CountMeIn: Evaluating Social Presence in a Collaborative Pervasive Mobile Game Using NFC and Touchscreen Interaction

Michael Wolbert
University of Amsterdam
Amsterdam, NL
michael.wolbert@gmail.com

Abdallah El Ali
University of Amsterdam
Amsterdam, NL
abdallah.elali@gmail.com

Frank Nack
University of Amsterdam
Amsterdam, NL
f.m.nack@uva.nl

ABSTRACT
This paper presents the motivation, design and evaluation of CountMeIn, a mobile collaborative pervasive memory game to revive social interactions in public places (e.g. a train station or bus stop). Two versions of CountMeIn were tested; an NFC-based and a touchscreen version. In a 2x1 within-subject (NFC vs. Touch) experiment (N = 20), post-experiment group interviews and findings indicate the NFC version led to increased perception of social presence while participants were more aware of others' actions and intentions (mode of co-presence). However, we did not find quantitative evidence that attributes of social presence were higher from the Social Presence Game Questionnaire. Together, our findings suggest that placement of a physical NFC interface does not necessarily increase perceived social presence when users play collaboratively. However, social expansion in mobile collaborative pervasive games can greatly benefit from people's mutual awareness from such an interface. This mutual awareness has the potential to both attract users and spectators, and reduce anxiety of users to invite spectators, or accept an invite from users.

Author Keywords
NFC, touchscreen, collaborative, urban, pervasive, games

ACM Classification Keywords
H.5.2 [**User Interfaces**]: Input Devices and Strategies;
H.5.3 [**Group and Organization Interfaces**] Collaborative Computing

INTRODUCTION
We've all felt bored, frustrated, or even aggressive while waiting for a bus or train to arrive, or for a traffic light to turn green. While the widespread adoption of mobile devices has alleviated some of the burden of these long waits, they also isolate us more than ever. To revive social interactions amongst people in such settings, we propose the use of pervasive games. Pervasive games try to eliminate spatial, temporal and social constraints by expansion of the play space to one's daily life. As noted by Montola et al. [15], these expansions break Huizinga's "magic circle" and blur the existence between play and ordinary life, resulting in a playful space, which is constructed as a social product. This "magic circle" is used metaphorically to define boundaries (spatial, temporal and social) in which a game takes place. While traditional games are confined within these boundaries, pervasive games are not and blend in with our daily lives.

From the work of sociologist Erving Goffman [6], public places, such as train stations or bus stops, can best be described as temporarily short situated social gatherings in which a person is involved in social interactions within the social situation. In these public social gatherings, people who are formally introduced to one another (acquainted) and strangers (unacquainted) are part of the interaction. These situated social gatherings, therefore, present an ideal ground for exploration of pervasive gaming in the direction of social expansion.

Ubiquitous technologies, such as mobile devices, could serve the purpose of pervasive games given their detachment from the previously mentioned spatial-, temporal and social constraints. Moreover, game environments which leverage ubiquitous technologies can extend people's experience by providing new experiences through these technologies that act as layers on top of our existing perceived environments [5,21]. The goal of this study is to create new experiences through mobile devices; to transform waiting in public places into a more fun, engaging and worthwhile social gathering. We hypothesize that incorporation of Near Field Communication (NFC) can not only lead to interesting physical interactions between users and objects, but also between users. The latter mode, also known as peer-to-peer NFC (P2P NFC), promises an interesting opportunity to evaluate these in context of social expansion in mobile pervasive gaming. This is because these interactions depend on proximate and physical (face-to-face) encounters between users within the social situation.

The rest of this paper is organized as follows. First we review related work in the RELATED WORK section. From there we frame our research questions. Section COUNTMEIN we elaborate the design of the collaborative pervasive mobile game designed, implemented and evaluated during this study. In the METHODS section, we present the methods and procedures for this study. In RESULTS, we present the study results. Finally, in the DISCUSSION, CONCLUSION and FUTURE WORK sections, we present a discussion of our work, a conclusion and the direction of future work, respectively.

RELATED WORK

Social Presence

The widespread adoption of mobile technology has drastically changed our perception of social presence [1, 23]. Historically, presence, and in particular social presence, has been widely studied in the fields of sociology [6] and psychology. While there is still considerable disagreement among authors on what social presence precisely entails, a definition that seems to be generally agreed upon is: "*the sense of being together with another*" [1, p.10]. Both Boccia et al [1] and Zhao's taxonomy [23], identify co-presence as an essential dimension of social presence. Zhao further distinguishes between two main types: sense of co-presence and mode of co-presence. With sense of co-presence, he refers to social presence as defined in [1] and elaborates: "*a subjective experience of being together with others.*" Mode of co-presence is based on the work of Erving Goffman [6], who wrote: "*co-presence renders persons uniquely accessible, available and subject to one another*" [6, p. 22]. Social presence therefore encompasses how people experience their interactions with others and refers to conditions that should be met in order to experience a sense of co-presence (i.e. mutual awareness).

Pervasive Games

Pervasive games have been around long before mobile devices, and perhaps the best known is *Killer [15]*. While many versions exist, this game has been played in numerous cities around the world and provided numerous fun, socially awkward, and rich experiences. In *Killer*, not only participants are partaking in a playful activity, but also spectators and bystanders. In pervasive games, it is not uncommon for participants to switch between these roles and the roles of others are unknown to the game's participants.

For example, in the pervasive game *Blowtooth* by Linehan et al. [13] each user has the individual goal to smuggle as much virtual drugs from one side of the security checks to the other at international airports. *Blowtooth* relies on Bluetooth technology to virtually stash and collect the virtual drugs on and off unaware fellow passengers, who (un)intentionally broadcast their device's Bluetooth address. Social expansion in *Blowtooth* thrives on the idea that others (unaware passengers) are used as part of the game's narrative without actually knowing they are part of the playful activity. In their study, Linehan et al. [13] found that the controversial narrative of the game was experienced indifferently by its users in terms of awareness of security at the airport, anxiety and awareness of other passengers.

Physical Mobile Interaction

Physical mobile interaction entails use of mobile devices in physical interaction between users and objects, but may also be performed between two users. In a user study by Nandwani et al. [17], they evaluated three physical mobile interaction types; touching, pointing, and scanning, it was found that users perceived touching and pointing as most appropriate and natural. Broll et al. [2,3] present a strong focus on NFC interfaces and study usability and performance issues, and cross-modal use of mobile screen and physical NFC interfaces. In [2] they present the classic *Whack-a-Mole* game, which is playable on a dynamic NFC interface. In this case, the object of interaction is part of a larger dynamic interface (e.g. a large public screen), which can be updated upon interaction with a mobile device. According to Nandwani et al., much of the research effort on physical mobile interaction is directed at users interacting with everyday NFC tagged objects [2,3], rather than interactions between users. Nandwani et al. [17] present two parlor games which feature P2P NFC interaction. Findings of their study suggest that users felt these physical interactions as 1) "*[...] a greater sense of connecting to the other person*", 2) "*similar to giving a physical handshake*" and 3) "*less mechanical [and] more intimate [than other wireless solutions]*" [17, p. 24].

RESEARCH QUESTIONS

Given the foregoing motivation and literature review, we specifically ask 1) Does a physical NFC interface positively influence users' perceived social presence in a mobile collaborative pervasive game over the use of mobile touchscreens? 2) How do users perceive P2P NFC interactions in a public place with the (un)acquainted? We hypothesize that a physical NFC interface does positively influence users' perception regarding "a sense of being together", because we expect an increased perceived mode of co-presence (i.e., mutual awareness) as a result of collaborating with other co-located users on a physical mobile interface.

COUNTMEIN

To address our research questions, we have developed a collaborative pervasive mobile game prototype we call CountMeIn. Our approach to designing CountMeIn closely followed the literature on pervasive game design [15], prototyping procedures (e.g., play-testing) and game evaluation [11,12].

Requirements

From our research questions and literature review, CountMeIn should have the following high level requirements:

- Provide a socially expanded experience (inviting others, encouragement for participation and collaboration) to revive social interactions among people in public places.

- Provide a collaborative goal for 1-4 users with clear reward and punishment game mechanics.
- Is playable in a relatively short time frame (e.g. while waiting at a bus stop) and should be easy to learn and engage its users directly.
- Support NFC tag-based interaction in NFC version of CountMeIn, and support similar game interaction on touch-screen only. Both versions use P2P NFC interactions for physical mobile interactions between users.

Prototype

Gameplay

CountMeIn can best be described as a collaborative pervasive memory game. The goal of CountMeIn is for players to reproduce sequences through the use of game boards (Figure 1 and Figure 2). In CountMeIn, a sequence is a collection of numbers from 1 to 5. Points are awarded when users reproduce the given sequence. More points are awarded when sequences get longer. Additionally, a multiplier is awarded for the number of users who collaborated on the same sequence. When a sequence is not correctly reproduced, the score for that sequence is subtracted from the user's score. We did, however, add an inversed score system for the collaboration multiplier. This way losing a sequence collaboratively is more advantageous than losing individually and therefore incentive is created to play collaboratively. A similar approach is taken when a sequence is won.

Upon the participant's election to play a sequence, she is able to see the sequence on the mobile screen for a configured amount of time (currently set at 10 seconds). Thereafter, users reproduce the sequence through a game board (either an NFC poster or touch-screen). While sequences with a length of 3 to 5 numbers are relatively easy to reproduce by a single person, longer sequences may become a cognitive problem [14]. This limiting factor of our cognitive function can be exploited to support collaboration. In this case, participants can invite others into a running game by holding the back of their devices together and "beaming" the sequence to the invitee through P2P NFC (Figure 3). For example, when a user is playing a sequence with a length of 9, but was not able to remember the last 3 numbers, she is able to invite others through this interaction. An invited user is now able to see the sequence for the configured amount of time (timer resets again to 10 seconds). Subsequently, more users can be invited to collaborate. Based on [14] and [22], we distinguished three levels of sequence difficulties (easy; 4-5 numbers, medium; 6-7 numbers; hard; 8-9 numbers).

Implementation of CountMeIn

Both NFC and touchscreen versions of CountMeIn were implemented on Android devices. These versions differ on interaction strategies for the tasks of initializing the game and playing a sequence. In the NFC

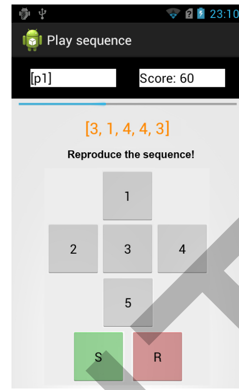

Figure 1. CountMeIn's Touch condition interface (S = Submit sequence, R = Reset sequence).

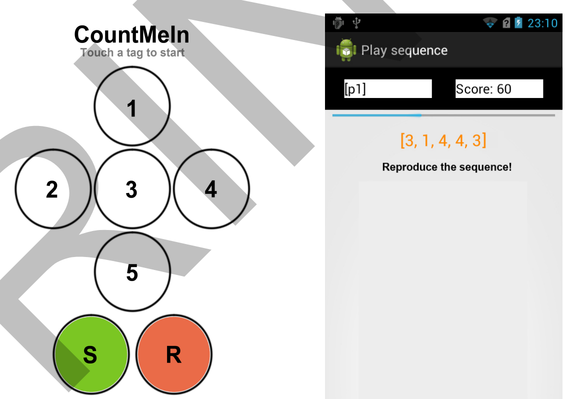

Figure 2. CountMeIn's NFC condition interface, left: NFC poster (S = Submit sequence, R = Reset sequence), right: mobile screen interface).

version, users can physically interact with an NFC poster to perform these tasks. The touchscreen version lets users initialize the game via the game launcher on the Android home screen and has a single screen to play a sequence. The invite mechanism for both versions is implemented through Android Beam facilities of the Android system. CountMeIn currently supports the following feedback: a) NFC-tag and P2P NFC interaction have stock Android feedback; tactile and sound b) Touchscreen button presses have no additional feedback other than stock visual cues c) In both versions, when a sequence is finished by a user (either won or lost) the correct sound is selected for playback d) Visual feedback (Figure 1 and Figure 2) for time left to see sequence (progress bar) scored points (dialog boxes and text), active players (text) and active sequence (text).

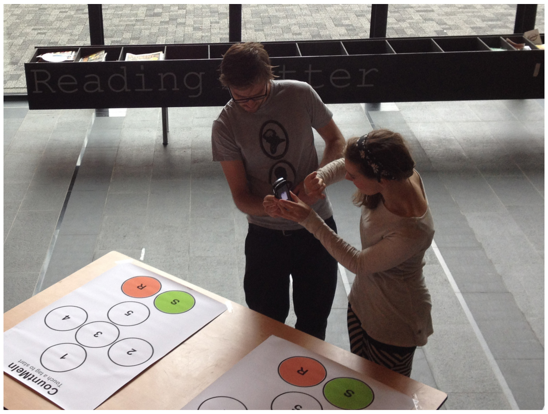

Figure 3. A participant (right) invites another participant (left) through P2P NFC in the NFC condition.

METHODS

Study Design

A counterbalanced 2x1 (NFC vs. Touch) within-subject design was used to investigate perceived social presence in a semi-controlled experiment set in a public place. The decision for a public place as experimental environment is two-fold. First, the relation of pervasive gaming to our everyday life could be compromised if tested in a closed environment (e.g., a lab). Ecological validity is therefore maintained. Second, we were also interested to what degree users perceive social presence towards spectators. This would be problematic in a closed environment without instructed spectators. The public place in which the experiment took place is the public area of the University of X's main building. In this area, mostly undergraduate students and university personnel roam the area while waiting for classes or when walking to other parts of the building. This area therefore presents similar features to that of a train station or bus stop.

To address our first research question, we used the Social Presence Game Questionnaire (SPGQ) [12], a module of the Game Experience Questionnaire [10,11][1]. Items on the GEQ are answered on a 5-point Likert-scale (range [0, 4], with "Not at all", "Slightly", "Moderately", "Fairly" and "Extremely" respectively) per item. The constructs measured in the SPGQ module are *empathy* (EMP; α = .93) towards other players, *behavioral involvement* (BIP; α = .91) towards other players and *negative feelings* (NF; α = .64) towards other players. Internal consistency of each construct proved to be highly to reasonably reliable. To include measurement of behavioral involvement towards perceived social presence with spectators, we reused altered BIP items that measure user's *behavioral involvement* towards spectators. The result was a new construct, BIS (shown in Table 1) with 6 items and a reasonable internal consistency (α = .64).

Behavioral Involvement Spectators (BIS)
My actions depended on spectators actions
The spectators' actions were dependent on my actions
The spectators paid close attention to me
I paid close attention to the spectators
What the spectators did affected what I did
What I did affected what the spectators did

Table 1. Questionnaire items for the added BIS construct.

To gain more insight, participants received a post-experiment semi-structured group interview. During these group interviews we asked how participants experienced playing both versions of CountMeIn in a public place (i.e. university hall or while waiting at a bus/train stop) and whether they preferred one version over the other given this setting. Additionally, for our second research question, we asked participants how they evaluated the invite interaction (implemented in both NFC and touchscreen versions) and how participants considered engaging in these types of interactions with others (with the focus on the unacquainted) in a public setting.

We included the core and post-game modules of the GEQ to get participant's feedback on overall perceived game experience. The core module consists of the constructs *competence* (CMP; α = .91), *challenge* (CHL; α = .72), *positive affect* (PA; α = .88), *negative affect* (NA; α = .65), *annoyance* (ANN; α = .85), *flow* (FLW; α = .83), and *sensory and imaginative immersion* (SNS; α = .76), were included. The post-game module consists of the following constructs; *positive experience* (PXP; α = .90), *negative experience* (NXP; α = .76), *tiredness* (TIR; α = .47), and *returning to reality* (RTR; α = .52). For the sake of brevity we did not include items of GEQ core and post-game constructs in this paper. Instead we refer to the literature [10,11] for a discussion of these items

Finally, since participants could become more physically involved, which could arguably lead to increased workloads, we included the NASA-TLX questionnaire [9] after each condition. The NASA-TLX questionnaire is a multidimensional assessment tool that rates perceived workload based on six subscales, namely mental demand, physical demand, temporal demand, performance, effort, and frustration.

For the demographic analysis, we asked participants about their personal mobile device type, self-reported computer

1 I Items for all used GEQ (including SPGQ) constructs can be found in the referenced literature.

skills ("*Novice*", "*Intermediate*", and "*Advanced*"), familiarity with NFC interactions, and whether they had a technical background.

Participants

20 participants (10 males, 10 females), aged between 21-51 ($M = 26.1$; $SD = 6.3$) were recruited through the researchers' networks. Our sample contained 5 different nationalities, of which Dutch was the most prominent (15/20). More than half (12/20) of the participants indicated they were familiar with NFC technology and knowledgeable of its capabilities. Participants indicated their personal mobile device was either Android (12/20), iPhone (7/20) or Blackberry (1/20). In our sample, slightly more than half of the participants (11/20) indicated they had a technical background. Additionally, participants' self-reported computer skills were novice (2/20), intermediate (7/20) and advanced (11/20).

Figure 4. Experimental setup with 3 participants playing the NFC version of CountMeIn.

Figure 5. Experimental setup with 4 participants playing the Touch version of CountMeIn.

Setup & Procedure

The experiment (**Figure 4** and **Figure 5**) was conducted in 6 test sessions where we set up 2 tables with 4 physical game boards and an additional table for our hardware. The physical game boards were printed on A1 paper and NFC-tags were attached to the design. To make the touchscreen version comparable to the NFC version, active sequences are shown and updated on each player's screen in real-time. To further balance versions, the number of game boards is equal for both versions.

All recruited participants were invited to schedule a timeslot with a limit of 4 participants per timeslot. While the number of scheduled participants per session was visible to potential participants, the names of these participants were anonymized in order to circumvent allocation bias of new participants choosing a specific session. At the start of each test session a general introduction to the study was given. Participants received an Android device and a snack while filling out forms for informed consent and additional background information. Participants were also asked to indicate to whom they were acquainted prior to our experiment. **Table 2** presents an overview of the allocation of participants and the number of participant pairs who were acquainted and unacquainted. For example, in session 1 (with 3 participants), P1 and P2 were acquainted (1 acquainted pair). P1 and P3, and P2 and P3 were unacquainted (2 unacquainted pairs).

Group	Session	N	Acquainted pairs	Unacquainted pairs
NFC	1	3	1	2
	2	3	1	2
	3	4	1	5
Touch	4	3	0	3
	5	4	6	0
	6	3	1	2

Table 2. Participant session allocation.

Participants in the NFC group initially received a tutorial of CountMeIn with NFC, while group Touch received the touchscreen version first. Thereafter, all participants were able to familiarize themselves with CountMeIn's rules and interactions for ~3 minutes in their assigned first condition and were encouraged to try several actions (e.g. resetting, submitting, inviting). Before the second condition started, we gave a shorter demonstration and emphasized that all previous rules of CountMeIn were similar to the first condition. Participants were told they were able to roam around the area (whilst staying connected to Wi-Fi) and that they were free to use any of the game boards available to them. No time limit was set for a game session, however the number of sequences played was always the same.

After each condition, participants received NASA-TLX and GEQ (including SPGQ) questionnaires. Both responses were filled out on the same questionnaire allowing for relative responses across conditions. When test sessions were completed, all participants in the session were interviewed (which we video recorded) as a group regarding both versions of CountMeIn, and how they evaluated interactions in public while playing the game. Afterwards, participants were thanked and were offered the promised monetary reward.

We configured CountMeIn to include 24 sequences (8 easy, 8 medium and 8 hard) ready to be played by our participants in each test session. All test sessions used the same sequences except for the demonstration session. On average participants played ~8.85 sequences, of which ~6.92 were won and ~1.92 were lost. Of the previous figure, participants played ~4.28 games collaboratively, of which ~3.75 were won and ~0.52 were lost.

Data Processing & Analysis

The responses for each GEQ construct were coded by taking the mean score of its items. We found several missing values for items of the SPGQ constructs: 1 item for EMP, 1 item for BIP and 2 items for NF. To resolve these missing values we removed these items from our initial recoding procedure and performed checks for internal consistency of each construct. Nevertheless, after this procedure good internal consistency was still met for EMP ($\alpha = .93$), BIP ($\alpha = .88$) and moderate consistency for NF ($\alpha = .63$).

We ran Shapiro-Wilk's tests for all measured constructs by condition to assess normality. For SPGQ's constructs, EMP, BIP and NF were normally distributed. BIS was corrected with a log transformation, but did not yield a normal distribution. Bartlett's test for homogeneity of variance was assumed for EMP, BIP and NF by condition. GEQ and NASA-TLX constructs were assessed as well. Depending on this assessment of normality the appropriate statistical test was used. Since our experiment design was 2x2, no posthoc tests were applied for the main statistical results.

Since the number of spectators varied across test sessions, it was not surprising that participants felt they were not able to fill out some of the BIS questionnaire items. While this resulted in missing values, we recoded these as "Not at all" (see Study Limitations for discussion). For all conducted tests, a confidence level of 95% was used.

RESULTS

SPGQ Responses

A one-way repeated measures ANOVA (N = 20) was conducted on each dependent variable EMP, BIP, and NF by condition (see **Figure 6**). Although slightly higher means were found across these constructs for NFC over Touch, no statistically significant difference was found for: EMP by condition NFC (M = 2.39; SD = 1.00) and Touch (M = 2.28; SD = 1.15); $F(1,19) = .19$, p = .67, BIP by condition NFC (M = 1.84; SD = 0.99) and Touch (M = 1.73; SD = 1.08); $F(1,19) = .41$, p = .53, and NF by condition NFC (M = .85; SD = .86) and Touch (M = .62; SD = .71); $F(1,19) = 2.38$, p = 0.15. A one-way Friedman's test was conducted for the BIS construct by condition NFC (Md = .92, IQR = [.29, 1.21]) and Touch (Md = .50, IQR = [0, .87]), however, no statistical significant difference was found; χ^2 (1) = 2.25, p = .13. These results led to provisionally accepting the null hypothesis; no difference was found for perception of attributes of social presence between the NFC and Touch versions of CountMeIn.

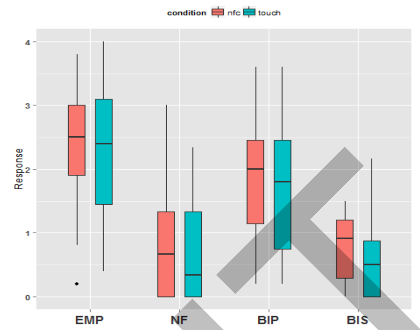

Figure 6. SPGQ response boxplots, including outliers 'dots'.

GEQ Core & Post-game Responses

For each construct in the GEQ core (**Figure 7**) and post-game modules (**Figure 8**), we conducted one-way repeated measure ANOVAs or Friedman's tests (N = 20), depending on whether the normality assumption was violated. We did, however, only find a statistical significant difference between NFC (Md = 2.30, IQR =[1.40, 2.85]) and Touch (Md =1.70, IQR = [1.20, 1.21]) for *sensory and imaginative immersion (SNS)*; χ^2 *(1) = 4.765, p = .03.*

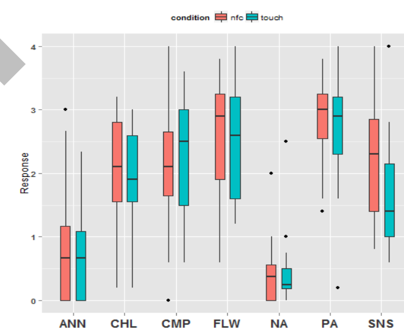

Figure 7. GEQ core module responses, including outliers 'dots'.

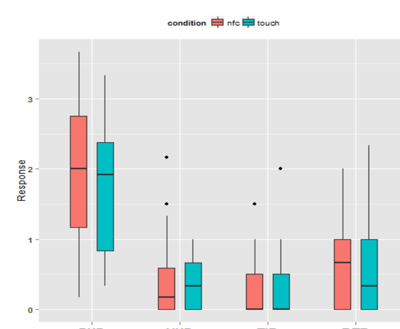

Figure 8. GEQ post-game responses, including outliers 'dots'

NASA-TLX Workload Responses

A Friedman's test was conducted for dependent variable *subjective load* (SL) by condition NFC (*Md* = 9.67; *IQR* = [4, 15.3]) and Touch (*Md* = 8.92; *IQR* = [1.7, 12]). A borderline statistical significance was found; $\chi^2(1)$ 3.20, p = .07. We proceeded analysis on the raw scores of NASA-TLX constructs (**Figure 9**) as discussed in [8]. A one-way repeated measures ANOVA revealed that indeed *physical demand* (PHY) was perceived higher for NFC (*M* = 6.95; *SD* = 4.29) over Touch (*M* = 4.00; *SD* = 2.81); $F(1,19)$ = 13.84, p < .01. Similarly, *mental demand* (MNT) was perceived higher in condition NFC (*M* = 12.45; *SD* = 3.99) over Touch (*M* = 11.25; *SD* = 3.70); $F(1,19)$ = 4.669, p = .04. Furthermore, a one-way Friedman's test revealed that construct *effort* (EFF) was perceived higher in condition NFC (*Md* = 12; *IQR* = [2,16]) over Touch (*Md* = 10; *IQR* = [1,13]); $\chi^2(1) = 13.23$, p < .001.

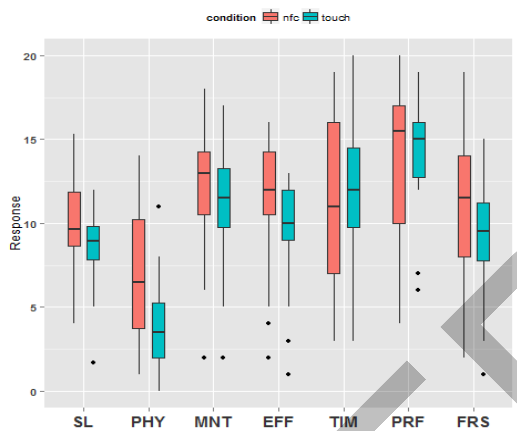

Figure 9. NASA-TLX responses, including outliers 'dots' (SL = Subjective Load, PHY = Physical Demand, MNT = Mental Demand, EFF = Effort, TIM = Time, PRF = Performance, FRS = Frustration)

Explanatory Analysis

For each SPGQ construct, *empathy* (EMP), *behavioral involvement towards players* (BIP), and *negative feelings* (NF), we ran multiple repeated measure ANOVAs each with one added between-subject factor from participants' demographic data. These tests did not yield significant main effects for the added between-subject factors, nor did we find significant interaction effects between condition and the added between-subject factors. Therefore, no additional posthoc testing was conducted.

Participant Group Interviews

CountMeIn in Public

When asked about playing CountMeIn in a public place with the unacquainted (i.e. a bus stop or train station), participants (11/20) indicated they would play such a game. Few participants (5/20) stated they would play to alleviate boredom while waiting (P18: "*When I'm waiting at a bus stop I usually play around with my phone, checking emails or playing a game. I think this game is interesting in this context"*). Few participants (3/20) stated they would do so in order to meet new people (P6: "*I would play such a game to kill waiting time. It is always fun to play a game and meet new people!*"). However, some participants (5/20) stated that they would rather not play with others in such settings because they enjoyed playing on their own (P3: "*I feel more comfortable keeping the game to myself, because otherwise you are dependent on others*"), or were simply not interested in playing in public with others at all (P19: "*When I'm on my own in a public place I don't feel the urge to play with other people in real-time*"). Another participant also argued he did not feel the urge to play in public and added that he might play at a more private place, such as a party or with friends at home. When asked about which version participants preferred in a public place, most participants (16/20) stated they preferred NFC. Few participants (3/16) stated that novelty of NFC was a reason for them to prefer this version (P1: "*I liked the poster version much more, it was new and different*").

Collaborative Pervasive Gaming

Interestingly, while all participants stated they enjoyed collaborating with one another in both versions, most participants (16/20) argued the physical NFC interface seemed more appropriate for collaborative pervasive gaming (P12: "*I find the NFC poster more interesting to be played together. Playing the touchscreen version is more fun on your own*"). It was noted by many participants (14/20) that the NFC version allowed users to be more involved with others (P2: "*I find it much more fun to actually see the others playing, cursing and being frustrated. It feels so much more alive*"). Moreover, participants argued that awareness of others' actions was perceived higher in the NFC version. (P15: "*If you are playing with your [touchscreen] only in a public place, then it is not entirely obvious to others that you are playing a game*").

Additionally, many participants (14/20) also indicated their perception of behavioral involvement towards spectators was higher when playing the NFC version (P14: "*With the poster everyone is able to see what others are doing. This makes it easier to go to others.*"). These participants stated this increased perception of involvement could attract aware spectators into the game (P18: "*Having the poster is nice since you go out of your screen, you just have more options to attract more people to join you*"). On the other hand, one participant stated this increased mutual awareness might backfire and lead to increased anxiety (P10: "*In the NFC version I feel much more like being watched and judged on my performance*"). Furthermore, a few participants (4/20) stated they preferred competitive - over collaborative pervasive play in public (P8: "*I prefer to play a competitive game. If you can only gain points together I expect the fun to be over soon.*").

Inviting the Unacquainted

When asked about participants' attitudes towards using the invite interaction of CountMeIn in a public place to invite the unacquainted, many participants (14/20) stated that they would do so in order to achieve the game's goal. However, when asked about an unacquainted other who invites, most participants (16/20) stated that performing the interaction depended on the social situation of which the presence of the unacquainted inviter was noted to be decisive (P14: "*If a shady person walks up to me I would say that my device does not support it.*"). Almost half of the participants (8/20) stated they were worried about their security and privacy when performing this type of interaction in a public place (P5: "*I would be suspicious. If you do not know this game then somebody might think: 'I will get all your data in the disguise of playing a game.*") (cf. [17]). These participants were mainly concerned the other would use the interaction to breach their device's security, get sensitive private data (e.g. contacts, emails, photos), upload a malicious application, or simply steal it (P14: "*I might find it troublesome if a stranger would hold up his device to perform the interaction. It may well be a scam or he even might steal it.*").

Many participants (16/20) indicated that being knowledgeable about the game would increase the chance of performing the invite interaction with the unacquainted. Participants argued that because the type of interaction was relatively new to them, intentions of the inviting party might not be clear. (P17: "*If I did not know the person and what [the interaction] was about I would find it weird, so I should either know the person or the game.*"). Interestingly, especially for the touchscreen version, a few participants (5/20) stated they needed functionality in the game itself to get whereabouts of possible other players of CountMeIn in order to approach and invite them (P1: "*I would like to have some kind of radar to see if other people around me have the game too.*").

When probe questions resulted in asking our participants about the nature of the invite interaction between two unacquainted persons, participants disagreed in their views. More than half of the participants (13/20) evaluated the interaction as close (P8: *"I feel like I connected to a person instead of a device."*), while others (7/20) did not perceive the interactions as being close (P11: *"I'm creating a connection from device to device, I do not touch the other person."*). Interestingly, from the latter view, participants disagreed to what extent this interaction was experienced as actually touching the other physically (i.e. making contact with hands or fingers during the interaction). Additionally, few participants (3/20) stated they perceived a mobile device to be extensions of themselves, and in this way experienced the interaction as physically touching others (P5: "*Nowadays devices are extensions of ourselves and via this interaction you literally make contact*"). Some participants (5/20) attributed analogies to the invite interaction in a public place with the unacquainted. For example, participants (3/5) stated it felt like asking for directions, while other analogies referred to physical involvement (P5: *"Shaking hands, but then holding hands a bit longer"*, P20: *"Highfiving!"*) (cf. [17]).

DISCUSSION

Triangulation of Methods

We found strong qualitative evidence from our group interviews suggesting that perceived social presence towards other users and spectators increased while playing CountMeIn on a physical NFC interface. However, we did not find evidence in our SPGQ responses supporting the qualitative findings.

These conflicting results force us to conjecture what could have been the cause of this discrepancy in triangulation of the methods used. One possible explanation is that although we counterbalanced our experimental design, carry-over effects could be present in repeated measures of SPGQ constructs. Perhaps, an asymmetrical carry-over effect [19] is present in our experimental procedure. For example, if participants in group "Touch" initially collaborated successfully, then this likely influenced their ability to collaborate even better in the following condition. On the other hand, group NFC may have collaborated badly initially, and even worse in the following condition as a result of their shared performance in the first condition. These carry-over effects could have biased our SPGQ findings. While carry-over effects (e.g. fatigue and learning) were expected to be counterbalanced, the dependence of repeated measurement of SPGQ constructs between conditions could have interfered with our inference. We suggest that a more representative – and larger sample is needed to obtain more conclusive findings on these constructs.

Also, similar slope directions of the two experimental groups for EMP, BIP and BIS) over time suggest that the chosen design might exhibit symmetrical carry-over effects. In this case, a possible explanation could be that during the introduction of our experiment, participants' perception of social presence towards each other was already firmly established and did slightly decrease as the experimental session unfolded.

Sensory & Imaginative Immersion

A statistically significant difference for construct *sensory and imaginative immersion* (SNS) was found, which was also confirmed during the group interviews. According to Nacke & Lindley [16] defining *immersion* is problematic because no clear definition seems to exist. They argue that *immersion* can be described as one's level of emergence into gameplay, and would therefore be related to the goal of the game. Therefore, it could be suggested that the collaborative goal of CountMeIn contributes to the difference in perceived SNS. Moreover, SNS would be a measure towards behavioral involvement to accomplish the goal of the game. Subsequently, this could explain the increased perception of behavioral involvement in the interview findings. However, in [4] it is argued that the human sensory system is part of *immersion*. Therefore, one explanation could be that the construct suffers from internal validity. A perhaps more plausible explanation could be that *immersion* was perceived differently because of the way the physical interface stimulated users' sensory system (i.e. not only audio-visually, but also increased coordination of body movements). This is supported by the increased *physical demand* and *effort* (cf. Workload Responses).

Mutual Awareness & Anxiety

From our interview findings we found that holding up a mobile device to perform a P2P NFC interaction, while blurting (although enthusiastically from game enjoyment) instructions to unacquainted others in public, is considered inappropriate (cf. Participant Group Interviews). Not surprisingly, during interviews we found that familiarity regarding this interaction by both parties was considered important. Additionally, the intention for the interaction was also suggested to be important, but was perceived as problematic because one would not know if the other is aware of the given game. In this case, when only using a mobile touchscreen, it is suggested that users perceived it to be difficult to assess whether others were involved in the same game.

Montola et al. [15] elaborate on three levels of awareness states regarding pervasive games, namely conscious state, ambiguous aware- and unaware states. The latter two states are of interest, because the soon-to-be-invited participant either knows something is going on (ambiguous state) or does not know what is going on at all. Montola et al. [15] argue these awareness states are a sliding scale from which our interview findings suggest that the use of a physical NFC interface could lead to a more rapid transition from unaware – to ambiguous state and consequently to conscious state.

The increased awareness and visibility of other users interacting with a physical interface over having only a mobile touchscreen therefore suggests an increased mode of co-presence (cf. Participant Group Interviews). In the public interfaces literature, this state transition is perhaps best be described as the honey-pot effect. This effect suggests that people are drawn to public interfaces while they can spectate (from a safe distance) and be enticed by others to join the activity [7]. Interestingly, this effect also suggests decreased anxiety for both parties, because the intention surrounding the game invitation becomes clear to the invited party (cf. Participant Group Interviews).

Study Limitations

When we balanced the comparison of NFC and touchscreen versions of CountMeIn, the design resulted in an equal number of physical- and touchscreen game boards. For a comparison on perceived social presence between versions, this is suggested to be appropriate for our study, because having a single physical interface would inevitably provoke increased perception of social presence. In the case of ecological validity, however, this design decision has limited the NFC condition, because a real-world setting would preferably have only one physical interface (i.e. a single NFC poster at a bus stop).

Because we decided to study this phenomenon in a real-world setting, a consequence was that spectators (e.g. number of spectators or their different behaviors) varied across test sessions. While we recoded missing values for BIS items to "Not at all", it could, however, be that there were no spectators around for our participants to report on. Participants in this study were not able to invite spectators, since NFC-enabled devices were limited, and the prototype of CountMeIn was not publicly available for download by non-participants (spectators) yet.

CONCLUSION

In this study we were interested in users' perception of social presence when participating in a mobile collaborative pervasive game when using either a physical mobile NFC – or mobile touchscreen interface. In our experiment, we did find qualitative evidence that suggests users perceived an increased sense of mutual awareness, behavioral involvement towards others and reduced anxiety when a physical NFC interface was used. However, we did not find conclusive quantitative evidence from our Social Presence Game Questionnaire supporting our hypothesis that perceived attributes of social presence increased when a physical NFC interface was used. Additionally, our qualitative findings suggest that the "invite" feature of CountMeIn, which has been implemented through peer-to-peer NFC, is promising for mobile pervasive gaming in the direction of social expansion. All in all, our findings suggest that placement of a physical mobile interface, over the use of a mobile touchscreen interface only, increases mode of co-presence. Consequently, this increase suggests a positive influence on perceived social presence. Social expansion of collaborative pervasive games can therefore greatly benefit from such interfaces in the direction of fun and engaging revival of our everyday social interactions in public.

FUTURE WORK

Future work should be in the direction of deploying CountMeIn in a real-world setting for further experimentation and evaluation in the wild. From a game design perspective, incorporation of real-life gamification (e.g. winning coupons for free train tickets or a free soft-drink at a train station) would be interesting. Such extrinsic incentive would stimulate people to play the game, and as a result create a larger user-base.

We also expect that designing peer-to-peer NFC interactions for mobile pervasive games, and mobile games in general, leads to interesting gameplay opportunities. For example, a mobile pervasive role-playing game could be extended with a peer-to-peer NFC interaction to trade goods, or combine in-game characters' powers while on a quest in an urban center. The design of such an interaction is obviously dependent on mutual participation of the users involved. Mobile HCI research should look into this dependency to further explore the limitation of NFC but also its possibilities from a technological and social perspective.

REFERENCES

1. Biocca, F., Harms, C., and Burgoon, J.K. Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence Teleoperators Virtual Environments 12*, 5 (2003), 456–480.
2. Broll, G., Graebsch, R., Scherr, M., Boring, S., Holleis, P., and Wagner, M. Touch to Play -- Exploring Touch-Based Mobile Interaction with Public Displays. *2011 Third International Workshop on Near Field Communication*, (2011), 15–20.
3. Broll, G. and Hausen, D. Mobile and physical user interfaces for NFC-based mobile interaction with multiple tags. *Human Factors*, (2010), 133–142.
4. Brown, E. and Cairns, P. A grounded investigation of game immersion. *Extended abstracts of the 2004 conference on Human factors and computing systems - CHI '04*, (2004), 1297.
5. Dourish, P. and Bell, G. The infrastructure of experience and the experience of infrastructure: meaning and structure in everyday encounters with space. *Environment and Planning B: Planning and Design 34*, 3 (2007), 414–430.
6. Goffman, E. *Behavior in public places: Notes on the social organization of gatherings*. Free Pr, 1966.
7. Hara, K.O., Glancy, M., and Robertshaw, S. Understanding Collective Play in an Urban Screen Game. *Proceedings of the 2008 ACM conference on Computer supported cooperative work 3*, (2008), 67–76.
8. Hart, S. NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (2006) 50*, 9 (2006), 904–908.
9. Hart, S.G. and Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload 1*, 11 (1988), 139–183.
10. IJsselsteijn, Kort, D., and Poels. The Game Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital games. Manuscript in Preparation. (2013).
11. IJsselsteijn, W. and Hoogen, W. Van Den. Measuring the experience of digital game enjoyment. *Proceedings of Measuring Behavior*, (2008), 26–29.
12. Kort, Y. de, IJsselsteijn, W., and Poels, K. Digital games as social presence technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). *Design*, (2007), 1–9.
13. Linehan, C. and Kirman, B. Blowtooth: pervasive gaming in unique and challenging environments. *Human Factors*, (2010), 2695–2704.
14. Miller, G. a. The magical number seven, plus or minus two: some limits on our capacity for processing information. 1956. *Psychological review 101*, 2 (1994), 343–52.
15. Montola, M., Stenros, J., and Waern, A. *Pervasive Games: Theory and Design*. Morgan Kaufmann, 2009.
16. Nacke, L. and Lindley, C. Affective ludology, flow and immersion in a first-person shooter: Measurement of player experience. *Measurement 3*, 5 (2010).
17. Nandwani, A., Coulton, P., and Edwards, R. NFC Mobile Parlor Games Enabling Direct Player to Player Interaction. *2011 Third International Workshop on Near Field Communication*, (2011), 21–25.
18. Paulos, E. and Goodman, E. The familiar stranger: anxiety, comfort, and play in public places. *Proceedings of the SIGCHI conference on Human factors in computing systems*, (2004), 223–230.
19. Poulton, E.C. unwanted assymetrical transfer effects with balanced experimental designs. *Psychological bulletin 59*, 1 (1962), 257–72.
20. Rukzio, E. and Leichtenstern, K. An experimental comparison of physical mobile interaction techniques: Touching, pointing and scanning. *UbiComp 2006: Ubiquitous Computing Lecture Notes in Computer Science 4206*, (2006), 87–104.
21. Williams, A. and Dourish, P. Imagining the City of Urban Computing. *Computer 39*, 9 (2006), 38–43.
22. Wolbert, M. and Abdallah El. Ali. Evaluating NFC and Touchscreen Interactions in Collaborative Mobile Pervasive Games, *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services*, (2013), 522–527.
23. Zhao, S. Toward a taxonomy of copresence. *Presence: Teleoperators and Virtual Environments*, (2003), 445–456.