



Designing and evaluating mobile self-reporting techniques: crowdsourcing for citizen science

Eman M. G. Younis¹ · Eiman Kanjo² · Alan Chamberlain³

© The Author(s) 2019

Abstract

In recent years, mobile phone technology has taken tremendous leaps and bounds to enable all types of sensing applications and interaction methods, including mobile journaling and self-reporting to add metadata and to label sensor data streams. Mobile self-report techniques are used to record user ratings of their experiences during structured studies, instead of traditional paper-based surveys. These techniques can be timely and convenient when data are collected “in the wild”. This paper proposes three new viable methods for mobile self-reporting projects and in real-life settings such as recording weather information or urban noise mapping. These techniques are Volume Buttons control, NFC-on-Body, and NFC-on-Wall. This work also provides an experimental and comparative analysis of various self-report techniques regarding user preferences and submission rates based on a series of user experiments. The statistical analysis of our data showed that pressing screen buttons and screen touch allowed for higher labelling rates, while Volume Buttons proved to be more valuable when users engaged in other activities, e.g. while walking. Similarly, based on participants’ preferences, we found that NFC labelling was also an easy and intuitive technique when used in the context of self-reporting and place-tagging. Our hope is that by reviewing current self-reporting interfaces and user requirements, we will be able to enable new forms of self-reporting technologies that were not possible before.

Keywords Mobile sensing · Mobile self-report · Pervasive computing · Citizen science · User interfaces

1 Introduction

Citizen science projects [14, 17] enable members of the public to be involved in scientific research and provide crowdsourced data. Despite an increasing interest in citizen science, some projects are reporting decreasing participation levels [42]; for

example, the California Water Resources Volunteer Climate network, established in the 1950s, currently has the lowest number of participants since it started [7]. Another example, the European project Evolution MegaLab had more than 6000 registrations, but only 38% of them submitted data [43].

According to a recent citizen science report, there appears to be a positive relationship between the number of ways of submitting data available and submission rates [41, 42]. Some submission and reporting apps are difficult to use or not practical.

When collecting data in “the wild” [5, 6] a participant could be doing anything from driving a car to cooking a meal. In particular, labelling while engaged in physical activities, such as walking is difficult. Taking into account, the nature of the current activity needs to be considered, both at UX design stage, for data providers and users, and at the experimental “set up” stage, before running the experiments.

An abundance of smart phone applications makes participation in citizen science projects an easy process that can fit in every day’s task. Especially, recent advances in pervasive technologies have allowed engineers to transform bulky and inconvenient monitors into relatively small, comfortable, and ergonomic research tools.

✉ Alan Chamberlain
Alan.Chamberlain@Nottingham.ac.uk

Eman M. G. Younis
Eman.Younas@mu.edu.eg

Eiman Kanjo
eiman.kanjo@ntu.ac.uk

¹ Information Systems Department, Faculty of Computers and Information, Minia University, Minia, Egypt

² Computing and Technology, Nottingham Trent University, Nottingham, UK

³ School of Computer Science, University of Nottingham, Nottingham, UK

Today, scientists are harnessing the power of citizen science using newer and smarter technologies; there are many techniques to choose from; these techniques depend on the types of the application and mode of interaction to collect the data, which can be categorised into three main categories:

- **Passive:** which relies on the sensing capabilities of the smart phone to enable automated collection of smartphone generated data, for example by the accelerometer, GPS data, and ambient noise levels [15]. Many of these tools are often used simultaneously to gather multimodal data from different sources. When accompanied with location tracking, these tools enable a broad range of applications ranging from weather observation and recording [22, 23, 25, 28] to health and well-being mapping [2, 19, 34], and environmental monitoring [15, 16, 27].
- **Active:** active sensing tasks that ask for an active contribution from users such as taking a picture, tagging a place, or sending a text message. These types of tasks can be performed using self-reporting techniques which are usually designed to allow simple and intuitive mechanisms to collect subjective data. These labels are often triggered by an event in the surrounding environment (e.g. finding a particular type of species or drop in weather temperatures).
- **Hybrid:** when the user is actively reporting subjective data while passively recording sensor readings [19].

This paper presents an overview of the recent trends in self-report techniques which are particularly useful when users are on-the-go in active or hybrid mobile modes.

We propose three new self-report techniques “in-the-wild” that can be utilised as alternative methods for on screen self-report. These techniques include the use of NFC-on-Body and NFC-on-Wall, as well as using Volume Buttons to make a quick user input, which is useful when users are busy doing other activities, while collecting data. For examples, reporting the type of Ladybird found on the way to work or number of Daffodils found near the school.

Unlike the traditional techniques, these newly proposed methods do not require the user to handle the screen of the mobile phone to annotate data manually. However, each of these techniques poses a challenge related to interference, usability, or data collection frequencies.

In the context of this paper, we reason beyond the ease of use of the interface and interaction mechanisms (which by themselves are important). We look at the wider system, regarding the ease of use for both the participant in terms of providing appropriate data and for research scientists who use the data to further science.

The next section briefly reviews the related work, in the area of mobile journaling and design consideration for mobile self-reporting. This is followed by a section describing the three new proposed self-report techniques. Then, we outline

the architecture of our prototype, followed by the results of the user study. We then reflect on and discuss the results and the challenges in the development of self-reporting techniques for data collection for citizen science.

2 Background

This section is divided into two parts. The first presents some of the current self-reporting techniques. The second part discusses the design considerations and challenges for the relevant mobile self-reporting systems.

Self-reporting is the most traditional method for gathering feedback from users for various user studies and applications [30, 36]. It is a subjective measure in which participants are asked to annotate data manually.

Recently, there has been a growing demand to gather subjective user input while collecting mobile and sensor information. This is associated with the increasing demand for machine learning developments for intelligent and context-aware applications such as human activity recognition and environmental monitoring, health, and smart cities applications [3, 8, 9, 11, 12, 16, 18–20, 32, 39].

In the past, researchers used surveys and interviews to elicit self-report information [36]. With the development of computers and mobile phones, new approaches have emerged. One of the most common self-report measurements is the verbal scale in which words are used to describe the participant’s feedback. Text could be collected using simple text buttons or swipe buttons as shown in Fig. 1:



Fig. 1 Screenshot of Mappiness application [24]

Mappiness [24] is a smartphone application that utilises text-based self-reporting to elicit users' feelings based on three emotional dimensions: "happy", "relaxed", and "awake". These are then used to investigate how their local environment (e.g. air pollution, noise, and green spaces) affects people's happiness. Text-based self-reporting is also employed by WiMo [26] to allow users to express and share their feelings about a particular outdoor place according to two scales: "comfortable" to "uncomfortable" and "like it" to "don't like it".

In addition to text, Emojis were used to collect self-report information, which have been adopted for individuals with low or no computing knowledge [10, 29].

Similarly, [37] employed mobile phones and wearable sensors to report pain using Emojis. While, Reid et al. [35] adopted mobile Emojis to capture users' feeling in response to music.

Pictograms or animated cartoons also have been used as an alternative to declarative words. Examples of such systems are the Self-Assessment Manikin [4] and Emocards [21].

Furthermore, few research projects have adopted screen touch and slide technique as the primary method of self-reporting. For example, Xiaoyi Zhang et al. [44] used self-report for quantify-self, comparing their approach to data collection with traditional diaries and notification-based reminders for health and well-being measures.

On the other hand, device gestural interaction based on hand movements are increasingly being used to register user input, specifically in games, for improving entertainment and providing experiences with the purpose of promoting better physical and well-being [40]. Jaime Ruiz et al. [38] conducted a usability study on the gestures as a way of interaction. They concluded that although using gesture is useful in some applications, some gestures may be socially unacceptable.

Similarly, self-reporting based on speech could offer a new input modality which removes the need for direct interaction with a mobile user interface. This method provides a naturalistic setting for data collection which might be the only possible way when users are full-handed (they provided an external microphone connected to the mobile directly). Although audio input might not be suitable for all data collection applications, some specific applications might benefit from this type of data collection on the go which relies on voice recording only [13].

Harada et al. [13] proposed a voice-based self-report system to collect voice labels from users while engaged in physical activities. Then, they developed an extra application to label the collected sensor data "offline" for activity recognition. Although using speech is appropriate for this type of application, it still requires an additional labelling step.

In relevant work [19], we utilised on screen self-report to collect user's feedback on outdoor places while passively collecting user physiological measurements and

environmental measurements for emotion recognition. The aim of the study was to utilise citizen science techniques and mobile sensing to explore and map the environmental impact of environment on health.

The Sense-it App [33] provides abstracted access to all sensors on Android smart phones. Though harnessing the sensing capabilities of mobile devices has proved a boon to Citizen Science, the issue of harnessing external sensors embedded within the environment is one that has not as yet received significant attention by the research community.

2.1 Design considerations of self-report apps

In general, self-report methods can be classified into two in relation to data collection mode:

- "Optional": gives participants a full control over what they write and the length of their report. The user may be given the option to write a report.
- "Forced": constrains participants to specific questionnaire items which could vary from simple check boxes to multiple choice items or pictures.

Also, it could be classified into two categories pertaining to the data labelling process:

- "Offline": The data is first collected and then the researchers, experts, or the participants are required to label the sensor data after it has been collected. This is an expensive and a time-consuming process.
- "Online": requires the participants to record their input in real-time while collecting other sensor data. This method is effective for developing rapid machine learning models; however, data can be noisy as a result of the mode of interaction.

Selecting one self-report technique from the aforementioned categories depends on the application and the surrounding environment. If the application requires huge amount of data feed, then real-time labelling is a better option as off-line labelling will be laborious and time consuming, especially when the collected data is in open space with high level of variations. Similarly, forced data collection method is more sensible option in this case since it allows an easy and quick input from a pre-determined list of labels.

Gathering mobile self-report is not a trivial process and can be particularly challenging when participants are moving due to the following reasons:

1. Participants have short time frame to think of a response to a questionnaire.
2. Participants will not be able to concentrate on mobile screen as they are walking.

3. This is particularly relevant to the automatic human activity recognition task.
4. If the data collection process involves the use of sensors, the act of providing a label can interfere with the sensing. For example, when the experiment requires contentious accelerometer data annotation. In this case, the self-report process will interfere with the accelerometer measurements.
5. The user might be carrying a bag, which makes it awkward to use two hands when interacting with the mobile.

An ideal data collection and labelling system should be unobtrusive, and the process of labelling the data should not affect the data being collected. But interacting with a Graphical User Interface (*GUI*) on a mobile device can be extremely obtrusive, especially when the device is physically attached to sensors such as accelerometers.

Interaction and recording constraints may not always be physical in nature. Social constraints may also affect obtrusiveness of an interface. For example, recording audio or video might be considered acceptable when performing physical activities; however, it might be considered obtrusive if collecting data during a meeting or in public spaces where others may fear being recorded.

3 New approaches for self-reporting

Current approaches for self-report are limited and often rely on the screen or voice input to collect user feedback. However, more innovative self-report techniques are required to keep up with the increasing demand for conducting user studies in the wild while the user is on-the-go.

To address these design considerations, we propose alternative approaches to traditional mobile labelling techniques when the user is engaged in any physical activity related to data collection. In this section, we present three new methods for mobile self-report which does not require direct screen manipulation.

3.1 Volume Buttons

Volume Buttons main functionality is to adjust the ringtone volume levels on the mobile phone. Here, we are proposing the use of mobile Volume Buttons as a method for collecting user self-report on the go. It is easily accessible for the user and non-obstructive. This technique does not require screen manipulation and interaction during data collection.

Usually, ringtone levels range between $\text{min} = 0$ to $\text{max} = 15$ (depending on the model) and can be logged

programmatically using Volume Control libraries to sense when the keys are up or down.

3.2 NFC-on-body

NFC is a short name for Near Field Communication, and it is a collection of close-range wireless communication standards. NFC-enabled phones have the ability to exchange information with each other. It is a sort of Identifier for a place or an object that can be easily identified by the phone. RFID short for Radio Frequency Identification used interchangeably with NFC. NFC is the new version of RFID; it is more secure as it allows the communications only from shorter distances. NFC tags are stickers that can be read by mobiles and devices equipped with that technology.

We chose Near Field Communication (NFC) as a technology, for the following reasons:

- NFC can be used to identify different body parts during data collection related to health experiments. For example, if the user was asked to report which part of the body is in pain at some point in time, it could be easy to identify different parts with NFC tags, and the user can just scan the intended one.
- NFC can be placed on the body as an accessory (e.g. bracelet necklaces), or directly attached to clothes or bags. For which the interaction paradigm is intuitive and does not need direct screen interaction, which makes it convenient for collecting user feedback while the user is moving or engaged in an intensive physical activity.

3.3 NFC-on-wall

To utilise NFC-on-Wall method, the tags should be placed at the intended place. The users are only required to scan the appropriate tag. NFC tags can be placed on a wall at a particular location or product as an ID.

The users might be asked about their ratings or emotions towards the place or product. In this case, they can just scan the tags of the places or products he likes. The limitation of this technique is that it cannot directly record multiple ratings for the same location or product, unless using multiple tags. It is suitable for binary ratings either like or dislike.

We have utilised NFC tagging previously to capture user proximity to outdoor places such as shops [1], as a part of our work on emotion mapping around urban areas.

4 Research methodology

4.1 Data collection

For this study, we have developed a data collection application named *EasyReport* which was built for Android platform. *EasyReport* is implemented for continuous and quick labelling while walking and collecting data. It passively records self-report input including audio, accelerometer data, screen swipe (left and right), button clicks, volume levels, and NFC scans (on-body and on-wall). Also, it records location coordinates based on *GPS* and logs time and date information. In order to get a meaningful recording, we asked the users to record their emotions ratings for the place, with scores ranging from low = negative to high = positive. When the user launches the application, mobile menu appears which offers seven different self-report interfaces, these are:

1. *Screen Buttons* (Image Buttons): five options ranging from very low (1) to very high (5).
2. *Screen Swipe* (left and right): left represents very low and right for very high.
3. *Device Gesture*: including two shakes (horizontal and vertical) based on the phone accelerometer motion. Horizontal shake represents higher value and vertical shake means lower value.
4. *Volume Buttons*: logs volume levels. Five levels are offered from one to five, one click means very low, and 5 means very high.
5. *Speech Labelling*: For this method, we used Android Speech API for voice recognition to recognise numbers.
6. *NFC-on-Body*: Two different NFC tags were positioned on the left and the right shoulder (scanning the right shoulder means you have positive emotions and the left one negative).

7. *NFC-on-Wall*: used to scan NFC tags placed on the wall (the tags were positioned in specific areas in the building. Two tags were positioned for each location, one for positive feeling and the other for negative).

We distinguish between the two *NFC* scan modes based on the *NFC ID*. A screenshot is presented in Fig. 2 showing the self-report system user interface.

4.2 Experimental setup

Thirty participants, 22 males and eight females, between the ages of 18–43 ($\mu = 26$, $\sigma = 5.4$) participated in the study. The participants all are students at Nottingham Trent University. The participants are invited to take part of this study during term time.

Each participant has been given a thorough introduction and a demo of the *EasyReport* application.

For simplicity, participants have been asked to perform the following tasks:

1. Use each labelling technique for 2 min while walking around a university building and try to record as much data as possible.
2. Record their preference of the seven self-report techniques based 5-levels Likert scale (5 highly in favour of the technique).

Participants were not explicitly told when to label a new activity; they had to remember.

Overall, we collected 5500 labels, 1912 button clicks, 720 screen touch, 419 volume changes, 446 gestures, 413 scan on-Body, and 404 *NFC* scan on-Wall.

Fig. 2 Two screenshots of the EasyReport Application

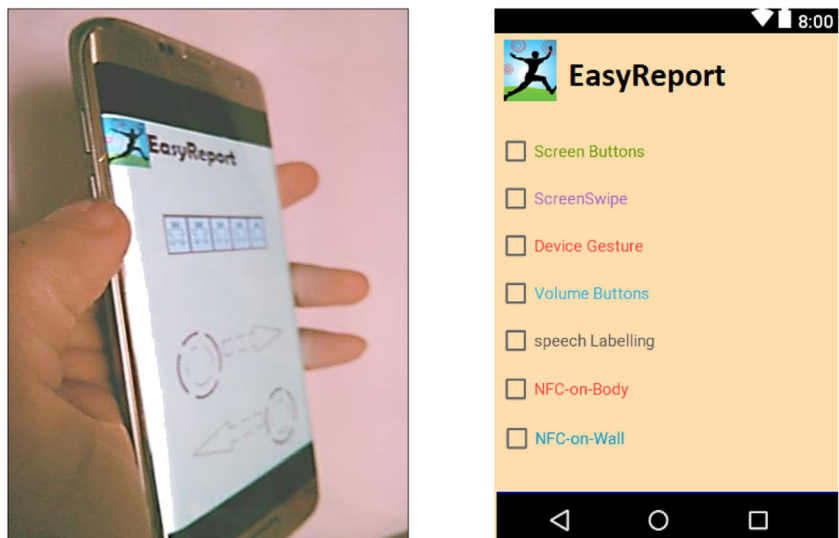


Table 1 User preferences for self-report methods

	Buttons	Screen	Volume	Gesture	Speech	NFC Body	NFC Wall
μ	3	3.03	3.41	2.13	3.10	3.10	2.23
σ	1.29	1.45	1.21	0.97	1.30	1.45	1.07
CV	0.43	0.48	0.35	0.48	0.42	0.47	0.48
Q1	1	2	3	2	1	2	1
Q2	3	3	4	3	2	3	2
Q3	4	4	4	4	3	4	3

They indicate the highest and the lowest values obtained

5 Results and analysis

5.1 Analysis of preferences

User preferences have been recorded for the seven self-report methods on a scale in the range (1–5).

Table 1 shows the main descriptive statistics including mean (μ), standard deviation (σ), coefficient of variations (CV), and quantiles (Q1, Q2, and Q3), comparing user preferences reported in the experiment.

A closer look at the descriptive statistics can reveal the following: *Volume Buttons* technique has the highest mean values compared to the other techniques. Moreover, it has the least coefficient of variation ($CV = 35$), whereas, *Device gesture* and *NFC-on-Wall* methods have the least mean values, and highest variation coefficient respectively.

Regarding the user preferences, we obtained Cohen's kappa Inter-user correlation coefficients in the range ($0.26 < ICC < 0.42$) with (p value < 0.0002) and 95% confidence interval for ICC Population Values. This is considered a moderate level of agreement between users. Cohen's kappa coefficient which is a statistic metric that measures the inter-rater agreement has been used. It is generally thought to be a more reliable measure than simple percent agreement calculation.

Without holding the assumption that user preference data are normally distributed, we decided at 0.05 significance level to test if the new self-report method (Volume Buttons) and the other

conventional methods such as screen touch, buttons, speech, or gestures have identical distribution of the user preferences.

The null hypothesis is that the Volume Buttons data of self-report have identical distribution with all the other methods. To test the hypothesis, we applied the *Wilcox*, which is statistical test function to compare the independent samples. As we obtained a p value of 0.0001, which is less than the 0.05 significance level, we reject the null hypothesis. So, we can conclude that the new proposed method of using Volume Buttons is significantly different in terms of the user preferences, in favour of the Volume Buttons method.

5.2 Analysis of labelling rates

Users labelling rates have been recorded for the seven self-report techniques. Table 2 shows the descriptive statistics (min, max, mean (μ), standard deviation (σ), coefficient of variations (CV), and quantiles (Q1, Q2, Q3), comparing labelling rates collected in the experiment. It can be noticed from the table that Screen Buttons technique has the highest mean values compared to the other techniques. Hence, it has relatively higher coefficient of variation compared to the other methods, whereas *NFC-on-Body* method has the least mean values. The Cohen's kappa Inter-user correlation coefficients in the range ($0.648 < ICC < 0.978$) with (p value < 0.00001) and 95% confidence interval for ICC Population Values. This is considered a very good level of agreement between users.

Table 2 Descriptive statistics for the labelling rates of the seven self-report techniques

	Buttons	Touch	Volume	Speech	Gesture	NFC Body	NFC Wall
min	26	26	19	20	8	10	10
max	95	54	32	56	21	18	16
μ	63.73	39.53	24	36.13	14.87	13.23	13.57
σ	23.16	8.16	4.12	11.35	4.45	2.81	2.04
CV	0.36	0.21	0.17	0.31	0.29	0.22	0.15
Q1	40.25	35.25	21	26	11	11	12
Q2	66	40	23	34	14.5	12.5	14
Q3	83.75	44.75	26.75	45.75	18	16	15

They indicate the highest and the lowest values obtained

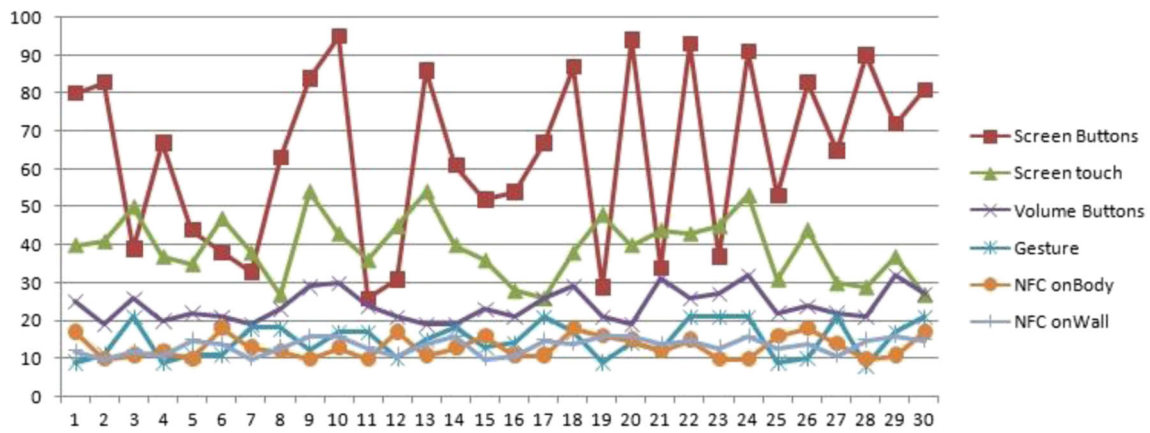


Fig. 3 Labels collected from 30 users across all the self-reporting approaches

Figure 3 shows a multiple line graph of the labelling rates for all the users across the seven labelling methods. It is noticed from the figure that screen buttons have the highest labelling rates and *NFC-on-Wall* has the least labelling frequency.

Without assuming that our data are normally distributed, we decided at 0.05 significance level to test if the new self-report methods (Volume Buttons) and the screen buttons have identical distribution of the frequencies of labelling.

The null hypothesis is that the Volume Buttons and the Screen Buttons input have identical distributions. To test the hypothesis, we applied the Wilcoxon Test to compare the independent samples. As the p value < 0.00005 , we reject the null hypothesis. This means that the Screen Buttons method has a significantly different distribution from the Volume Buttons method. Figures 4 and 5 show a visual analysis of the statistics of the seven self-report methods. The figures show that Screen Buttons technique has the highest labelling rate amongst the other methods, while Volume Buttons technique has the highest user preferences.

6 Discussion

The results from our user studies suggest that the utilisation of Volume Buttons and NFC tags as real-time self-report input are viable options depending on the required labelling rate and the

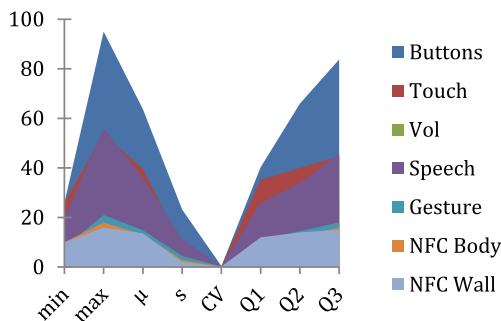


Fig. 4 The descriptive statistics of labelling, mean (μ), Sd (σ), CV, Q1, Q2, and Q3 for all self-report methods for all users

other sensor data to be collected in the study. Volume Buttons is the most practical method, which was favoured by the users.

For example, user 16 said:

I didn't like the idea of walking and staring on the mobile screen to record my input, using Volume-Buttons makes the whole process a lot easier and safer.

On the other hand, some of the users reported that *NFC-on-Body* self-report method offers a good and alternative option, but only if the labels are required at low rate. Also, they felt that they will be put off using these techniques when they are surrounded by people, for example user 10 noted:

I am concerned that I will look weird if I start tapping my body with my mobile phones constantly. I feel as if I need to explain my self all the time.

Similarly, users felt that *NFC-on-Wall* option can be easier and more enjoyable to use, only if there is a clear and large sign of the tag to be seen by the participants. For example, user 25 said:

I really enjoyed being able to tab the NFC signs on the wall, it gave me more flexibility, but I have missed few as the signs were not clear enough to me.

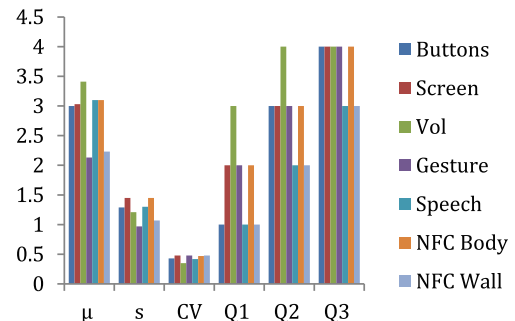


Fig. 5 The descriptive statistics of user preferences, mean (μ), Sd (σ), CV, Q1, Q2, and Q3 for all self-report methods

Although it may be beneficial for cognitive, affective capture and modelling, such real-time reporting has several limitations such as self-deception, intrusiveness, the subjective nature of the report, and issues relating to the generalisation of results. However, as we have discussed in the literature review, the use of self-reporting as a tool (as part of a collection of resources) to enable a researcher or designer to fully understand interaction, gather requirements, or even gather opinions is very powerful, particularly as a mechanism that one might employ to gather qualitative data. Nonetheless, like any methods or approaches that are employed to understand human behaviour, there are issues that anyone who carries out studies must be aware of; these issues or challenges (as we shall call them) can impinge upon the design of a data collection system, and ultimately could skew one's understanding of the data. Some of these issues are methodological, but in this section, we would like to address some of the more technical, practical use considerations and what we would call social issues that impact upon the use and design of mobile self-reporting systems.

In terms of understanding qualitative data, it is important that the researchers understand the context of use, as part of this there is a degree of inter-subjective understanding that needs to exist in order for the researcher to fully understand what is occurring—what the participant is doing and why? This in many respects is a methodological discussion that resides beyond the remit of this paper; however, there are practical considerations that must be brought to bear to inform the design of such systems and it is important that these are fully appreciated. Labelling is a key issue, understanding and quantifying, even restricting labelling can be an issue. Providing the correct tools to support labelling and the right input/output modality can be tricky, and while it may be the case that a lot of data may prove useful for understanding some contexts, controlling and understanding the rate and amount of reporting that participants do or can do is a feature that must be appreciated when the system is designed. Obviously, in some contexts, the system may only allow reports based on limited choice, but in others, a full response may be required.

NFC stickers need to be attached to each place or on the body to tag places or body parts. Recording user feedback in unconstrained settings is more challenging than in controlled environments since data collection is subject to noisier environment, and also, we must remember that we are developing tools for people to use while they are mobile, and the movement of the user can be a real issue and can impact upon the quality and the frequency of data collection and understanding the data.

There are multiple challenges when dealing with speech. These challenges range from the clarity of spoken words, different language/slang use and problems relating to background ambient noises that make speech hard for humans to

hear, such as wind noise. However, tools might be employed that post-process such audio in order to make it clearer, or that offer a degree of auto-label correction.

Also, there were some practical issues related to Volume Buttons methods, some users noted that they will not be able to adjust the volume if they listen to a media file if they were to utilise the buttons as a self-report tool. For example, user 7 said:

...always listen to music and adjust the volume constantly if walking around noisy places, alternative methods should be chosen by those who need to use the Volume-Buttons while collecting data

Further challenges in regard to speech relate to not being able to speak in circumstances where speaking might not be socially acceptable, and in circumstances where the focus of the study is on sensor data relating specifically to gathering background noise, such as occurs in environmental soundscape recordings. *Privacy* may be a core concern of the participant, and in some contexts, they may want to switch off the system if there was something that they did not want to report. It must be remembered that this in itself is data and in many respects needs to be reflected upon when mobile self-reporting systems are developed. How does one design for privacy, and how do we know if users are truthful? How do we as researchers and designers develop systems that engender truthfulness and trust; are there ways that we might mask the user identity?

7 Conclusions and future work

This paper proposes three new self-report methods which are useful for examining and discussing a series of contemporary self-reporting techniques for citizen science on-the-go, including Volume Buttons based responses and *NFC-on-Wall* and *NFC-on-Body* systems. To compare and understand the performance of these techniques, a user study was carried out to collect self-reported labels. Both the user labelling rates and user preferences are recorded for each method. Our data analysis showed, that while pressing screen buttons and screen touch allowed for higher labelling rates, Volume Buttons proved to be valuable when users are engaged in other activities which made using a mobile touch screen difficult, e.g. walking. We found that *NFC* labelling was also an effective technique when used in the context of self-reporting and place-tagging. However, using an *NFC*-based technique requires *NFC* tags to function in respect to a given scenario (e.g. tags need to be placed on the body or on street furniture prior to recording). The higher labelling rate of the screen use allows timely and constant data feed from users. In particular, it is useful when we need fast updates of the changes and

patterns in the user reporting. In addition, the screen buttons are also useful for optional user data collection, such as text boxes for data entry. On the other hand, the other techniques are only suitable for forced data collection, where, the user input should be specified in advance.

The new proposed methods can be utilised in many citizen science applications and data collection activities, where the user is on-the-move and submitting data. Future developments will focus on examining the potential of these new approaches in respect to real-time data collection in different application scenarios.

Also, we will consider adding vibration and sound effect as forms of feedback to users to confirm the recording of their input (e.g. the mobile phone will vibrate (with specific vibration pattern) when the Volume-Button is pressed. These types of instant feedback will allow users to change the way they use of their mobile phones in case of accidental input (e.g. Volume Buttons are pressed by mistakes inside a pocket).

By introducing new self-report and labelling techniques, we allow more natural and timely interaction that makes collecting citizen science data with mobile phone easier and less time consuming. This in returns will encourage more people to taking parts in scientific projects while moving and without stopping them from carrying out everyday tasks.

Acknowledgements Dr. Chamberlain's input into this work was supported by the Engineering and Physical Sciences Research Council [grant number EP/L019981/1] Fusing Semantic and Audio Technologies for Intelligent Music Production, and Consumption [grant number EP/M000877/1] Living with Digital Ubiquity.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alajmi N, Kanjo E (2013) Shopmobia: an emotion-based shop rating system. *Affective Computing and Intelligent Interaction (ACII)*, Humaine Association Conference on. IEEE
- Al-Barrak L, Kanjo E, Younis EMG (2017) NeuroPlace: categorizing urban places according to mental states. *Plos One J* 12: e0183890
- AlZuhair MS, Najjar AB, Kanjo E (2014) "NFC based applications for visually impaired people - a review". *IEEE International Conference on Multimedia and Expo Workshops*, p1–6
- Bradley MM, Lang PJ (1994) Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry* 25(1):49–59
- Chamberlain A, Crabtree A, Rodden T, Jones M, Rogers Y (2012) Research in the wild: understanding 'in the wild' approaches to design and development. *Conference on Designing Interactive Systems 2012 ACM DIS 2012*: 795–796
- Crabtree A, Chamberlain A, Grinter R, Jones M, Rodden T, Rogers Y (2013) Special issue on "The Turn to the Wild" with authored introduction. *ACM Transactions on Computer-Human Interaction - ToCHI* 20(3):13:1–13:4
- Department of Water Resources (2012) Analysis of the Department of Water Resources Volunteer Climate Cooperator Network (memorandum report). California Natural Resources Agency, California Retrieved from http://www.water.ca.gov/climatechange/docs/VccnAnalysis_20121206.pdf. Accessed May 2017
- Desmet P (2003) Measuring emotion: development and application of an instrument to measure emotional responses to products. *Funology*. Springer, Netherlands, pp 111–123
- Desmet P, Overbeeke K, Tax S (2001) Designing products with added emotional value: development and application of an approach for research through design. *Des J* 4(1):32–47
- Fuentes C, Rodríguez I, Herskovic V (2015) EmoBall: a study on a tangible interface to self-report emotional information considering digital competences. *Ambient Intelligence for Health*. Springer, Cham, p189–200
- Gaggioli A, Pioggia G, Tartarisco G, Baldus G, Corda D, Cipresso P, Riva G (2013) A mobile data collection platform for mental health research. *Pers Ubiquit Comput* 17(2):241–251
- Hancke GP, Hancke GP Jr (2012) The role of advanced sensing in smart cities. *Sensors* 13(1):393–425
- Harada S, et al. (2008) VoiceLabel: using speech to label mobile sensor data. In: *Proceedings of the 10th international conference on Multimodal interfaces*. ACM, p 69–76
- Herodotou C, Villasclaras-Fernández E, Sharples M (2014) The design and evaluation of a sensor-based mobile application for citizen inquiry science investigations. In: *European Conference on Technology Enhanced Learning*. Springer, Cham, p 434–439
- Kanjo E (2010) Noiseply: a real-time mobile phone platform for urban noise monitoring and mapping. *Mobile Networks and Applications* 15(4):562–574
- Kanjo E, Bacon J, Roberts D, Landshoff P (2009) MobSens: making smart phones smarter. *IEEE Pervasive Computing* 8(4):50–57
- Kanjo E, Steve B, Mark P, Alan C, Woodgate D (2008) MobGeoSen: facilitating personal geosensor data collection and visualization using mobile phones. *Pers Ubiquit Comput* 12(8): 599–607
- Kanjo E, Al-Husain L, Chamberlain A (2015) Emotions in context: examining pervasive affective sensing systems, applications, and analyses. *Pers Ubiquit Comput* 19(7):1197–1212
- Kanjo E, Younis EM, Sherkat N (2018) Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach. *Information Fusion* 40: 18–31
- Kanjo E, Ang K, Siang C (2017b) NotiMind: utilizing responses to smart phone notifications as affective sensors. *IEEE Access* 5, 22023–22035
- Kong X, Yang Y (2009) Measuring emotions in interactive contexts. In: *2009 IEEE 10th International Conference on Computer-Aided Industrial Design & Conceptual Design*, IEEE, p 526–530
- Kwan V, Hagen G, Noel M, Dobson K, Yeates K (2017) Healthcare at your fingertips: the professional ethics of smartphone health-monitoring applications. *Ethics Behav* 27:615–631
- Lottridge D, Chignell M (2009) Emotrace. Tracing emotions through human-system interaction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 53. No. 19. SAGE Publications
- MacKerron G (2017) Mappiness.org.uk. Accessed May 2017
- Mannini A, Sabatini AM (2010) Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* 10(2):1154–1175

26. Mody RN, Willis KS, Kerstein R (2009) WiMo: location-based emotion tagging. Proceedings of the 8th international Conference on Mobile and Ubiquitous Multimedia. ACM
27. Moran S, Pantidi N, Rodden T, Chamberlain A, Griffiths C, Zilli D, Merrett G, Rogers A (2014) Listening to the forest and its curators: lessons learnt from a bioacoustic smartphone application deployment, ACM CHI 2014, Conference on Human Factors in Computing Systems, April 26th - May 1st, in Toronto, Canada. ACM
28. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, Deleeuw W (2010) Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *J Med Internet Res* 12(2):e10
29. Nagel F et al (2007) EMuJoy: software for continuous measurement of perceived emotions in music. *Behav Res Methods* 39(2): 283–290
30. Nakagaki K, et al. (2016) Hydromorph: shape changing water membrane for display and interaction. Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction. ACM
31. Niforatos E, Karapanos E (2015) EmoSnaps: a mobile application for emotion recall from facial expressions. *Pers Ubiquit Comput* 19(2):425–444
32. Osmani V, Balasubramaniam S, Botvich D (2008) Human activity recognition in pervasive health-care: supporting efficient remote collaboration. *J Netw Comput Appl* 31(4):628–655
33. Posner J, Russell JA, Peterson BS (2005) The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev Psychopathol* 17(03): 715–734
34. Read, Robin, and Tony Belpaeme. (2013), Using the affect button to measure affect in child and adult-robot interaction. Proceedings of the 8th ACM/IEEE international conference on human-robot interaction. IEEE Press
35. Reid SC, Kauer SD, Dudgeon P, Sancu LA, Shrier LA, Patton GC (2009) A mobile phone program to track young people's experiences of mood, stress and coping. *Soc Psychiatry Psychiatr Epidemiol* 44(6):501–507
36. Renninger KA, Hidi S (2011) Revisiting the conceptualization, measurement, and generation of interest. *Educ Psychol* 46(3): 168–184
37. Rodriguez I, et al. (2016) B-ePain: a wearable interface to self-report pain and emotions. Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. ACM
38. Ruiz J, Li Y, Lank E (2011) User-defined motion gestures for mobile interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI). ACM, New York, NY, USA, 197–206. DOI=<https://doi.org/10.1145/1978942>
39. Samadani A-A et al (2013) Perception and generation of affective hand movements. *Int J Soc Robot* 5(1):35–51
40. Simor FW et al (2016) Usability evaluation methods for gesture-based games: a systematic review. *JMIR Serious Games* 4:2
41. Truong KN, Shihpar T, Wigdor DJ (2014, April) Slide to X: unlocking the potential of smartphone unlocking. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM p 3635–3644
42. West SE, Pateman RM (2016) Recruiting and retaining participants in citizen science: what can be learned from the volunteering literature? *Citizen Science: Theory And Practice*, p.15
43. Worthington JP, Silvertown J, Cook L, Cameron R, Dodd M, Greenwood RM, McConway K, Skelton P (2012) Evolution MegaLab: a case study in citizen science methods. *Methods Ecol Evol* 3:303–330
44. Zhang X, Pina LR, Fogarty J (2016) Examining unlock journaling with diaries and reminders for in situ self-report in health and wellness, Proceedings of the 2016 CHI conference on human factors in computing systems, Santa Clara, USA

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.