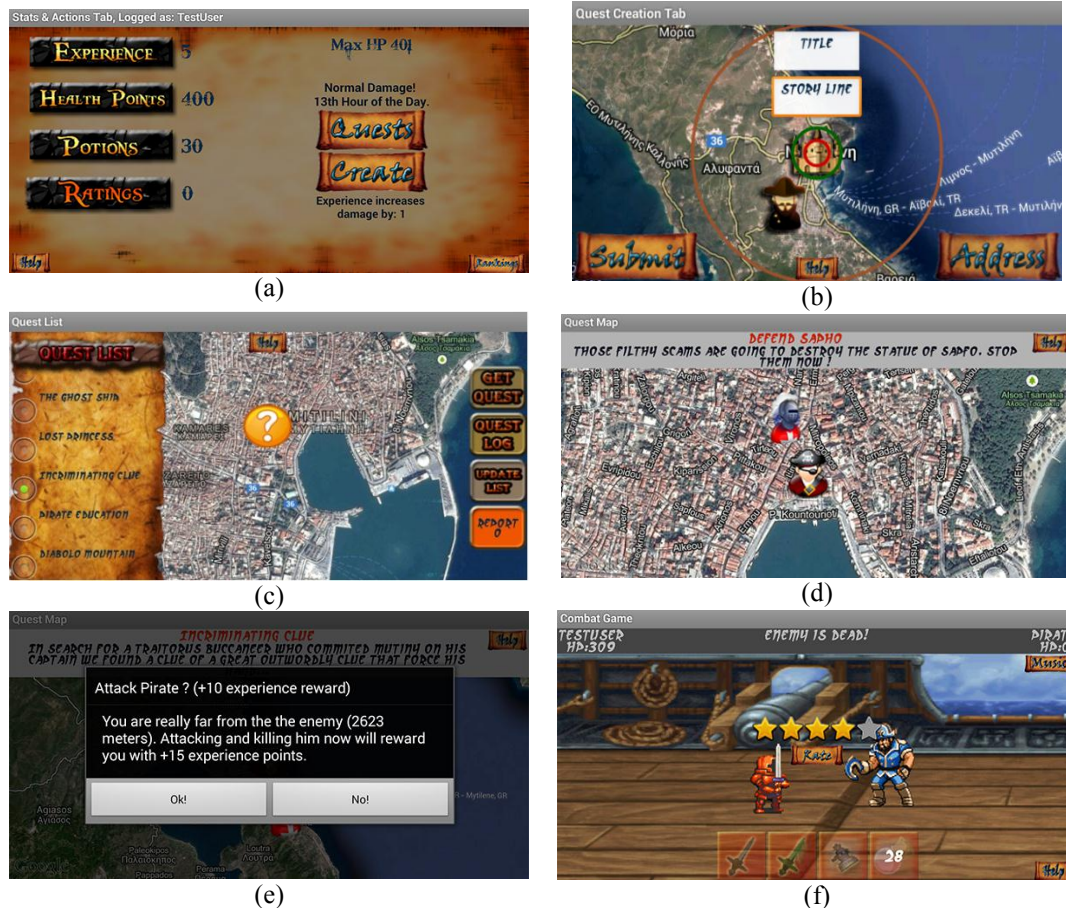
Figure 3. Screenshots taken from the *Barbarossa*: The *Conqueror* mobile (Android) client application**.**

4.2.   **Phase II: The Interplay**

The *Outlander* and *Insider* players ranked first in the first phase are contacted by the developers (via e-mail) and invited to participate in the second (main) game phase. Upon accepting the invitation, the developers introduce the *Outlander* and *Insider* players (exchanging their contact details), who are then free to use any communication means they prefer to contact directly with each other. A third player (a guest) is invited by the *Insider*. Furthermore, a meeting between the *Insider*, her guest and the developers of *Barbarossa* is arranged to brief the players on the second phase. All players are also provided information about the custom Android applications they have to use to pursue their scenarios. Below, we briefly discuss the implementation details relevant with the three player roles involved in the main phase of *Barbarossa*.

Treasure Hunt

The *Insider* player who is invited into the second phase pursues the *Treasure Hunt* scenario, aiming at finding the hidden treasure chest, utilizing a custom Android

application provided by the developers. The *Treasure Hunter* moves around the city of Mytilene scanning QR-codes and utilizing a variety of tools (like SunSPOT[2] sensor node devices) to reveal clues that will direct her towards the hidden chest. To scan a QR-code the player uses the QR-code scanner (zxing[3]) embedded into the custom Android application (see Figure 4a). Then, the game engine provides her with clues in the form of images and text.

For every scan the *Treasure Hunter* consumes energy points based on the average sound level of its environment while the scan is performed (the energy points subtracted are proportional to the sound level). To recharge energy points a *Treasure Hunter* can switch to 'sleep' mode and return to the normal operation mode in order to scan for a new clue. Finally, the *Treasure Hunter* can utilize *Hints* for assistance. *Hints* convey additional information about the clue the *Treasure Hunter* is currently on. The number of the *Hints* available to consume depends on the performance of the *Pirate* and *Knight* players, as explained below.

ManHunt

To pursue the *ManHunt* scenario and kill *Ilyas* before he flees the city the *Outlander* needs to install a custom Android application (see Figure 4b) available from Google Play[4]. When the player launches the application the game engine requests her location via GPS. When a GPS fix is acquired the player may start a new game session. The game engine selects two random points of interest (among many retrieved from Google Places service[5]) in proximity to the user (being at least 200m away from each other). Then, via a Google Directions web service[6] invocation, it retrieves walking directions in between the selected points of interest. The steps included into the service's response are utilized to animate the marker of Ilyas form the first towards the second point of interest (the latter is supposedly the intended escape point of Ilyas).

The moving speed of Ilyas is based on the environmental temperature at the player's location retrieved from the Yahoo Weather Service[7] (e.g., when cold, Ilyas moves fast to shorten the duration of the game session). To finalize the game session the *Knight* player has to chase Ilyas, approach him (<60m) and shoot him (before Ilyas reaches his escape point) using a cannon icon; subsequently, she recovers the combination lock.

Finally, a maximum of two *Hints* are transferred to the *Treasure Hunter* to assist her into finding the treasure chest. The number of *Hints* earned is proportional to the distance between Ilyas and his Traitor ally, so as to reward the player for her effort. Figure 4b presents the *ManHunt* application interface.

Set Me Free!

In order to free Aruj and recover the combination lock that unlocks the treasure chest, the guest player (invited by the *Insider* player into the second game phase)

---

2   http://www.sunspotworld.com/
3   https://github.com/zxing/zxing
4   Following the end of the evaluation trials, ManHunt has been freely available for anyone to play and enjoyed more than 8000 downloads.
5   https://developers.google.com/places/documentation/
6   https://developers.google.com/maps/documentation/directions/
7   https://developer.yahoo.com/weather/

should complete the *Set Me Free!* scenario, acting the *Pirate*. *Set Me Free!* involves three players.

Two *Guards* control the movement of Aruj rolling two dices, changing shifts when needed. Upon meeting with the players the developer undertakes the role of one *Guard* while the *Insider* player acts the second *Guard*. The two guards utilize a custom Android application called the *Shift Adapter* (see Figure 4c) where they roll two virtual dices. Upon the start of the game the two Guards take turns in rolling the dices until one of them scores a double. Then the other *Guard*, considered as loser, taps inside a specified area (represented by a circle) to set the position of Aruj. Next, the same *Guard* starts rolling the dices until she scores a double. When that happens, she hands over the application to the other *Guard* so that she can set a new position for Aruj. In this manner, Aruj changes position whenever the *Guards* change shifts.

Meanwhile, the outdoors *Pirate* player uses a custom Android application which utilizes GPS and augmented reality (AR) to indicate the position of Aruj and his distance from the player (see Figure 4d). The *Pirate* has to run after Aruj to approach him (<25m) and tap on Aruj AR marker to set him free from his *Guard* and recover the second combination lock of the treasure chest. Similarly to the *ManHunt* scenario, a maximum of two hints are transferred to the *Treasure Hunter* based on the speed recorded for the *Pirate* player while trying to fee Aruj.

Figure 4. Screenshots taken from the (a)*Treasure Hunt*, (b) *ManHunt*, (c) *Shift Adapter*,(d) *Set Me Free!* mobile (Android) client applications.

5. Evaluation

5.1. Methodology

So far, questionnaires, interviews and log data (i.e. data capturing the mobility and interaction activity of players throughout the game sessions) have been the methods most commonly employed in pervasive games evaluation [26]. The same

practice has been followed in *Barbarossa*. We have conducted user evaluation trials using all the three abovementioned evaluation methods; log data have been a critical element in the evaluation process in order to cross-check them against player answers (as compiled by questionnaires and interviews) and extract more safe and reliable conclusions. Below we describe the evaluation process in full detail. For the questionnaires we used likert scale and yes/no questions.

The evaluation process of *Barbarossa* commenced in October 29th, 2013 by releasing the invitational game mode though Google Play as well as a website wherein the players could check their rankings and search useful information and game instructions[8]. We provided players sufficient time (21 days) to play the first game phase; thereafter, we started contacting the highest ranked players among the *Outlanders* and the *Insiders* in order to form the 3-player teams required to proceed to the second game phase. We have invited one team at a time, a practice that enhanced competition among Phase I players who wished to participate in the second game phase. Ten teams (of 3) have participated overall in the main phase. On average, each team required around4days to complete the game. Note that none of the main phase participants has been socially related with any of the investigators.

Prior to proceeding to the second game phase we have asked two members of the 3-player teams that qualified from Phase I (i.e. those acting as *Treasure Hunter* and *Knight* in Phase II) to complete a questionnaire about their experiences in the first phase. Then, we introduced the players into the second game phase and allowed them a week to play the game session and collaboratively locate and unlock the treasure chest. Having completed the second phase, we have invited all players to complete an additional questionnaire tailored to the scenario they pursued in the second game phase. Finally, each player has been interviewed about her overall game experience. In parallel, we have collected detailed log data about the player's game actions throughout the game sessions(e.g. total completed and created quests, game session duration, total distance covered, average speed).

Currently (November 2014) *Barbarossa* features more than 1700 downloads and an average rating of 3.96/5 in Google Play. Furthermore, 941 players actually registered in *Barbarossa*; according to the log data, 290 among them performed at least one in game action, such as obtaining a quest.

5.2.   Evaluation results

The assessment of the overall, generic impression of the players with respect to the preparatory phase of the game comprised the first step of our evaluation process. As illustrated in Figure 5 all players reported a positive position towards *Barbarossa* indicating their appreciation towards the game scenario and their keenness in playing again.

---

8   http://www.BarbarossaRPG.com
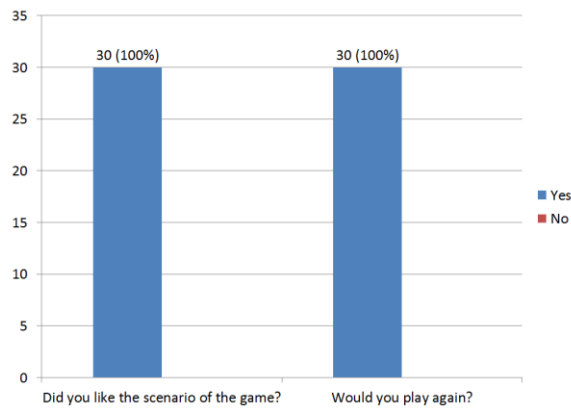
Figure 5. Players'attitude towards the game scenario and willingness for future participation.

Figure 6 suggests that all the players enjoyed participating in *Barbarossa*, at least to an average extent. This feedback is supported by the fact that, without exception, all *Insiders* and *Outlanders* players accepted our invitation to the join the second phase of the game. Moreover, none of the main game phase players, including the - invited by the *Insiders- Pirate* players, dropped out of the game, while all the second phase teams completed successfully their game goals.
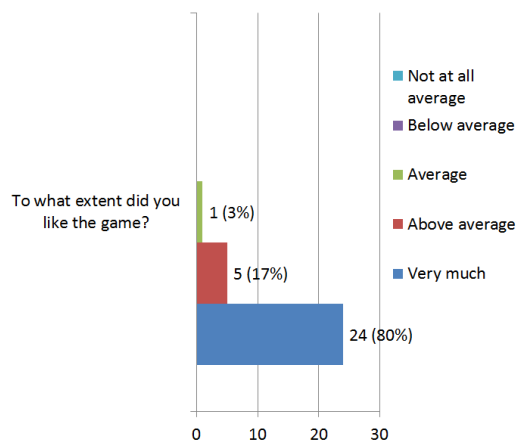
**Figure 6.** Players' attitude towards *Barbarossa*.

Impact of the invitational game phase.

Prior to investigating the impact of the invitational phase on the overall game experience, we evaluated the usability and learnability aspects of the invitational game mode to ensure that no serious technical flaws disrupted the players overall game experience (one of our main design objectives has been to enable participants to gradually develop game awareness, thus, it is important that the preliminary phase is trouble-free and easy to familiarize with). Figure 7presents the responses of evaluators with respect to their gaming background. Almost all players stated that they are regular video game players, while only two -first phase- participants have had previous experience with games similar to *Barbarossa*.
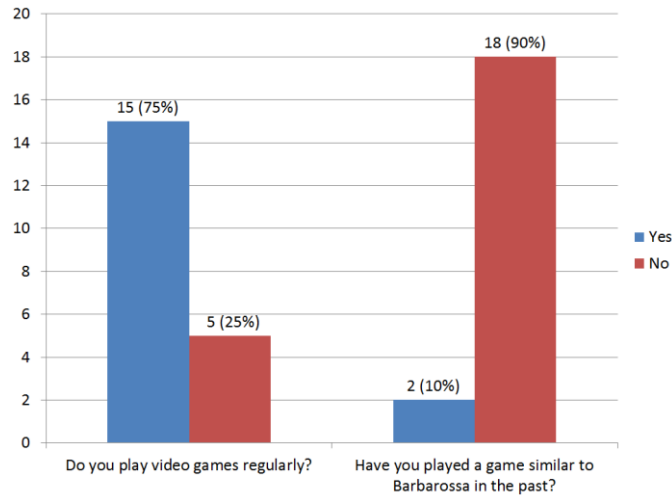
Figure 7. Player profile (demographic) questions.

We have also addressed several questions to participants to understand their perception of game usability and learnability aspects of the first phase game mode (see Figure 8).
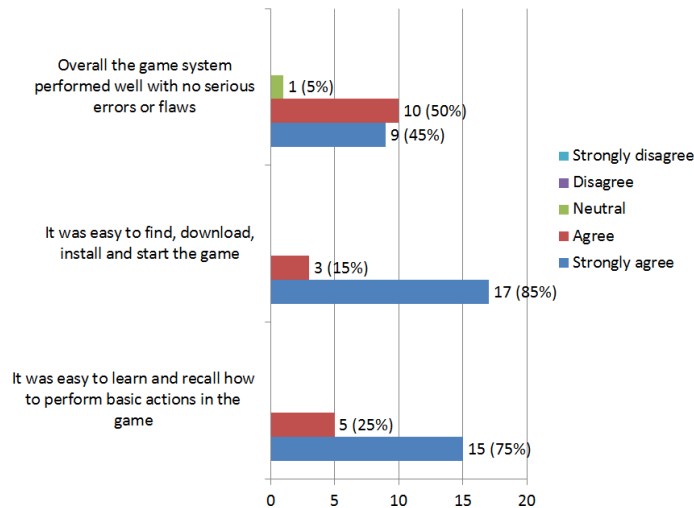

Figure 8. Easiness and usability-relevant questions.

Our evaluation results revealed that the invitational game mode performed well without any serious flaws affecting the players' experience. The above finding is backed by the only 6 error reports submitted by players to the Google Play Developer Console. Further, all players responded positively with respect to the game's learnability (i.e. their ability to recall how to perform basic game actions).

On the completion of the second game phase, we invited the players to complete questionnaires about the scenario they pursued on that phase, followed by an interview. We tried to assess the impact of the invitational mode to the players' view of the game goals and also check their commitment towards their team and the team's goals. Along this line, we asked the players to express their perception about the clarity of the game goals and also to comment on how responsible they felt to complete those goals in order for their team to succeed. As illustrated in Figure 9the players (even the *Knights* who did not met their co-players in person and played far from the game stage location) were aware of the game goals and also felt responsible to complete their mission and support the success of their team.
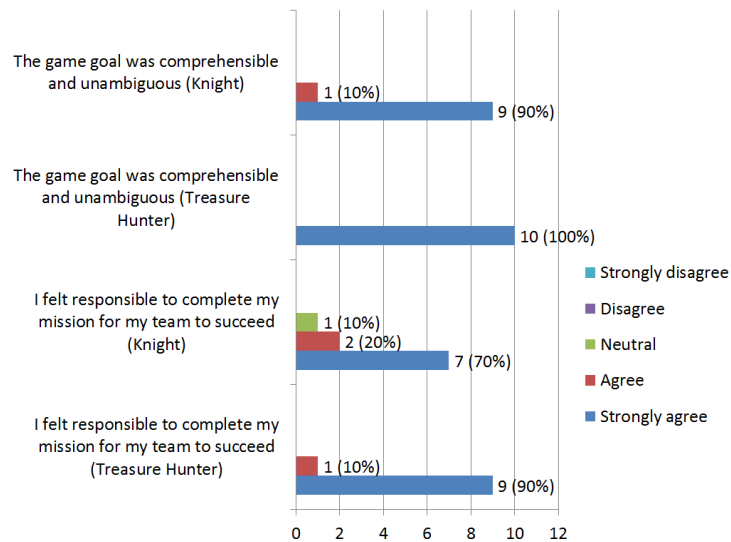
Figure 9. Questions about the overall game play experience & social/multiplayer aspects.

The most interesting findings with respect to the impact of the first game phase have been compiled from the interview, where in the participants have been encouraged to freely and fully express their views. Figure 10 illustrates players' opinion on the utility of the invitational phase. Participant answers are presented as a percentage of positive and negative responses received by interviewed *Treasure Hunters* and *Knights*. Note that the *Pirate* players have not been inquired about the first game phase as they had not participated in it.
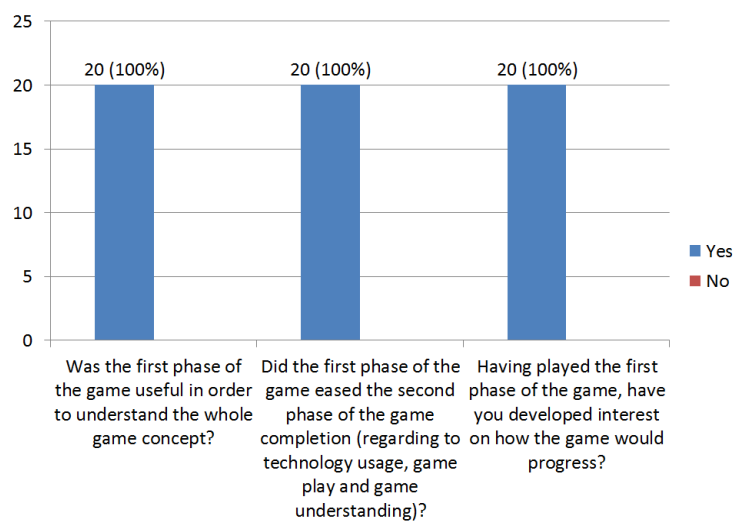
Figure 10. Interview questions about the invitation phase.

All *Knight* and *Treasure Hunter* players regarded the preliminary game phase as particularly useful for understanding the whole game concept. The first phase has also been found helpful for evaluators in completing the main game phase as it allowed them to familiarize with crucial game aspects (e.g. technology usage and game play) and increased their interest in following the game progress. In the sequel, we list some representative quotes extracted from the evaluators' interviews:

"The first phase helped me to grasp the whole game concept and game play."

"The first phase was useful for understanding the game spirit and getting introduced to its second phase."

"…the further I got into playing the game, the more I picked up the concept. I noticed that the knight had multiple capabilities to choose from, so I tested them all. The weapons, the use of color and the way that the outfit changed were amazing. I've been into role-playing games since I was 6 years old, so familiarizing with Barbarossa was not that difficult. Such kind of game can attract the attention of a certain crowd."

"It surely increased my interest to the whole game, I was eager to see how the game scenario would progress".

Impact of the recruitment of guests by co-players.

Recall that the *Treasure Hunters* have been invited to call a friend (guest player) into the game to undertake a rather critical role (*Pirate*) for the whole team to succeed. We firstly examined the familiarity of the guest players with the technologies utilized in *Barbarossa*. Thereafter, we evaluated the usability aspects of the *Set Me Free!* game mode to ensure that no serious errors or flows affected the game experience of the *Pirate* players. Last, we inquired these participants with respect to their perception of the game play and scenario.

Figure 11 shows that most of the guest players have been already familiar with the technologies utilized in *Barbarossa*; also, no serious errors and flows have been reported. None of the interviewees reported any difficulty in getting accustomed with basic game actions and interpreting the information displayed on screen while playing. These evaluation results have been confirmed by examining log data as all 10 players but one, who interrupted the game to request assistance in using a particular game feature, succeeded their game goal within a single game session.
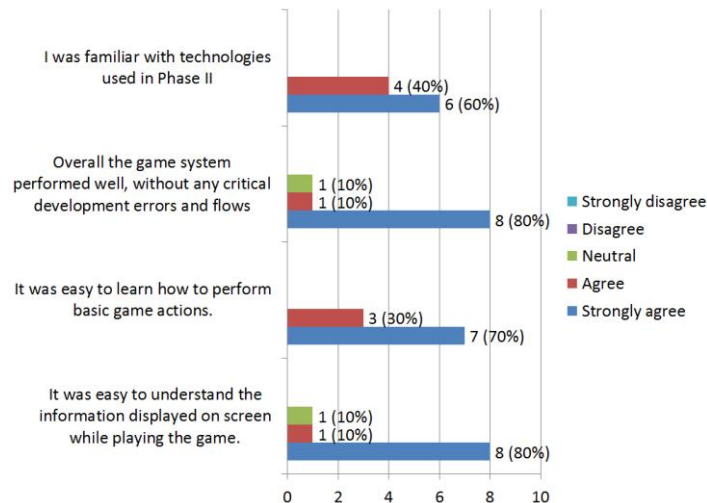

Figure 11. *Set Me Free!* usability & user profile (demographic)-relevant questions

We also prompted the guest players to express their perception on the game goals and comment on their game commitment, in order to assess how fast they digested the basic game principles and measure their willingness to put effort in contributing to their team's cause.

Indeed, as shown in Figure 12 all guests understood well the game goal and felt responsible towards their teammates to achieve it. Moreover, all guest players

argued that being invited by already committed players affected positively their attitude towards the game.
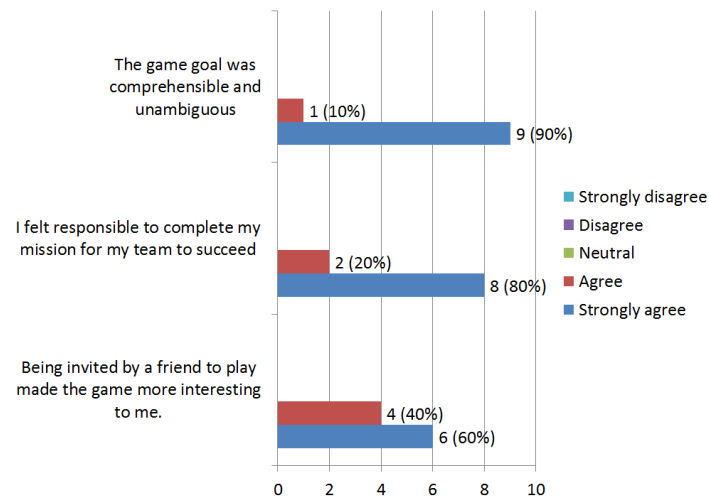
Figure 12. Questions about *Set Me Free!* game play experience & social/multiplayer aspects.

This recruitment method has also been perceived positively by the *Insiders* who argued that being able to invite a friend increased their enjoyment and sense of inclusiveness into the game.
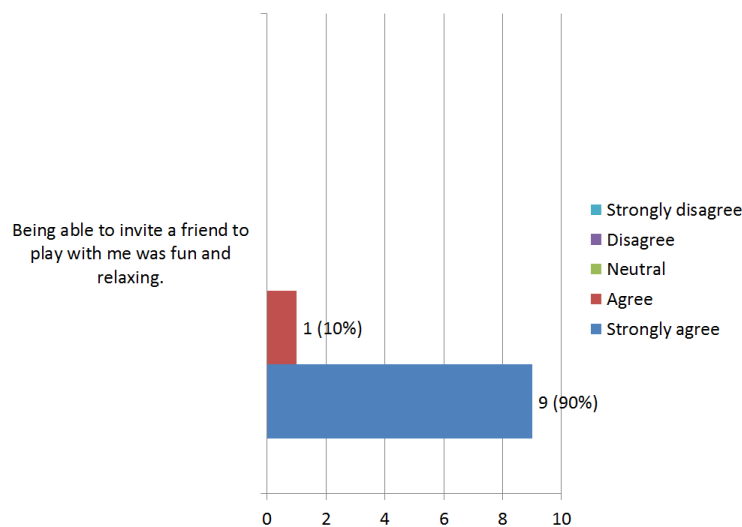

Figure 13. Questions about *Set Me Free!* social/multiplayer aspects.

5.3.   Summary of Evaluation Process and Discussion

The evaluation method proposed in this work involves a multi-phase process. The opening phase of comprises the design and development of the actual game phase. Thereafter, developers identify the key game aspects to be reproduced in the orchestration-free preliminary game phase (e.g. key technologies to train participants upon, scenario links between phases, etc) so as to increase the game awareness of players. In this step the developers should also decide on the ranking criteria upon which the most committed participants will be chosen (e.g. experience points or active game time). Having finalized and tested the preliminary game phase prototype, its official release is publicized via various communication channels so as

to reach both potential focus groups and the general public and ensure the participation of a large number of players. Participants should be allowed sufficient time to play. The preliminary game phase should separate the players that drop out of the game from the committed participants who belong into its potential target group; the latter is subsequently invited to act as evaluators in the main execution phase. Our evaluation method is illustrated in Figure 14.
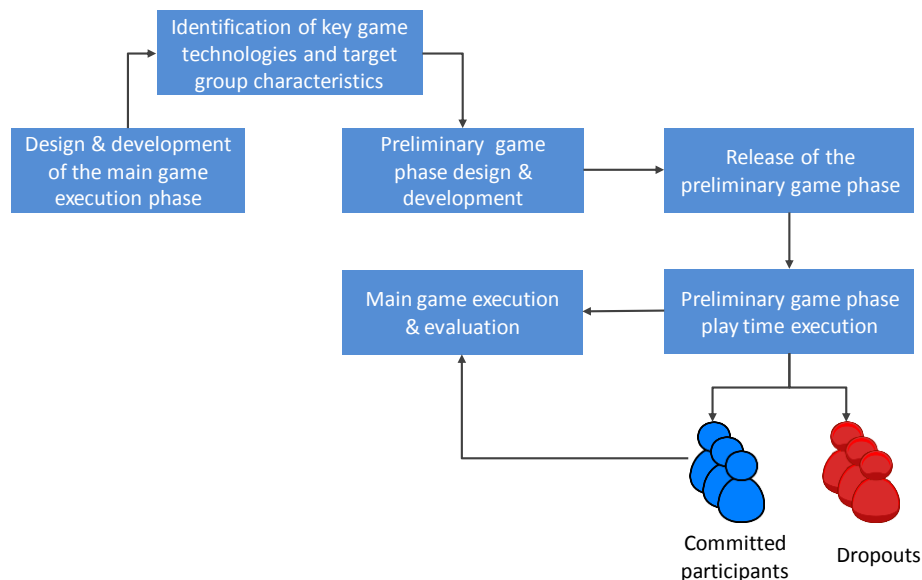
Figure 14. Evaluation process.

The evaluation results presented in the preceding subsections demonstrated that the invitational game mode (serving as a qualification round) has had significant impact in raising game awareness and training players in the technologies utilized during the main game phase of *Barbarossa*. Moreover, the invitational game mode triggered players' interest on the game's progression, thereby enhancing their keenness to participate in Phase II.

Confirming our first research hypothesis, the execution of the preparatory game designated, at an early stage, a group of participants with unarguable interest and positive attitude towards the game concept. This pool of players can be regarded as a representative sample of the potential target group of the game, which implies their suitability to act as evaluators of the main game product in the field trials, especially when requiring long-term commitment.

As regards the invitation of 'external' guest players by other (committed) players, the evaluation results of *Barbarossa* indicate that those have been largely familiar with the technologies used into *Barbarossa* and found it easy to grasp the game concept and goals. They have also demonstrated a high degree of commitment to completing their part of the game and contributing to their team's success. Besides, the social bond among the experienced and the guest player has shown to mutually increase their enjoyment, interest towards the game and sense of inclusivity.

Moreover, the combined usage of the above invitation methods has been shown to drastically reduce the possibility of player dropouts in the evaluation phase, facilitating the smooth execution of field trials and mitigating the risk of additional investments.

We argue that our proposed methodology for carrying out field trials should save considerable human resources (e.g. orchestration teams and evaluator trainers) and time, due to enabling a smoother execution of evaluation trials and reducing potential trial reiterations. We also assume that the attainment of valid, reliable and qualitative evaluation results should also have quantitative impact with regards to market success (i.e. financial revenue) of evaluated games. Nevertheless, the accurate quantification of resource savings is not feasible as it would require a direct comparison of our methodology against existing approaches upon a common pool of game prototypes.

6. Related Work

An outlook of existing pervasive game prototypes' evaluation reports reveals a variety of invitation methods which have been utilized for recruiting field trial participants. *EM II* (Epidemic Menace II) [16] is a cross media, augmented reality multiplayer game wherein players try to eliminate a humankind-threatening virus epidemic. The game has been evaluated by 29 evaluators (school students, journalists and FIT employees) invited using mailing lists. The evaluators of *Treasure* [18] have also been recruited via emails. *Treasure* is a game which exploits the "design-in-play" concept when played within players' daily living environments to enhance the variability of a game in mixed-reality environments.

In *Blowtooth* [23] players try to smuggle drugs through airport security using their Bluetooth enabled mobile phone devices. *Blowtooth* has been evaluated by 6 participants, recruited by personal contacts (e.g. friends of friends) of the developers. *URAAY* (Uncle Roy All Around You) [3] utilized live performance and pervasive technology to form a mixed reality game where players search of an elusive character named Uncle Roy. *URAAY* has been evaluated by a total number of 674 participants who answered to its advertising campaign. Finally, in *Insectopia* [19] players using a Bluetooth enabled mobile phone were challenged to acquire and maintain a collection of inspects, spawned at discoverable Bluetooth devices near the player. *Insectopia* has been publicly announced in the web, attracting 109 players in total. 11 players answered a questionnaire about their game experience and 7 of them also agreed to participate into an interview.

Jones and Marsden catalogued a list of advantages and disadvantages inherent in the above recruitment methods [21]. Notably, none of the above referenced evaluated pervasive games utilized invitation methods that enhance players' game awareness. That is, experiences are missing in assessing evaluation methods which actually introduce the players into the game. Even more so, no methods have been proposed to allow game designers monitoring and 'screening' the players and provide them the means to select the most eligible players to participate into the evaluation process.

7. Conclusion

Game evaluation trials represent a valuable tool for measuring the enjoyment and immersion perceived by players in the physical space, far from a supervised

laboratory environment [2, 3, 8-10, 16, 20, 23, 29]. However, they typically involve a lengthy and expensive process (preparatory activities, recruitment of evaluation subjects, trials orchestration and monitoring, execution and compilation of surveys, etc). As a result, user trials should be well prepared and carefully orchestrated to ensure their flawless execution. The composition of evaluator groups is critical to ensure that the conclusions of the evaluation process are qualitative and reliable. This paper proposes a methodological approach for performing user evaluation trials for pervasive games; our main focus has been the creation of a performative environment aiming at selecting highly qualified evaluation subjects in order to maximize the usefulness and investment payback of field trials.

Pervasive game prototype developers traditionally relied on emails, personal contacts and announcements/advertising to invite participants in order to perform evaluation user trials. In *Barbarossa*, we investigated a novel invitational method by developing a cost-effective preliminary game mode which framed a group of players with manifested interest into the game. The most committed players of that game mode have been invited into the main game phase, in effect, the 'official' game trials.

The evaluation results confirmed that the execution of a preparatory game mode, when applicable, can help developers to recruit highly qualified participants, truly enthusiastic to playing the game. Further, invitational game modes may serve as a useful instrument for developers to train evaluation participants on any technological equipment used in the game and also enhance their awareness on the overall game goal, scenario and gameplay. We have also found that allowing experienced, committed players to invite other participants may also be a useful means of participant recruitment as guest participants have shown to easily adapt and develop commitment and team spirit. Our evaluation results demonstrated that these two recruitment methods drastically reduced dropouts during field trials.

References

[1] M. Bell, M. Chalmers, L. Barkhuus, M. Hall, S. Sherwood, P. Tennent, B. Brown, D. Rowland, S. Benford, M. Capra, Interweaving mobile games with everyday life, Proceedings of the SIGCHI conference on Human Factors in computing systems, ACM, 2006, pp. 417-426.

[2] S. Benford, A. Crabtree, M. Flintham, A. Drozd, R. Anastasi, M. Paxton, Can You See Me Now?, ACM Transactions on Computer-Human Interaction, 2006, pp. 100-133.

[3] S. Benford, M. Flintham, A. Drozd, R. Anastasi, D. Rowland, N. Tandavanitj, M. Adams, J. Row-Farr, A. Oldroyd, J. Sutton, Uncle Roy All Around You: Implicating the City in a Location-Based Performance, ACM Advanced Computer Entertainment, 2004, pp. 1-10.

[4] C. Bezerra, R. Andrade, R.M. Santos, M. Abed, K.M. de Oliveira, J.M. Monteiro, I. Santos, H. Ezzedine, Challenges for usability testing in ubiquitous systems, Proceedings of the 26th Conference on l'Interaction Homme-Machine, ACM, 2014, pp. 183-188.

[5] H. Bial, S. Brady, The Performance Studies Reader, Routledge 2007.

[6] B. Brown, S. Reeves, S. Sherwood, Into the wild: challenges and opportunities for field trial methods, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2011, pp. 1657-1666.

[7] S. Bryan, The Gamer Tribes: An Ethnography, 2013.

[8] L. Chen, G. Chen, S. Benford, Your Way Your Missions: A Location-Aware Pervasive Game Exploiting the Routes of Players, International Journal of Human-Computer Interaction, 29 (2013) 110-128.

[9] A.D. Cheok, E.T. Khoo, Age Invaders: Inter-generational Mixed Reality Family Game, The International Journal of Virtual Reality, 5 (2006) 45-50.

[10] A.D. Cheok, A. Sreekumar, C. Lei, L.M. Thang, Capture the Flag: Mixed-Reality Social Gaming with Smart Phones, IEEE Pervasive Computing, 5 (2006) 62-63.

[11] N. Davies, Proof-of-concept demonstrators and other evils of application-led research: A position statement, Ubicomp, 2005.

[12] A.d.S. e Silva, L. Hjorth, Playful Urban Spaces A Historical Approach to Mobile Games, Simulation & Gaming, 40 (2009) 602-625.

[13] A.d.S. e Silva, D.M. Sutko, Digital cityscapes: Merging digital and urban playspaces, Peter Lang 2009.

[14] K.D. Eason, Towards the experimental study of usability, Behaviour & Information Technology, 3 (1984) 133-143.

[15] U. Ekman, Ubiquity, The Fibreculture Journal, (2011).

[16] J. Fischer, I. Lindt, J. Stenros, Final Crossmedia Report (part II) – Epidemic Menace II Evaluation Report, Integrated Project on Pervasive Gaming, 2006.

[17] J. Fritsch, T. Markussen, Exploring affective interactions, The Fibreculture Journal, (2012).

[18] B. Guo, R. Fujimura, D. Zhang, M. Imai, Design-in-play: improving the variability of indoor pervasive games, Multimedia Tools Appl., 59 (2012) 259-277.

[19] S. Hannamari, J. Kuittinen, M. Montola, Insectopia Evaluation Report, Integrated Project on Pervasive Gaming, 2007.

[20] I. Herbst, A.-K. Braun, R. McCall, W. Broll, TimeWarp: Interactive Time Travel with a Mobile Mixed Reality Game, Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services, ACM, 2008, pp. 240-242.

[21] M. Jones, G. Marsden, Mobile Interaction Design, Wiley 2006.

[22] V. Kasapakis, D. Gavalas, N. Bubaris, Pervasive games research: a design aspects-based state of the art report, Proceedings of the 17th Panhellenic Conference on Informatics, ACM, Thessaloniki, Greece, 2013, pp. 152-157.

[23] B. Kirman, C. Linehan, S. Lawson, Blowtooth: A Provocative Pervasive Game for Smuggling Virtual Drugs through Real Airport Security, Personal and Ubiquitous Computing, 16 (2012) 767-755.

[24] J. Kjeldskov, M. Skov, B. Als, R. Høegh, Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field, in: S. Brewster, M. Dunlop (Eds.) Mobile Human-Computer Interaction - MobileHCI 2004, Springer Berlin Heidelberg 2004, pp. 61-73.

[25] Y. Rogers, K. Connelly, L. Tedesco, W. Hazlewood, A. Kurtz, R. Hall, J. Hursey, T. Toscos, Why It's Worth the Hassle: The Value of In-Situ Studies When Designing Ubicomp, in: J. Krumm, G. Abowd, A. Seneviratne, T. Strang (Eds.) UbiComp 2007: Ubiquitous Computing, Springer Berlin Heidelberg 2007, pp. 336-353.

[26] H. Saarenpää, Data Gathering Methods for Evaluating Playability of Pervasive Mobile Games, (2008).

[27] J. Stenros, A. Waern, M. Montola, Studying the elusive experience in pervasive games, Simulation & Gaming, 13 (2011) 339-355.

[28] A.I. Wang, H. Guo, M. Zhu, A.S. Akselsen, K. Kristiansen, Survey on attitude towards pervasive games, Games Innovations Conference (ICE-GIC), 2010 International IEEE Consumer Electronics Society's, IEEE, 2010, pp. 1-8.

[29] R. Zender, R. Metzler, U. Lucke, FreshUP - A pervasive educational game for freshmen, Pervasive and Mobile Computing, 14 (2014) 47-56.