Eliciting situated feedback: A comparison of paper, web forms and public displays

Jorge Goncalves *, Simo Hosio, Yong Liu, Vassilis Kostakos

University of Oulu, Department of Computer Science and Engineering, Pentti kaiteran katu 1, P.O. Box 4500, FI-90014 Oulu, Finland

ARTICLE INFO

ABSTRACT

Interactive displays are becoming an increasingly popular civic engagement mechanism for collecting user feedback in urban settings. However, to date no study has investigated (i) how the situatedness of public displays affects the quantity and quality of collected feedback, and (ii) how public displays compare to traditional paper or web feedback mechanisms. We answer these research questions in a series of lab and field studies. We demonstrate that because people tend to approach this technology with no specific purpose in mind, the feedback collected with public displays is noisier than web and paper forms. However, we also show that public displays attract much more feedback than web and paper forms, and generate much more interest. Furthermore, we found that users appropriated our technology beyond its original purpose. Our analysis provides implications on the tradeoffs of using public displays as a feedback mechanism, and we discuss ways of improving the collected feedback using public displays.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Public interactive displays offer a unique opportunity to promote civic engagement in urban settings. The visibility of these displays is an advantage that can be leveraged by city officials when collecting feedback and in an attempt to enhance civic engagement [16]. In this sense, the attractiveness and inherent situatedness of this medium have been suggested as catalysts in promoting civic engagement within a community [19,28]. Researchers have so far been mostly concerned with the usability and design aspects of providing feedback on such displays [e.g., 2,29,36], for example considering how to use mobile devices to minimize on-screen typing on public displays [e.g., 28,35]. However, they have mostly overlooked the issue of content: *what kind of feedback do such displays elicit, and of what quality*? In previous work, researchers have either claimed success in their deployment [e.g. 7] although their experiments are conducted in an artificially controlled environment, or have simply ignored noisy or irrelevant feedback without exploring the issue [e.g. 18]. We argue that a systematic investigation is crucial to our understanding of the reasons for, and potential solutions to, this question.

Furthermore, it is not clear *what are the added benefits of public displays in relation to more traditional mechanisms* for gathering feedback and eliciting engagement, such as paper forms or web forms. Is the increasing popularity of public displays justified, or do they introduce some bias in the way they engage users?

Answering these questions is challenging due to the "in the wild" nature of these public interactive displays, and because the public is a diverse audience that tends to give mostly "noisy" feedback [18].

In this paper we present the first, to our knowledge, series of lab and field studies that investigate how public interactive displays are used for the purpose of providing civic feedback. Our investigation of public feedback mechanisms is far from novel, however no studies have considered this in the context of public interactive displays. In this paper we seek to measure the *quality* and *quantity* of feedback collected on public displays, and investigate whether the situatedness of this technology has an effect by considering two other mechanisms of collecting feedback from the public: paper and web forms.

2. Background

2.1. Public feedback promoting civic engagement

Eliciting feedback from the public is a worthwhile cause. It has been shown that citizen participation and feedback leads to positive outcomes both for people and institutions and also for the society of which they are part of [8]. Not only are decisions made through broader civic participation better, they are also more likely to be accepted [16]. Others have argued that citizen participation can increase one's social wellbeing by reinforcing their perception as being socially integrated and accepted and by strengthening their belief in being beneficial to themselves and society [27].

* Corresponding author. Tel.: +358 8 404140304; fax: +358 8 5532 612.
E-mail address: jgoncalv@ee.oulu.fi (J. Goncalves).

Feedback is also an *essential* part of civic engagement. In fact, a great number of studies have explored civic engagement through reporting from mobile phones [e.g., 12,13]. However, definitions on what civic engagement actually means or encapsulates can vary. However, we note two definitions found in literature [9] claiming that: (i) engagement usually refers to participation in activities that benefit themselves and, often, their communities or society as a whole; and (ii) it refers to participation and a sense of belonging in community, school, the workforce, and other aspects of life.

It may be argued that engagement is both a process and an outcome. As a process, engagement and feedback in a community are a positive development leading to long-term social inclusion [9]. Civic engagement can then be viewed from the perspective of benefits to be gained and costs to be borne. Participation in activities, communities, and institutions, enhances positive identity development, social responsibility, and the development of a wide range of life skills and competencies. Participation is a powerful tool at the disposal of citizens that wish to engage in democracy pursuing positive social change to the benefit of the individuals involved and to society as a whole [9]. In today's society citizens are normally encouraged to adopt more active roles [6], however civic engagement can only be fostered on the basis of reciprocal trust between people and responsible institutions [1].

2.2. Feedback mechanisms of public interactive displays

A number of previous studies have investigated the use of public interactive displays for the purpose of gathering feedback and fostering civic engagement [3,7,18]. Recently, Hosio et al. [18] reported on *Ubinion*, a service that employed large public interactive displays to enable young people to give personalized feedback on municipal issues to the local youth workers. While successful in engaging the youth, the study reported a high level of appropriation of the public displays. The intended purpose of the system – to give feedback on youth related matters – took a back seat to user self-expression and playful interactions.

In another study, Brignull and Rogers [7] report on *Opinionizer* designed to encourage socializing and interaction. Participants could add their views and give opinions onto a shared display, which they and others could observe and then add further add comments to if they felt inclined to do so. During the study the authors found that a major deterrent preventing people from participating is social embarrassment. They suggest that a way to avoid this is by making this public interaction have a purpose behind it. They also claim that the environment, both on and around the display, produces strong physical and social affordances, which people can easily and unambiguously pick up on. Hence they argue for facilitating the public in its needs to rapidly develop their conceptions of the purpose of the social activity, and to be able to move seamlessly and comfortably between being an onlooker and a participant.

Finally, another study that considered public displays as feedback mechanisms was *TextTales* [3]. Here the authors attempted to explore the connection between story authorship and civic discourse. They did this by installing a large, city-scale, interactive public installation that displays a 3-by-3 grid of image–text combinations. A discussion on a certain topic would then be born from SMS sent by users that then created captions for pictures present on the system. This study reflects many other studies focusing on moving text entry away from the display [e.g., 28,38]. However, due to some people's lack of willingness or capability to use such proposed systems [19] one ends up with a substantially smaller potential sample of contributors. We argue that by allowing the text entry mechanism to remain on the display we do not exclude anyone from the process.

In the above examples, the authors report on the importance of deploying these feedback mechanisms in authentic environments, or living laboratories [34]. Public interactive displays provide a unique medium that fosters opportunistic feedback gathering by passersby to understand situated and contextually relevant information, leading to better insight in feedback [5]. Coupled with the honey-pot effect [7] displays can foster the appearance of "mini-communities" around them, participating in a dialog about the topic at hand as reported in several previous studies [e.g., 30,33].

An important issue that most previous studies have overlooked, however, is establishing how this new medium compares to existing and well-established channels for feedback. While studies have claimed that people do indeed use these displays for providing feedback, no study has directly compared displays to other feedback mechanisms. Advances in display technology have reduced costs, but still the cost of maintaining a display and software in public can be high [31]. Therefore, we seek to establish the extent to which public displays can be reliable in terms of quality of feedback but also cost-effective mechanisms.

3. Studies

In a series of studies we seek to establish how public displays affect the quantity and quality of collected feedback, and how this compares to paper and web feedback mechanisms. We conducted our studies within the context of two ongoing debates in our city: the major reconstruction of a central pedestrian street, and the quality of education of our university. We note that the public displays used in our studies have been part of the city's infrastructure for about 4 years and therefore we did not expect any effect of novelty in our findings.

(a) Study 1 was conducted in the field, and its purpose was to replicate and validate previous studies [3,7,18] that used public displays to collect situated feedback. This study was conducted in the context of the pedestrian street reconstruction, and we sought to measure the quality of the collected feedback.

(b) Study 2 was a controlled lab study conducted in the context of the quality of education in our university. Here we assessed participants' attitudes towards providing feedback using a public display, a web form, and a paper form.

(c) Study 3 used identical apparatus as the second study, but was conducted in the wild over a 2-month period without explicitly recruiting participants. Here we sought to validate our findings from Study 2 in a more naturalist setting.

3.1. Study 1

During initial meetings, city officials they expressed their ongoing struggle to collect feedback from citizens regarding a major pedestrian street reconstruction. The reconstruction had received frequent attention in the online version of the local newspaper but in a non-constructive and negative manner. Although the city officials had set up a web-based application for this purpose, and allowed citizens to give feedback using paper forms, these mechanisms were not getting used at all. Our goal, therefore, was to funnel the public's interest in a more constructive manner, and redirect it to city officials using a medium they could control. Our first step was to better understand why exactly citizens did not provide feedback to the city officials directly. We conducted semi-structured interviews ($N = 20$) around the city in public settings by interviewing citizens (7 male, 13 female) aged 16–60 ($M = 29.9$, $SD = 11.3$). The questions were: (1) Demographics, (2)

have they ever provided any feedback concerning the renovation to the city officials and if yes by what means, if not for what reason, (3) what [other] means of providing feedback do they know, (4) could they provide an alternative way for citizens to give feedback on such matters. None of the participants reported having submitted any feedback to city officials on the renovation or otherwise. Regarding the existence of means to provide such feedback, participants either assumed they exist ($N = 13$, e.g., website and city officials' office) or simply reported not knowing about any ($N = 7$). Amongst those that reported the city officials' office as a means to provide feedback, the majority ($N = 9$) claimed not being willing to physically visit the location just for the purpose of providing feedback. Interestingly, the majority of interviewees ($N = 15$) suggested the use of the citywide deployment of public interactive displays for the purpose of giving feedback on the pedestrian street reconstruction. They claimed the displays' location near the actual reconstruction site would remind and motivate them to give their thoughts.

Following-up on the previous question, we asked what input method for giving feedback on the display they would prefer. The majority of interviewees ($N = 18$) said that a mechanism to input text directly on the public display would be the ideal way for them to give feedback on the displays. Suggestions of audio or video recording were immediately dismissed due to privacy concerns. Interviewees also claimed they would not want to spend money for sending SMS or spend time with a pairing mechanism, and the majority also claimed not to have a data plan. These sentiments reflect the latest survey conducted in 2011 by the Official Statistics in our country [blind], in which 33% of respondents said they use their mobile phones for Internet access and only 22% claimed to use it on a weekly basis. Given this, we chose not to experiment with different types of entry mechanisms. Furthermore, while previous research has repeatedly reported that text entry on interactive displays is very challenging [e.g., 28,38], the fact that our potential users cannot be trained before using any novel text entry mechanism can reduce people's willingness or capability to use the system [20].

While our intention from the beginning was to suggest the use of the citywide interactive public display deployment, as we believed the situatedness and attractiveness of this medium suited our purposes, the answers given by the citizens increased our confidence and that of the city officials in the idea. This deployment consisted of 16 public interactive multi-purpose displays installed in a number of indoor and outdoor public locations, as shown in Fig. 1.

Fig. 1. One of the public interactive displays used to collect feedback in Study 1.

3.1.1. Design

We deployed an application that collected free-form text input from citizens, and also gave them the option of answering two Likert-scale questions. We attempted to replicate previous research that has reported on such feedback techniques [18].

In terms of how citizens would input text, using *physical* keyboards is impractical on such interactive public displays, as there is no good place to "park" them or it would require each person to bring their own. Taking this into consideration and the comments obtained during the interviews we decided on the use of a *soft* keyboard [25,26], which is a virtual keyboard that is displayed on the screen instead of having a physical form. This allowed us to avoid the problems related to placement of the keyboard. The soft keyboard was designed following the recommendations of [23] to minimize error rates. Initially we considered allowing users to be more creative while providing feedback (i.e., scribbling [10]) using the public display (as well as the web form). However, we decided against this mainly due to two reasons: (1) Encouraging scribbling or other forms of creativity would bias people towards appropriation of the public displays, and (2) We wanted to keep the traditional feedback mechanisms realistic, so by allowing scribbling on the public displays the same would have to be allowed on the web forms going against this very principle.

In addition we used two Likert scale questions as we felt they are an easy and fast way to gather feedback, and have worked well according to previous work [4]. The statements that we used in the Likert questions were originally in Finnish, and their translations are:

- "*The large renovation project in <location> is topical and necessary for <City>!*" and
- "*<City> officials are informing citizens sufficiently about the renovation project!*".

To assess the validity of the Likert scale results, we alternated the above statements with their opposite phrasing:

- "*The large renovation project in <location> is irrelevant and not necessary for <City>!*" and
- "*<City> officials are not informing citizens sufficiently about the renovation project!*".

We expected that the alternation between the two sets of statements would give us a clearer understanding on how reliable these scores were, and we expected a roughly symmetrical average score. Furthermore, previous research has argued for the use of item-specific instead of agree/disagree as response options [37]. However, we intended for this method of giving feedback to be engaging and quick. Therefore we chose the use of "smileys" as the scale of agree/disagree response options. To verify that the statements were not misleading we conducted a small pre-study with 10 native speakers who agreed that the difference between the positive and negative statements was clear.

We deployed our system "in the wild" on multiple public interactive displays around the city for 1 month. They were mostly located relatively near to the actual reconstruction site in downtown. The Likert scale data was averaged across all entries. The open-ended text responses obtained were manually labeled as either relevant or non-relevant to the city's reconstruction. A meeting with the city officials then followed the deployment in which we assessed the deployment's success.

3.1.2. Results

A total of 17 textual comments were inserted during this time in addition to 46 answers to the Likert-scale questions. We compared timestamps of the answers and all 46 responses were given

at distinct timeframes (as opposed to in rapid succession) so we therefore assumed that we had 46 unique respondents to the survey. The results obtained from the 5-point Likert scales statements were:

- 4.44 (SD = 1.23, Mdn = 5) and 4.32 (SD = 1.11, Mdn = 5) for positive statements (N = 25).
- 3.86 (SD = 1.71, Mdn = 5) and 4.14 (SD = 1.42, Mdn = 5) for negative statements (N = 21).

3.1.3. Findings

While a perfectly symmetrical average score between the positive and negative statements could not be realistically expected, the very close proximity of the scores suggests that many respondents simply did not read the statements, or it was hard to read in the context of the displays. This was further corroborated by the fact that the median was 5 for both positive and negative statements. However, we acknowledge that these results may have simply been due to problems inherent to the positive versus negative balancing method used such as acquiescence, social desirability and extremity bias [14]. Regardless of the reasons behind the results, we can still infer that the use of Likert scales as feedback mechanism, at least in this context, was not reliable. Furthermore, both ourselves and the city officials felt it was limited in richness of feedback and was thus abandoned in subsequent studies.

Additionally, contrary to interviewees' claims of wanting a mechanism to directly input feedback on the public displays, the results showed that the comments were relatively "noisy". Although the sheer number of citizen feedback entries (17) was characterized as "promising" by the city officials, it was mainly noise (e.g., "*zzzzzz*", "*nananana*") with only 2 comments being relevant to the issue at hand (e.g., "*More workers to get this reconstruction done faster please*").

These results raised the question whether an interactive public display is suited for this type of feedback collection. However, several previous studies have reported using this medium to successfully gather feedback from the community [e.g., 3,18]. Hence, to better understand our own findings we followed with Study 2 in a controlled lab setting.

3.2. Study 2

In this study we contrasted situated feedback on an interactive public display with two more traditional mechanisms for giving feedback: paper and web forms. We did this to better understand if the ratio of noise was due to the *medium* or the *situatedness* of feedback. The role of the medium was highlighted in comparing the public display versus paper forms, while the role of situatedness was assessed through the comparison between the public display and the web form.

3.2.1. Design

The study followed a within-subjects experimental design, with each participant having to write pre-determined feedback messages on a public display, a paper form, and a web form. We prepared a set of three feedback statements relating to different type of feedback, namely: Suggestion, Quality of Service, and Infrastructure.

- Statement 1 (Suggestion): "There should be more choices of languages to study."
- Statement 2 (Quality of Service): "The level of teaching is not up to European stands!"
- Statement 3 (Infrastructure): "The sports facilities for students are insufficient and old."

We made these statements purposely negative as to gauge the level of comfort each participant would have writing the different types of feedback on each medium. We instructed all participants to write each feedback statement using all three mediums: an interactive public display with a soft keyboard and textbox, a paper form, and a web form using a desktop computer. All three mediums had the same basic interface layout. All trials took place in the same public setting (Fig. 2 top) to address any location-based bias with the desktop component being conducted in a collocated private booth (Fig. 2 bottom) as to mimic the privacy conditions at a participant's home.

3.2.2. Measures

In both the desktop and the public display application we logged the following: *participant id, medium, typing time* (from the first character typed to the last) and *number of backspaces/deletes*. For the paper version we only gathered *participant id* and *writing time* by using observation and a stopwatch. Upon completing all tasks they answered a questionnaire with demographics and a single 7-point Likert scale in which they *stated how comfortable they were* while giving each statement on the different mediums. Finally, we conducted semi-structured interviews with all participants to get further insights. The questions were: (1) Demographics, (2) which medium would they prefer, if any, for writing each type of feedback, (3) which medium would they prefer using in general, (4) their opinion regarding advantages and disadvantages
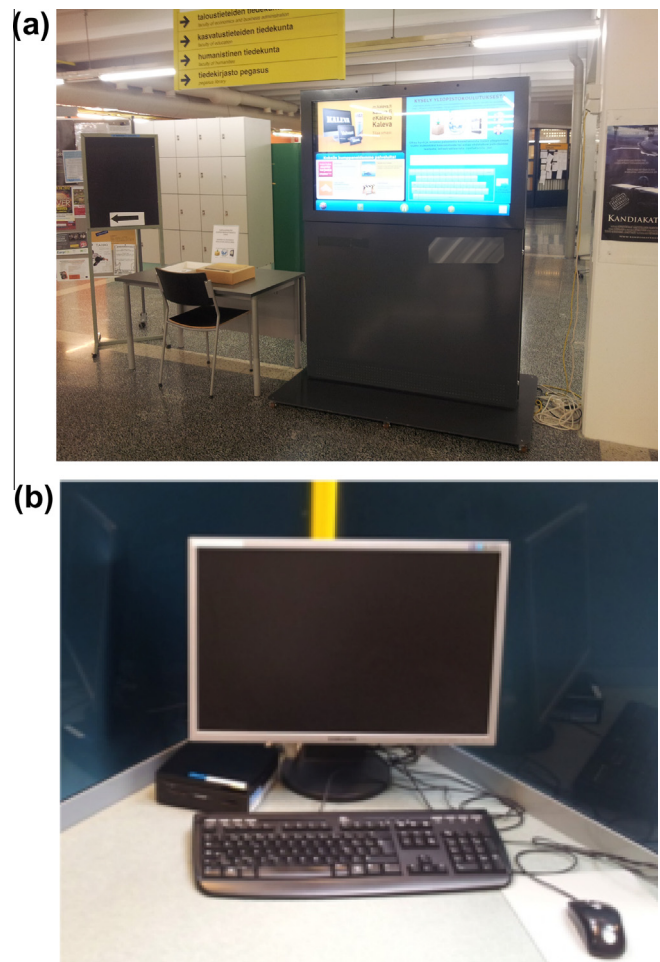
Fig. 2. Left: The public display used in Study 2, with the paper forms on top of the table to the left of the display. Right: the booth PC used to complete the web form in Study 2.

of each medium, and (5) their opinion regarding having other type of input mechanisms (e.g., video, audio, and using phone).

3.2.3. Participants

In total 18 participants (6 female, 12 male) aged 20–37 (M = 27.3, SD = 4.7) were recruited. They were all recruited amongst passersby in front of an interactive public display located near the main cafeteria in our university. All participants were briefed on the details of the study and given a movie ticket as compensation.

3.2.4. Apparatus

A 57-inch touch screen was used as the public display, mounted vertically at 1.2 m from the floor (Fig. 2 left). This is the same hardware used during the deployment in Study 1. A 15-inch desktop computer with a three-key external mouse was used to access the web-based application, placed on a separate private booth to the public display (Fig. 2 bottom). A table with the A4 size paper forms was placed next to public display with a vertical plastic sign notifying its presence (Fig. 2 right).

3.2.5. Procedure

Each participant was briefed on the process and the approximate time the whole experiment would last. We handed them three separate pieces of paper with each one having a different feedback comment we wanted them to input on each medium. The order of medium used by the participants was counter-balanced, as was the order of feedback statements.

3.2.6. Results

Each participant entered the three feedback comments on each medium. This resulted in each participant completing 9 tasks, for a total of 162 tasks in the study. Figs. 3–5 summarize the self-reported level of comfort when typing each of the 3 pre-defined message on each of the 3 mediums.

Detailed results are broken down by medium in Table 1. Saphiro–Wilks tests of normality showed that for each statement the answers from at least one of the mediums were not normally distributed ($p < .05$) leading us to use non-parametric tests in further analysis. There was a statistically significant difference in level of comfort depending on which medium was used while typing the Suggestion statement ($\chi^2(2) = 19.40$, $p < .01$, Friedman test), the Quality of Service statement ($\chi^2(2) = 23.81$, $p < .01$, Friendman test) and the Infrastructure statement ($\chi^2(2) = 25.24$, $p < .01$, Friedman test). We chose not to run similar analysis for time spent as the measurement for the tasks done on paper was not precise but recorded manually using a stopwatch. Overall, our results show that participants felt less comfortable when giving feedback on the public display.

The majority of participants stated that they would only use the public display to give feedback on an event that just happened. As one participant claimed

"*It all depends heavily on the mood I am, maybe I could give even angry feedback in public if I am mad and disappointed in something that happened just now*"

or to avoid forgetting about it:

"*If I just finished a horrible lecture, I'd definitely use a public display or a ballot box to give feedback. Maybe when I get home, I probably would not care about it anymore. Would like to just spit it out right away.*"

Another concern by participants was the sheer length of the comments they would want to make in which participants would only use the public display short and spontaneous feedback:

"*The type of content matters a bit, not that much. The length is a bigger issue – for short and spontaneous feedback the public displays is ok! For longer and more thoughtful whining, using online app from home is the best.*"

while avoiding it for more in-depth and emotional comments:

"*I would like to have my own time and edit my text carefully before submitting anything – maybe at home is the best then at least for more in-depth and emotional content. So it depends on the content also, for something lightweight a public display or ballot box is ok.*"

Interestingly, participants did not think it would be a good idea to use video or audio to give feedback on the public display mainly because of privacy concerns:

"*I would prefer text on both mainly because of privacy issues so that nobody knows it was me regardless of what type of feedback I'm giving*"

but also due to awkwardness:

"*I would find it very awkward to talk to a machine in video or audio feedback, so text is still the best although it really does not work on a public display.*"

3.2.7. Findings

This study was not meant to be an exhaustive experimental investigation, but rather was used to shed light on the findings of
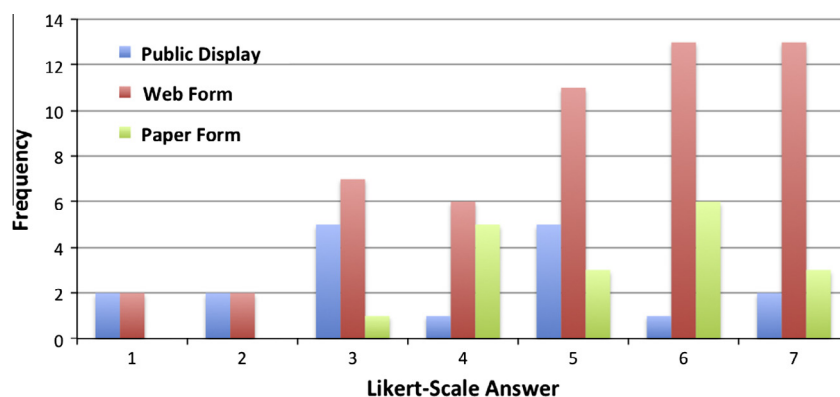

Fig. 3. Distribution of answers for pre-defined statement 1 (Suggestion) for each medium.
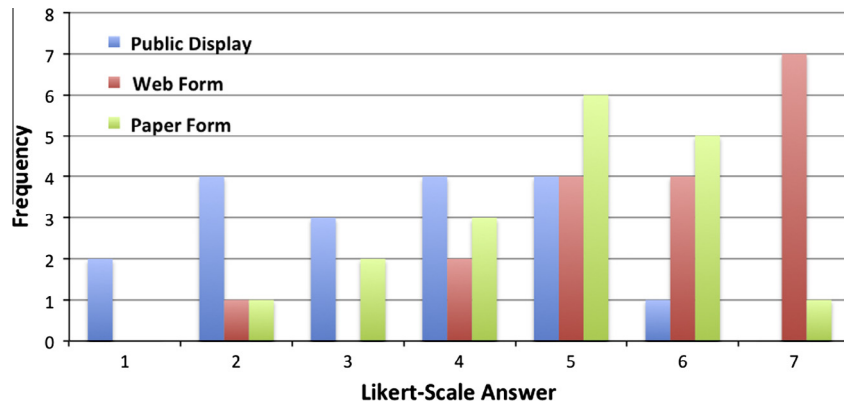
Fig. 4. Distribution of answers for pre-defined statement 2 (Quality of Service) for each medium.
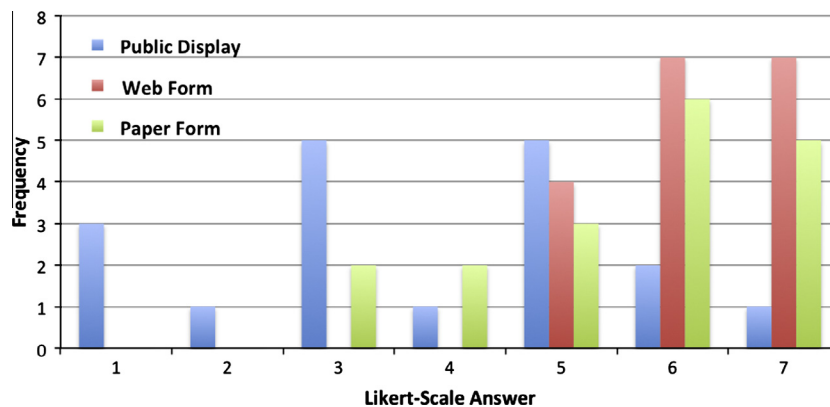

Fig. 5. Distribution of answers for pre-defined statement 3 (Infrastructure) for each medium.

Table 1
Results obtained from Study 2 for all three mediums.

	Public display	Web form	Paper form
Total comments	54	54	54
Avg. typing time (in seconds)	53.1	34.3	~42.1
Avg. number of backspaces/ deletes	6.9	8.8	N/A
Avg. comfort (statement 1)	3.9	6.1	5.3
Avg. comfort (statement 2)	3.4	5.7	4.8
Avg. comfort (statement 3)	3.8	6.2	5.6

Study 1, particularly on the issue of noisy and irrelevant comments. Study 2 highlighted that giving text feedback on a public display:

- requires more effort,
- is less comfortable,
- can be more convenient at times.

Despite the difficulty of inputting text directly on the public display, participants were resistant to the idea of using video or audio due to privacy or awkwardness. The study highlighted multiple reasons and situations in which users claimed they would use the public display for giving feedback. All these took advantage of the *situated nature* of the display. However, it was important to investigate what would happen in a more naturalistic setting. Would the participants' claim translate into action when the interaction stopped being mediated by a researcher? We addressed this in the next study.

3.3. Study 3

3.3.1. Design

We conducted a 1-month "in the wild" deployment to measure the quantity and quality of feedback each of the 3 mediums (display, web, and paper) would elicit. The topic remained identical (i.e., education) and we used the same public display as in Study 2 on the University campus. For this study the feedback application was active on-screen at all times. We took this decision in order to make it possible to have both paper forms and public display available at the exact same location simultaneously. For the web form we did not setup a dedicated computer as we wanted people to give their feedback on their own device in order to mimic other web-based feedback forms. The web form was also optimized to allow those with Internet connectivity on their mobile phone to submit feedback through the given URL. We decided against having a dedicated mechanism for mobile phones (e.g., SMS gateway or app) due to feedback gathered during our interviews. Additionally, as with Study 2, the interface for all three mediums was as consistent as possible (Fig. 6).

Another concern was to make sure that the three mediums were promoted equally. We addressed this in two ways:

- we distributed flyers across the whole campus in which we mentioned the ongoing debate on education and described the 3 available methods of giving feedback (Fig. 7),
- we cross-promoted on each medium the other two mediums, in order to give respondents the possibility to choose in which they would prefer to give their comments.

Fig. 6. The interfaces used in all three mediums of Study 3. From left to right: Public interactive display, web form, paper form (English version on the reverse side of the sheet).

Fig. 7. One of flyers used to promote Study 3.

In a last-minute experimental manipulation we decided to extend the study for a second month, but this time we removed the public display feedback application and only used the other two mediums. This way we wanted to establish whether there was any effect between mediums. We continued promoting the two remaining mediums between each other.

3.3.2. Measures

All metrics captured in Study 2 were also logged for the public display and web form in this study. We added a further measure, *time between last character typed and submission*. With this additional metric we intended to establish whether people pondered or checked what they wrote before submitting their feedback.

3.3.3. Results

At the end of the deployment we gathered the feedback from every medium two independent raters classified each feedback statement as being relevant (e.g., "*Libraries should be open longer and even on weekends.*") or irrelevant (e.g., "*qwerty*", and "*beep beep I'm a jeep*"). There was 100% agreement between the 2 raters. We then measured the level of noise for each medium by calculating the ratio of irrelevant comments to the total number of comments.

We also measured the average number of characters typed for each medium. The results obtained for this deployment can be seen in Table 2. Furthermore, we show the progression of feedback obtained throughout the initial 30 days of deployment for each of the 3 mediums in Fig. 8. In Fig. 8 it is also possible to observe that there was a steady amount of feedback getting submitted during the 30 days, suggesting no significant effect of novelty. As for the second month of deployment (i.e., after we had removed the public display application) we found that, surprisingly, the paper and web forms received *no feedback* at all.

3.3.4. Findings

In Study 3 we found that even though the public display had by far the most responses, they were mostly "noise". On the other hand, the web form had fewer reports but they were all relevant. Additionally, respondents using the web form spent more time to construct well-thought out comments and were more concerned with correcting mistakes they made. This is also reflected on the average amount of characters typed with the web forms being far beyond those typed on the other two mediums. Respondents on the web also seemed to ponder for some time about what they wrote before submitting their comment.

Table 2
Results obtained in Study 3 for all three mediums.

	Public display	Web form	Paper form
Total comments	72	20	13
Avg. typing time (in seconds)	30.9	1967.5	N/A
Avg. time between last character typed and submission (in seconds)	2.3	194.2	N/A
Avg. number of backspaces/deletes	3.9	160.5	N/A
Avg. number of characters	22.9	571.5	43.2
Noise	95.8%	0%	23.1%
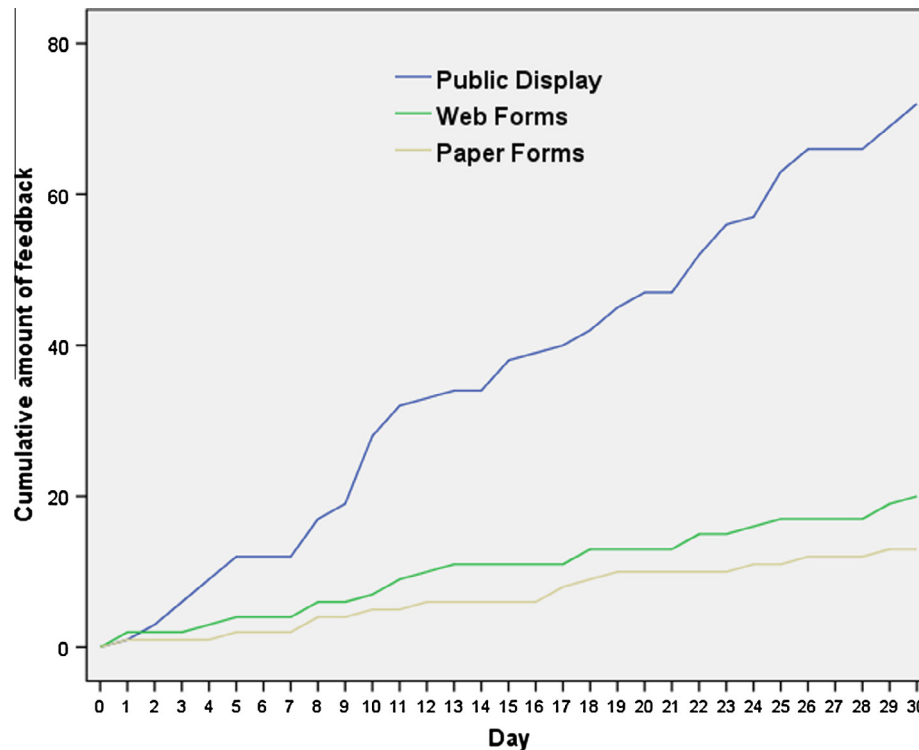
Fig. 8. Progression of feedback gathered throughout the initial 30 days of deployment of Study 3 for each of the 3 mediums.

Regarding the paper forms, we found that while they were located adjacent to the public display they received much less attention than the screen. At the same time, however, a larger portion of the paper feedback statements was relevant. This indicates that the majority of people that wanted to give more thoughtful and appropriate feedback in situ chose to do so on the paper forms.

However, our most interesting finding came from our last-minute manipulation to extend the study for a second month without the use of the public display. Our results suggest that the absence of the public display diminished participation in all other mediums. One interpretation of these results is that a number of participants first became interested in the topic by encountering it on the public display, but chose to use an alternative medium to provide their feedback. Another interpretation, even though unlikely to be the sole cause due to the steady influx of feedback during the first month (Fig. 8), is that people simply ran out of things to say. Furthermore, the high amount of noise on the public display suggests that people who did encounter the topic but did not get interested were likely to provide "noisy" feedback on the display itself.

4. Discussion

This paper set out to investigate (i) how the situatedness of public displays affects the quantity and quality of collected feedback, and (ii) how public displays fare against traditional paper or web feedback mechanisms. Considering our results across the three studies we found that while public interactive displays can be a powerful medium to promote civic engagement it can also be rather "noisy" when deployed in a natural setting. This finding contrasts previous research on feedback collection using public displays which have been conducted in a contrived setting [7]. The likelihood of participants typing in gibberish or noisy comments is much lower in such contrived settings as opposed to a natural setting.

Furthermore, we acknowledge that a variety of text input mechanisms exists, but it is not in our interest to test these on public displays; this has been thoroughly investigated [e.g., 19,26,38]. Rather, we are interested in identifying how public displays are used in *genuine* settings. An important challenge in this approach is that users cannot be coached or trained, and therefore experimenting with exotic text entry mechanisms is beyond the scope of our work. Rather, we chose to deploy a conventional text entry mechanism to investigate the *quality* of feedback we received, not the number of spelling mistakes.

4.1. Irrelevant yet purposeful feedback

Study 1 showed that while people we interviewed have a clear idea of good ways to give feedback on interactive public displays, these do not necessarily work in a naturalistic setting. For example

the Likert scale responses, while fast and easy, appear to produce extremely unreliable results. These results imply that the majority of people probably did not read the statements or did not choose to answer in a serious manner.

Furthermore, the soft keyboard and free text combination, while described by all stakeholders as an ideal mechanism for providing concrete and objective feedback, also failed to produce feedback of high quality. We attempted to investigate *why* this was the case by designing Study 2 where we contrast a public display to two traditional mediums of giving feedback: paper and web forms.

One of our main findings is that public displays elicit noisy and non-serious feedback, regardless of the instructions given by the system. To date this has not been reported or measured explicitly, but rather Ubicomp researchers have been somewhat forgiving of the data they collected. For instance, researchers mention "appropriation" [18] of the technology in this context, but do not explicitly acknowledge that this leads to results and data that are simply not useful for the purposes of the deployment. Therefore, we argue that our first finding in terms of noisy feedback on public displays is non-trivial, and substantiated by a series of our results. Additionally, an in-depth analysis of these messages reveals several latent social phenomena, which we discuss next.

4.1.1. Playing with technology

Several messages collected in Studies 1 and 3 consisted of random characters, such as "v811", "l00000l", or "zadffffgghhjjkkjhbbb". We argue this was caused by users *playing with the technology* and creating a mutual social event around the display rather than focusing on the application purpose itself. This is often observed with groups interacting with mobile devices in public [21], and public displays' inherent encouragement of social and performative use [31,33] strengthens this effect. Similarly, previous work [18] has reported groups of users socializing together around public displays as a major factor in obtaining high numbers of (arguably noisy) feedback.

4.1.2. Self-representation and expression

Self-representation and expression were identified in several feedback messages. These are commonly observed in public spaces in the context of photography [41] and refer to an individual's needs for highlighting his/her activities, humor, or any unique identifiable angles around oneself, i.e. bringing oneself forward. In this case even the visually modest textual feedback channel was enough for users to submit their names or nicknames, affiliations, or mental and physical states. Comments like "I'm Sniff Dogg and I'm a wild guy!", "Beep beep I'm a jeep", and comments with names or nicknames all represent the strong desire that people feel for expressing and advertising themselves in their appropriation of new communication technologies [15].

4.1.3. Documenting rule-breaking

The third observed phenomenon was *documentation of rule-breaking*, a social need according to Schwarz [39]. Users submitted messages of breaking social rules and norms, or ridiculing the authority that receives the messages, the University – without intention to discuss the educational issues. Messages like "I really don't appreciate you…", "It stinks like s∗∗t here, f∗∗k you all!", "I'm way too drunk to give you any constructive criticism, sorry a∗∗holes.", and random swear words all indicate acts of documenting the breaking of rules of social behavior and norms. A free form, anonymous channel like ours is likely to receive abuse like this when deployed in the wild.

4.1.4. Storytelling and discussions

Storytelling and discussions are series of submitted messages that complement or continue the previously submitted message.

Storytelling was illustrated, for example, in the following three comments submitted sequentially in Study 1 during 3 min: "I am a 12 year old girl.", "I am a 12 year old girl, from <Location>!", and "I am a 12 year old girl, from <Location>, but it was not me! It was <nickname>". Self-expression and storytelling are both clearly distinguishable from such sequences. Considering that typing with virtual keyboard was judged cumbersome and frustrating in Study 2, such social play around a display has great potential to overcome those difficulties.

4.2. Users' perceptions of public displays

The majority of interviewees in Study 1 stated they would prefer to use the public display to give their feedback because it was interesting and "cool", something also reported in [33]. However, we collected contradicting remarks from those who participated in Study 2, where most participants claimed that they would prefer to give feedback at home if it required a well-thought out comment. The drawback of this is that people are likely to forget and not be motivated to give their feedback upon arriving home, unlike during a serendipitous encounter [24] with a public display.

Previous work has reported that providing feedback on public interactive displays is challenging [3,18], and Study 2 gave us insights on explaining why and in what situations displays would be the preferred medium. For instance, some participants mention discomfort in providing feedback in public:

> "If I wanted to give feedback right there targeted to someone, i.e. had to give names, I would feel uncomfortable using the public display for it and would just use the paper forms"

Another participant suggests that this was caused by the privacy drawbacks of the public display:

> "Prefer to use paper forms for personal / sensitive feedback. For more casual feedback public display is fine, although I'd prefer to have a smaller text-box so that I could occlude the text I'm typing with my body."

We wish to highlight that the above 2 statements were of particular importance to the participants if they wanted to give negative feedback. These concerns took a back seat if their feedback was not negative. This can further explain the low quantity of relevant feedback given on the public display, as all feedback given across the 3 mediums was of negative nature. There is the possibility that for occasions where more positive feedback is generated that the ratio of "noise" on the public display would be considerably lower.

As for positive side of the affordances of public displays, some participants mentioned that the situatedness of the public display could be appealing in certain circumstances:

> "If I just finished a horrible lecture, I'd definitely use the public display or ballot box to give feedback. I probably would not care about it anymore when I got home. I would like to just spit it out right away."

> "The type of content matters a bit, not that much. The length is a bigger issue – for short and spontaneous feedback the public displays is ok! For longer and more thoughtful whining, using online app from home is the best."

In certain cases, however, the situatedness of the display imposes a sense of "urgency" that discourages potential users:

> "I would like to have my own time and edit my text carefully before submitting anything – maybe at home is the best then at least for more in-depth and emotional content. So it depends on the content also, for something lightweight a public display or ballot box is ok."

4.3. Impact of physicality and situatedness

Study 3 was designed to explicitly test the impact of situatedness and medium on the quality and quantity of collected feedback – in a naturalistic setting.

An unexpected result was the discrepancy in the results between the public display and the paper form, even though they were *co-located*. The public display attracted about five times more comments than the paper form. This discrepancy highlights the general attractiveness and lure of a public display that is often noted [e.g., 17,28,32]. However the actual number of *relevant* comments was higher for the paper forms: 10 compared to the 3 obtained by the public display. We argue that this may have been caused by several different reasons, which we attribute to the affordances of each medium. Specifically, one explanation is that some participants might have wanted to give feedback "right there, right now" but in a more private way. This interpretation is supported by the comments given by participants of Study 2.

Furthermore, the web form elicited much lower volume of feedback, but in fact a higher number of relevant comments than the public display. As also noted by participants in Study 2, the higher volume of feedback on the display can be mostly attributed to its situatedness and strong physicality. Since these two mediums were *not co-located*, we assume that those that sent their feedback through the web did not feel the need to comment "right there, right now". Those who choose the web form took significantly more time to write their comments including a significant reflection period between finishing writing and submitting. Additionally, these participants typed significantly more characters in their comments, over 20 times more than those on the public display and over 10 times more those paper forms. This suggests that the web form is the ideal vehicle for thoughtful and insightful feedback, as also supported by comments given by our participants.

To summarize, we found that the public display's situatedness resulted in more feedback than the web form, but also more noisy feedback. However, situatedness on its own does not fully account for this difference: we found that the questionnaire, which was also highly situated, in fact received even fewer responses than the web form. Therefore, we argue that a combination of the hedonic appeal of the public display, and its strong physical presence, account both for the large volume of feedback and high ratio of noise on public displays. The differences between the mediums can to some extent also be explained by the *prequalification* [42] that exists in online environments: users have a clear intent when accessing certain websites. This in turn prequalifies the user as someone that accessed the platform with the purpose of leaving feedback unlike those that stumble upon the public display and paper forms.

Finally, in Study 3 we saw a drastic change in our data after the removal of the feedback application from the public display. We saw the other two feedback mechanisms going from a steady rate of feedback (Fig. 8) with a high quality during 1 month to absolutely no feedback at all the next month once the public display ceased to be a factor. While in general we would expect interest in our study to diminish over time, the volume of feedback fell straight to zero as soon as the public display was removed from our study. In this sense, the flatlining of feedback volume highlights the importance of the public display as an interest generator even though this may have not been the only reason for this to happen. For instance, the flatlining of feedback could have also been caused by users that had any particular opinion to express having done so already during the first month of deployment.

4.4. Design implications

We believe there is a substantial "missed opportunity" on public displays. While they generate much more interest than websites and paper forms, they suffer from playfulness or lack of seriousness on the users' part. We therefore consider that a potential research challenge is to identify ways to harness the interest and opportunities that public displays create, and minimize the feedback noise by "convincing" or "nudging" respondents to give useful rather than playful responses. Barriers to filter noise have been shown to be very effective in crowdsourcing markets. For example, researchers have proposed that crowdsourcing tasks make use of fact-checking questions whose answer is known to researchers (e.g., "who is the president") but also is apparent to responders that the answer is known to researchers [22]. Such noise filtering mechanisms on public displays have been shown deter non-serious responders and dramatically reduce noise [11].

Furthermore, our results indicate that it is crucial to avoid lengthy typing on the display as documented in previous research [e.g. 28,38]; however using an external device to facilitate this is not always possible [19]. Interviewees in Study 1 expressed concerns in spending time with pairing mechanisms on the public display or spending money in the case of an SMS-based solution. One potential solution would be to introduce an auto-complete mechanism on these public displays that has been shown to be successful in desktop environments [40]. By providing a list of words the user might be trying to write, it would potentially lower the amount of time needed and therefore reduce the frustration experienced by users.

4.4.1. Limitations

We note a number of limitations of this study. First, as expected during Study 1 some of the public displays suffered periods of downtime due to malfunctioning touch screens or temporary shutdown caused by the reconstruction going on around them. However, these were all quickly fixed therefore not having a significant impact in the results we obtained.

Furthermore, the Study 1 deployment was done on multipurpose public displays in which our application was one of many. While this might have affected the total amount of feedback gathered during Study 1, it was necessary in order to observe people's interest and openness in using our citywide installation to provide feedback. This limitation led us to a fully controlled study (Study 2), followed by an "in-the-wild" study but with only one display with the feedback application on-screen at all times in order to mitigate the problem mentioned previously.

We also recognize that while we attempted to promote all three mediums in Study 3 equally, the public display ultimately got more attention due to its "self-promoting" caused by its very nature [18,24,28]. Despite this, we took the appropriate measures to attempt to mitigate this issue by setting up flyers all around campus including locations far away from the display and ballot box, and by adding cross-promotion in all three mediums.

Another limitation was the available input methods. The decision to mostly use direct text entry was mainly motivated by the feedback gathered during our initial interviews and to keep the comparison between mediums as fair as possible. Furthermore, we chose not to overly complicate this initial study, leaving the exploration of further input methods for future work.

Finally, we did not explore if and how different demographics of users reacted to the 3 choices of medium given. Unfortunately, in such in-the-wild studies obtaining users' demographics is not an easy task. While possible on the display and web form, we would not have been able to guarantee that the information provided was reliable, and it would also be an extra barrier to participation. As for the paper, we would have no way of "forcing" users to give the information. Due to these reasons we opted to avoid probing the differences between demographics.

5. Conclusion

Investigating public feedback mechanisms is by no means new. However, no studies have systematically investigated this topic in the context of public interactive displays. The findings in this paper give valuable insight regarding the use of public interactive displays to elicit situated feedback.

First, we show that **feedback on public displays is noisy**, and while in the case of text feedback this is easy to filter gibberish, it is not possible when using other instruments like Likert-scale questions. Furthermore, we show that compared to web and paper forms, public displays are more likely to receive noisy feedback, but do **manage to attract more comments**.

However, not all news is bad for public displays. In our study we have strong evidence suggesting that such **displays are instrumental in generating interest** on a particular topic, and funneling respondents to other mediums. Hence, an appropriate strategy is to attempt to mirror a public display feedback mechanism online, and promote this additional medium on the public display itself.

References

[1] J.R. Alford, We're all in this together: the decline of trust in Government, 1958–1996, in: What is it about Government that Americans dislike? Cambridge University Press, 2011, pp. 28–46.

[2] F. Alt, T. Kubitza, D. Bial, F. Zaidan, B. Ortel, B. Zurmaar, T. Lewen, A. Shirazi, A. Schmidt, Digifieds: insights into deploying digital public notice areas in the wild, in: Proc. MUM '11, ACM, 2011, pp. 165–174.

[3] M. Ananny, C. Strohecker, TexTales: creating interactive forums with urban publics, in: M. Foth (Ed.), Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City, IGI Global, 2009, pp. 68–86.

[4] A. Bajaj, S. Nidumolu, A feedback model to understand information system usage, Inf. Manage. 33 (1998) 213–224.

[5] P. Battino, A. Vande Moere, V. Barsotti, Situated & social feedback in the city, in: Workshop on Large Displays in Urban Life from Exhibition Halls to Media Facades, CHI'11, ACM, 2011.

[6] N. Bendapuni, R.P. Leone, Psychological implications of customer participation in co-production, J. Market. 67 (1) (2003) 14–28.

[7] H. Brignull, Y. Rogers, Enticing people to interact with large public displays in public spaces, Proc. INTERACT 2003 (3) (2003) 17–24.

[8] R.B. Cialdini, N.J. Goldstein, Social influence: compliance and conformity, Annu. Rev. Psychol. 55 (2004) 591–621.

[9] M. Cooper, G. Cooper, Overcoming Barriers to the Positive Development and Engagement of Ethno-Racial Minority Youth in Canada, Canadian Heritage, 2008.

[10] E.F. Churchill, L. Nelson, Interactive community bulletin boards as conversational hubs and sites for playful visual repartee, in: Proc. HICSS'07, IEEE, 2007, pp. 76–85.

[11] J. Goncalves, D. Ferreira, S. Hosio, Y. Liu, J. Rogstadius, H. Kukka, V. Kostakos, Crowdsourcing on the spot: altruistic use of public displays, feasibility, performance, and behaviours, in: Proc. Ubicomp'13, ACM, 2013, pp. 753–762.

[12] J. Goncalves, V. Kostakos, S. Hosio, E. Karapanos, O. Lyra, IncluCity: using contextual cues to raise awareness on environmental accessibility, in: Proc. ASSETS'13, ACM, 2013, article 17.

[13] Goncalves, J., Kostakos, V., Karapanos, E., Barreto, M., Camacho, T., Tomasic, A., Zimmerman, J. Citizen Motivation on the Go: The Role of Psychological Empowerment. Interact. Comput. (2013), online first.

[14] D.P. Green, J. Citrin, Measurement error and the structure of attitudes: are positive and negative judgments opposites?, Am J. Polit. Sci. (1994) 256–281.

[15] R. Harper, Texture: Human Expression in the Age of Communication Overload, MIT Press, 2011.

[16] T.A. Heberlein, Principles of Public Involvement, University of Wisconsin, Department of Rural Sociology, Madison, 1976.

[17] S. Hosio, J. Goncalves, V. Kostakos, Application discoverability on multipurpoe public displays: popularity comes at a price, in: Proc. PerDis'13, ACM, 2013, pp. 31–36.

[18] S. Hosio, V. Kostakos, H. Kukka, M. Jurmu, J. Riekki, T. Ojala, From school food to skate parks in a few clicks: using public displays to bootstrap civic engagement of the young, in: Proc. of Pervasive'12, Springer-Verlag, 2012, pp. 425–442.

[19] S. Hosio, H. Kukka, M. Jurmu, T. Ojala, J. Riekki, Enhancing public interactive displays with social networking services, in: Proc. MUM '10, vol. 23, 2010.

[20] R. Johnson, Y. Rogers, J. van der Linden, N. Bianchi-Berthouze, Being in the thick of in-the-wild studies: the challenges and insights of researcher participation, in: Proc. CHI '12, ACM, 2012, pp. 1135–1144.

[21] T. Kindberg, M. Spasojevic, R. Fleck, A. Sellen, The ubiquitous camera: in-depth study of camera phone use, IEEE Pervasive Comput. 4 (2) (2005) 42–50.

[22] A. Kittur, E. Chi, B. Suh, Crowdsourcing user studies with Mechanical Turk, in: Proc. CHI '08, pp. 453–456.

[23] S. Ko, K. Kim, T. Kulkarni, N. Elmqvist, Applying mobile device soft keyboards to collaborative multitouch tabletop displays: design and evaluation, in: Proc. ITS '11, ACM, 2011, pp. 130–139.

[24] H. Kukka, H. Oja, V. Kostakos, J. Goncalves, T. Ojala, What makes you click: exploring visual signals to entice interaction on public displays, in: Proc. CHI'13, ACM, 2013, pp. 1699–1708.

[25] I. MacKenzie, S. Zhang, The design and evaluation of a high-performance soft keyboard, in: Proc. CHI '99, ACM, 1999, pp. 25–31.

[26] I. MacKenzie, S. Zhang, R. Soukoreff, Text entry using soft keyboards, Behav. Inf. Technol. 18 (4) (1999) 235–244.

[27] T. Mannarini, A. Fedi, S. Trippetti, Public involvement: how to encourage citizen participation, J. Commun. Appl. Soc. Psychol. 20 (2010) 262–274.

[28] J. Müller, F. Alt, D. Michelis, A. Schmidt, Requirements and design space for interactive public displays, in: Proc. MM '10, ACM, 2010, pp. 1285–1294.

[29] J. Müller, M. Jentsch, C. Kray, A. Krüger, Exploring factors that influence the combined use of mobile devices and public displays for pedestrian navigation, in: Proc. NordiCHI '08, ACM, 2008, pp. 308–317.

[30] J. Müller, R. Walter, G. Bailly, M. Nischt, F. Alt, Looking glass: a field study on noticing interactivity of a shop window, in: Proc. CHI '12, ACM, 2012, pp. 297–306.

[31] T. Ojala, V. Kostakos, H. Kukka, T. Heikkinen, T. Linden, M. Jurmu, S. Hosio, F. Kruger, D. Zanni, Multipurpose interactive public displays in the wild: three years later, IEEE Comput. 45 (5) (2012) 42–49.

[32] T. Ojala, H. Kukka, T. Lindén, T. Heikkinen, M. Jurmu, S. Hosio, F. Kruger, UBI-Hotspot 1.0: large-scale long-term deployment of interactive public displays in a city center, in: Proc. ICIW'10, 2010, pp. 285–294.

[33] P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, P. Saarikko, It's mine, don't touch!: interactions at a large multi-touch display in a city centre, in: Proc. CHI'08, ACm, 2008, pp. 1285–1294.

[34] 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: Proc. UbiComp '07, ACM, 2007, pp. 336–353.

[35] Y. Rogers, Interaction design gone wild: striving for wild theory, Interactions 18 (4) (2011) 58–62.

[36] E. Rukzio, M. Müller, R. Hardy, Design, implementation and evaluation of a novel public display for pedestrian navigation: the rotating compass, in: Proc. CHI '09, ACM, 2009, pp. 113–122.

[37] W.E. Saris, M. Revilla, J.A. Krosnick, E.M. Shaeffer, Comparing questions with agree/disagree response options to questions with item-specific response options, Survey Res. Methods 4 (1) (2010) 61–79.

[38] C. Sas, A. Dix, Designing and evaluating mobile phone-based interaction with public displays, in: Proc. CHI'08 EA, ACM, 2008, pp. 3941–3944.

[39] O. Schwarz, Subjectual Visibility and the Negotiated Panopticon: on the Visibility-Economy of Online Digital Photography, 2011.

[40] T. Stocky, A. Faaborg, H. Lieberman, A commonsense approach to predictive text entry, in: Proc. CHI '04 EA, ACM, 2004, pp. 1163–1166.

[41] N.A. Van House, Flickr and public image-sharing: distant closeness and photo exhibition, in: Proc. CHI'07 EA, ACM, 2007, pp. 2717–2722.

[42] P. Zhang, G. von Dran, Satisfiers and dissatisfiers: a two-factor model for website design and evaluation, J. Am. Soc. Inf. Sci. Technol. 51 (2000) 1253–1268.