The value and acceptance of citizen science to promote transit accessibility

Aaron Steinfeld\textsuperscript{a,}\textsuperscript{*}, Rafae Dar Aziz\textsuperscript{a}, Lauren Von Dehnen\textsuperscript{a}, Sun Young Park\textsuperscript{a}, Jordana L. Maisel\textsuperscript{b} and Edward Steinfeld\textsuperscript{b}
\textsuperscript{a}Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA
\textsuperscript{b}IDEA Center, School of Architecture and Planning, SUNY at Buffalo, Buffalo, NY, USA

\textbf{Abstract.} We propose that citizen science methods can engage riders with disabilities and others in improving public transportation accessibility by documenting and assessing problems and good solutions throughout the system. This will empower riders, resulting in a greater understanding of the transportation system, and improve the feedback loop between rider and provider. This paper includes findings on how riders prefer to report such observations through an experiment designed to compare the modalities of the Notes (text, audio) and Media (none, photo, video). The results from two use: groups, those without disabilities and those who use wheeled mobility devices, suggest that text with photo should be supported and that use of video may not have additional value to end users. The generally positive responses suggest that end users are open to participating in such communities and that feedback is important.

Keywords: Citizen science, public transportation, accessibility, transit riders, disabilities

1. Introduction

Reports of barriers to accessibility in existing transit systems indicate that the lack of implementation of known best practices is a more serious problem than the lack of technology solutions to removing barriers. Improvements to existing technologies are certainly needed, but, on the short term, the slow adoption rate of best practices has a more significant impact on people with disabilities. Adoption and maintenance of best practices are hindered by the sheer size and complexity of transit systems and the limited effort agencies can dedicate to accessibility problems given their consistent funding challenges.

Likewise, consumers regularly report little to no feedback when submitting reports of problems and in many cases it is not clear to the customer service agent where to route the problem. Positive feedback and suggestions for accessibility improvements faces similar barriers. We believe that two-way feedback between riders and providers is the key to supporting best practices and propose that technology can be used to streamline this interaction. In particular, we see real promise in the use of citizen science – the application of rich media evidence to civic advocacy [17].

Accessible public transportation is critically important. It allows individuals with disabilities, especially those with severe disabilities, to have independent access to works sites, educational programs, health facilities, and social and recreational activities. In a mobile culture, full social participation hinges on accessibility of transportation systems. However, the current state of accessible public transportation is a barrier to social participation and, particularly, employment. More than half a million people with disabilities in the United States cannot leave their homes because of transportation difficulties [6]. One-third of people with disabilities have inadequate access to transportation [14]. Consequently, four times as many people with disabilities

\* Address for correspondence: Aaron Steinfeld, Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA. Tel.: +1 412 268 6346; Fax: +1 412 268 6436; E-mail: steinfeld@cmu.edu.

ISSN 1055-4181/10/$27.50 © 2010 – IOS Press and the authors. All rights reserved
as people with no disabilities lack suitable transportation options to meet their daily mobility needs [13].

Such difficulty leads to numerous consequences. According to one study, 46% of people with disabilities, compared to 23% of people without disabilities, reported feeling isolated from their communities [15]. Individuals with disabilities were five times more likely to report dissatisfaction with their lives than were their non-disabled counterparts, and a majority of those surveyed said that lack of a full social life was a reason for this dissatisfaction. For example, with disabilities were about half as likely to have heard live music, gone to a movie, or attended a sporting event or concert over a one-year period [9,15]. Inadequate transportation limits access to these activities for individuals with disabilities. People with disabilities, both in urban and rural areas, frequently cite a lack of local transportation as hindering their ability to find employment. Lack of transportation (29%) was only second to a lack of appropriate jobs being available (53%), as the most frequently cited reason for being discouraged from looking for work [12].

1.1. System wide assessment

Transit providers and consumer advocates in many locations are working effectively to develop good solutions to many common problems but these best practices are not being adopted in other locations. For example, the National Council on Disability published a report entitled “The Current State of Transportation for People with Disabilities” [13]. Of the 11 different problems with fixed route systems identified in that report, seven of them were related to service delivery and policies as opposed to vehicle and building technology issues: reliability of stop announcements, maintenance problems with lifts, compliance with lift operation policies, planning accessibility to stations (i.e. accessibility beyond key stations), wheelchair securement policies, elevator maintenance problems, and continued use of poor accessibility solutions like mini-high platforms. Methods are needed to identify problems as they arise, assess the impact of those problems on people with disabilities, and bring this information to the attention of service providers and policy makers as part of operations and planning.

Systematic research at the local level, implemented as part of continuous quality improvement, can identify problems in a timelier manner and collect rich information on systems, policies, and practices that work. Knowledge re-use comes in the form of disseminated best practice solutions at a national and international level (e.g., Eastern Seals Project ACTION; Access Exchange International). However, most transit agencies do not have the resources to initiate systematic research using conventional approaches nor do such approaches adequately address their needs and the needs of riders.

Surveys are a common method to assess a metropolitan area but they have limitations. For example, survey methods make it difficult for respondents to define the issues from their own perspective [21], data collection is time consuming and protracted, and there are numerous barriers to independent analysis and advocacy by end users. Options like guided tours, in which end users identify problems and solutions to researchers in naturalistic settings, are extremely valuable because the exposure to real settings and products is very effective in prompting detailed responses from the end user [8, 10]. Interviews and focus groups allow individuals to define issues from their own perspective [1]. However, due to high cost and logistics these methods are often constrained by small sample size problems and are difficult to maintain on a continuous basis. They are more suitable for periodic assessments, as a prelude to other research, and to elicit end user input.

1.2. Enabling the end user to make a difference with technology

A key limitation of many approaches is that the process of data collection is owned and operated by researchers. While this reduces bias and noise, it does not empower end users nor support their immediate needs. Finally, these methods often do not offer service providers evidence in a form that is directly applicable to their needs. For example, a survey of riders with disabilities or focus groups may identify that lift breakdowns are a frequent problem but these methods suggest some lift models and bus routes are more prone to problems than others. In fact, by the time data is obtained, the problem may already be fixed.

In the transit context, the ideal scenario for accumulating information that can be applied rapidly in practice and policy development is to combine the immediate, personalized, and localized information obtained in guided tours with techniques that are more cost effective than a large sample survey. The widespread availability of personal electronics provides an opportunity to implement this scenario using the end users’ own cameras and mobile phones. For example, Fig. 1 was taken by a member of the team while en route to a meeting with a transit agency. These ubiquitous con-
Fig. 1. An example mobile phone image of a maintenance problem.

Consumer products are often capable of multimedia recording. Thus, they provide an opportunity for end users to collect data that can be used as a research database. A good participatory action design [7] approach is to go further and enable real time access to the data, thereby enabling continuous interaction between researchers and end users.

There are many examples that demonstrate the value of using this form of data collection. Major network news channels, accident reconstruction teams, courts, and law enforcement routinely utilize information obtained by citizens on camera phones and camcorders. The popularity of YouTube and similar sites demonstrates the potential for using these methods as a means of civic engagement and public discourse. For example, YouTube videos of transit bus features are regularly produced by amateurs and popular enough to accumulate thousands of views each.

Paulos points to citizen science work in air quality and public parks [17]. Multimedia is much more powerful than dry statistical data. For example, a local bicycling advocacy group recorded GPS and air quality levels while riding a circuit of downtown Pittsburgh and documented road sections of the city with heavy pollution and packaged it in a form readily understood by decision makers. Pittsburgh and other municipalities now have iPhone applications for reporting problems directly to city offices (e.g., iBurgh). Similar efforts have occurred in documenting general accessibility barriers at a citywide scale using MMS (Fig. 2).

ParkScan.org in San Francisco is a model for the application of citizen science to the improvement of public parks. This domain is a good metaphor for citizen science in public transportation given the similar issues with physical and organizational complexity. In 2007 alone, ParkScan had 425 registered users, 1,531 observations, and 68% of the issues identified by end users were addressed by the City [16]. Rich multimedia evidence provides extremely persuasive evidence for end users to promote change in their communities and in policies at a national level. Likewise, the use of such methods of data collection is very compatible with widely adopted management tools. For example, surveillance video is used widely to maintain security and loss control and facilities managers use digital photography and video to document accidents and other events.

Given this potential, we are working towards large-scale deployment and evaluation of a citizen science system for public transit. In parallel to software development, the team is examining some key end user questions through experimental methods. The first study, reported here, is focused on desired media reporting methods and general issues related to willingness to file reports.

2. Experiment method

2.1. Participants

During recruitment, the participants were informed they would be completing surveys and documenting items in the laboratory. All participants were paid volunteers and fully consented. There were two participant groups: wheeled mobility device users (WMD) and people without disabilities (Control). Each group had 12 participants. While the latter group may seem unnecessary, we believe that the best way to deploy a citizen science system is to appeal to the full spectrum of potential end users. This universal design approach means that riders without disabilities can use the same system to report problems that matter to them (e.g., vandalism, broken glass, etc). As with ParkScan.org, the system would be a key tool for providing general customer service and developing a collaborative relationship with all users.

Participants were recruited from local universities and the general public using a local recruitment website, contacts in local organizations, and community email lists. Participants were required to be 18 years of age or older, fluent in English, and not affiliated with the project. The inclusion criteria for the WMD group were anyone who met the general requirements and used either a wheelchair or scooter.
2.2. Design and materials

The study was a 2 × 3 design. The first effect was the modality corresponding to the Notes about the problem. These were either in Audio or Text form and corresponded to either a voicemail or email/web documentation of the problem. Audio reports were recorded on a digital voice recorder while text reports were typed into a text editor on a computer. Since it was not possible to have a full array of accessible computer modification devices, participants who could not type narrated to an experimenter who typed their entry verbatim under the participant’s supervision. Editing of text was allowed.

The second effect was the Media used to support the notes. These were None, Photo, and Video. The middle was documented using a consumer grade 10-megapixel Canon digital camera while the latter used a 720p Flip digital video recorder. WMD participants who could not hold the camera or video recorder in a stable manner were given the option of using a flexible stalk mount (The Smart Pod) which was attached to their wheeled device using a clamp.

The permutations of the 2 × 3 design led to six conditions. To prevent bias, each condition was tested with one of six different paired information and spatial problem stimuli. Half of each participant group used a different Set of stimuli. Set “A” consisted of surface conditions related to bus shelter upkeep (e.g., accumulated snow, accumulated ice, trash, dirt pile, broken glass, etc) and bus information conditions related to the bus stop sign (e.g., no labels, vandalized, worn out labels, labels on wrong side of sign, etc). Set “B” consisted of spatial conditions related to movement in and around the shelter (e.g., full bench, narrow passage, permanent blockages, narrow sidewalk, etc) and bus information conditions related to the bus schedule sign inside the shelter (e.g., water damage, graffiti, broken glass, stickers, etc).

All conditions and stimuli were run within-participants in a counterbalanced order. Counterbalancing also included the order in which participants reported on information and spatial stimuli. Details on the counterbalancing and statistical checks for bias can be found elsewhere [18].

2.3. Storyline and procedure

A simulated bus shelter was created inside the laboratory for this experiment (Fig 3). This included an advertising poster, a bench, a tempered glass panel on the upstream side of the shelter, a place to mount route information signs (inside, above the end of bench), and a bus stop sign with interchangeable route label panels.

Control participants were told that they were approaching the bus shelter to take a bus with a friend who uses a wheelchair or scooter. WMD participants were told they were approaching the bus shelter to take a bus. Both groups were told this was an unfamiliar stop. In Set A, the WMD user intended to park under the bus shelter and needed to confirm that this was the correct bus stop for their desired bus. In Set B, the WMD user wanted to maneuver around the shelter to board the bus and needed to see if they had missed the bus based on the schedule. All participants were asked to stay on the
ratings for (a) experiment condition preference and (b) regarding the merit of each condition for documenting broken lifts/ramps on vehicles and problems with not getting on a bus due to overcrowding. These two problems were examined with the post-test survey since they would be difficult to recreate using a full-scale simulation.

Within the study, each participant completed an identical post-condition survey (Table 1) after each condition. Responses to each question within each post-condition survey category (Ease of Use, Usefulness, and Social Comfort) were flipped to have the same positive/negative direction, with higher as better, and averaged as a group. The three resulting Index groups all surpassed the 0.7 reliability acceptance threshold used in the literature [18]. Part of this survey was based on a previously validated survey for desktop computing [19] but new questions were added to collect additional data on perceived usefulness and on perceived social discomfort. The latter were developed during review of other studies on perceived vulnerability (e.g., [5,11]). Participants were not shown the index labels.

3. Results

3.1. Pre-test survey

The final block of questions in this survey focused on consideration of technology for assistance in daily tasks. Participants were asked to rate facts like ease of use, cost, attractiveness, visibility to others, impact on privacy, and safety on a scale of 1-10 (not important at all to extremely important). There were significant, strong correlations between visibility and privacy (0.526, p = 0.008) and visibility and safety (0.585, p = 0.003). These trends reinforce the notion that social comfort is an important consideration.

3.2. Post-condition survey

The effects of Notes and Media were analyzed in a full factorial ANOVA along with Group (Control, WMD) on Ease of Use. While Audio reports had a slightly higher mean than Text notes for WMD participants (6.1 vs. 5.3), the Control participants had the inverse relationship (5.7 vs. 6.4). This resulted in a significant interaction between Group and Notes for Ease of Use ($F = 5.9, p = 0.017$). However, the Tukey HSD post-hoc revealed no differences among the four com-
Table 1
Post-condition survey (7-point scale from strongly disagree to strongly agree)

<table>
<thead>
<tr>
<th>Ease of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Learning to use this method was easy</td>
</tr>
<tr>
<td>2. Becoming skillful with this method was easy</td>
</tr>
<tr>
<td>3. I had no problem physically using this method</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Usefulness, part 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Using this method would improve my performance in reporting observations</td>
</tr>
<tr>
<td>5. Using this method for reporting observations would increase my productivity</td>
</tr>
<tr>
<td>6. I feel this method is too slow for everyday use (r)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Social Comfort</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. I felt uncomfortable using this method when people were around in public (r)</td>
</tr>
<tr>
<td>8. When I use this method, I feel like other people are looking at me (r)</td>
</tr>
<tr>
<td>9. Using this method in front of strangers embarrasses me (r)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Usefulness, part 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. I like the idea of using this method</td>
</tr>
<tr>
<td>11. I would have done as good a job without using this method (r)</td>
</tr>
<tr>
<td>12. Carrying items to be done this method on daily trip is such a hassle to me (r)</td>
</tr>
</tbody>
</table>

(r) = scale reversed for index averages and analysis.

Fig. 4. Media results for Social Comfort (± 1 std error).

Fig. 5. Group results for Social Comfort (± 1 std error).

There were no other significant differences for Ease of Use.

Social Comfort ($F = 20.4, p < 0.0001$) was higher for Text notes when compared to Audio notes (Fig. 4). There was also a significant difference in Social Comfort for Media type ($F = 4.2, p = 0.018$) with post-hoc results showing significant differences between None and Video. There were also differences in Group ($F = 21.1, p < 0.0001$) and the interaction between Notes and Group ($F = 5.2, p = 0.02$). The post-hoc revealed that the Control group reported significantly reduced Social Comfort for Audio as compared to the other three combinations (Fig. 5).

Usefulness mimicked some of the trends seen in other metrics but did not result in significant differences at the $p < 0.05$ level. Note type was not significant.

There was a slight significant difference for Media on question 12 (hassle) with Video showing a full point higher hassle than None in the post-hoc ($F = 24.1, p = 0.047$).

3.3. Post-test survey

There were no significant differences in Lifts/Ramps in Notes ($F = 26.7, p = 0.004$) and Media ($F = 24.5, p = 0.023$) but no Group or interaction effects. Post-hoc analysis showed Photo was significantly higher than Video.

Ratings for documenting problems with broken lifts/ramps and getting on a bus due to overcrowding...
were generally high for the photo and video media types (Fig. 7). While there was a visible trend for broken lifts/ramps for Media, only Group showed a significant difference with WMD participants providing lower ratings than Control participants (4.8 vs. 5.8, $F = 10.8$, $p = 0.001$). For overcrowding, there was a significant Media effect ($F = 5.7$, $p = 0.004$) with None being significantly lower than the other two options. There was also a significant difference for Group ($F = 4.9$, $p = 0.029$) with lower ratings from WMD participants (4.7 vs. 5.3). The lower responses from the WMD group is not surprising in that these events are fast paced, thus making it hard for WMD participants with limited dexterity to capture photo or video evidence. In fact, a subsequent t-test comparing Control and WMD responses to these two questions on only the

None conditions revealed no significant differences.

Only two Control participants had filed a complaint about a transit problem while 7 of 12 had in the WMD group. Three of the WMD group had filed over 20 complaints, mostly over the phone. The Control participants did not receive feedback and the WMD participants who did, reported low rates. Interviews by the team [20] with the local transit agency indicated that the customer service representatives rarely have time to return calls for issues other than lost and found. Only two participants reported timely resolution of complaints and those rates were around 50%. It is unclear whether the reported problems were issues within the jurisdiction of the transit agency. For example, damage to the sidewalk outside of the immediate bus stop area is the responsibility of the local municipality.

There was concern that comfort with uploading photos or videos to websites would incur an unmanageable negative bias on the data. The post-test survey included a question to this effect, “Have you uploaded photos or videos before? If yes, are you comfortable doing upload photos or videos?” One Control participant and three WMD participants responded they had not uploaded photos or videos. All other participants indicated comfort with the process except for one WMD participant. While only two thirds of the WMD participants reported comfort with such activity, the quantity seems within reason.

There was additional supporting data for the photo modality. The post-test survey included the Yes/No question: “When explaining an experience, I usually prefer to include visual information, instead of just text.” Over 70% of the participants responded positively. Also, the survey ended with a free form section where participants could provide comments and suggestions. Some of the remarks aligned with the other results:

“... For some reason I felt it was more on the spot with audio.”

“I liked the camera and text approach the best. But I hate text messaging on the cell phone because my vision is not the best and texting is laborious. For me I would like to take a picture then go back to my office or home and write the complaint out. This gives me time to calm down after I have been upset by some incident.”

4. Discussion

4.1. Does the data suggest citizen science has potential appeal in the public transit realm?

The consistent, favorable results for the text with photo condition suggest this combination has merit.
This is supported by the large quantities of MMS-based reports collected during citizen science accessibility efforts in Barcelona and elsewhere. Unlike these prior implementations, the team suspects that without feedback, problem reporting systems can easily appear to be "black holes" where complaints go in and nothing happens. This is apparent in the data presented here where participants reported limited feedback and resolution for reported problems. Our colleagues on the iBurgh team have also anecdotally reported that lack of feedback in the early versions of their application was a problem. This black hole problem can lead to low perceived benefit for engaging with the local transit agency.

Parallel interviews and interaction concept testing by the team [20] has revealed that riders rarely encounter infrastructure problems that meet the perceived cost-benefit threshold for reporting. As a result, the team’s view is that observation reporting should be part of a larger system that includes valuable, frequent information (e.g., arrival times, vehicle fullness, and dynamic route changes), thereby streamlining the infrequent desire to report observations and lowering the cost. This approach gives riders many different reasons to remain engaged with the system and supports pre-loading of important real-time details (e.g., current location, route, etc).

4.2. How do acceptability factors like added hassle and perceived social comfort vary by reporting method?

Conditions with text were viewed as having considerably higher social comfort than audio conditions. Likewise, video scored slightly worse than the other media options and was reported to be more of a hassle than the no media approach. These findings align with the high incidence of text messaging in public and the social implications of video recording around strangers. There may also be a perceived risk to using an apparently more expensive device — even though the camera and video recorder used in this study were of comparable value and far cheaper than some smart phones.

There were no apparent differences in perceived hassle or social comfort when comparing no media to photo, thus implying that documenting a problem with some types of rich multimedia is similar to traditional methods for such metrics.

4.3. Next steps

Future analysis will examine the relative quality of the documentation collected and provide input on the ability of end users to collect useful multimedia data. There are also plans to examine the value of audio for riders who are blind, have low vision, or poor manual dexterity.

As mentioned, the work described here is a portion of a larger effort towards developing, implementing, and evaluating a software infrastructure for this purpose. The project will culminate in a large-scale field test over a whole metro area.

Acknowledgments

The authors would like to thank colleagues in the local Pittsburgh area who facilitated participant recruitment. Sun Young Park joined the Department of Informatics, University of California, Irvine after contributing to this work.

The Rehabilitation Engineering Research Center on Accessible Public Transportation (RERC) is funded by grant number H133E080019 from the United States Department of Education through the National Institute on Disability and Rehabilitation Research.

References


Special Issue: Quality of Life Technology:
Intelligent Systems for Better Living
Guest Editor: Asim Smailagic

ASSOCIATION FOR THE ADVANCEMENT OF ASSISTIVE TECHNOLOGY IN EUROPE
TECHNOLOGY AND DISABILITY
Volume 22, Numbers 1,2, 2010

Abstracted/Indexed in: Cumulative Index to Nursing & Allied Health Literature, CINAHL database, EBSCO's database, Educational Research Abstracts online (ERA), EMBASE/Excerpta Medica, EMCare, Ergonomics Abstracts, Exceptional Child Education Resources (ECER), MasterFILE, REHABDATA, SCOPUS

Special issue: Quality of Life Technology: Intelligent Systems for Better Living
Guest Editor: Asim Smallagic

CONTENTS

Guest Editorial 1

J.T. Matthews, S.R. Beach, J. Downs, W.B. de Bruin, L.P. Mecca and R. Schulz
Preferences and concerns for quality of life technology among older adults and persons with disabilities: National survey results 5

L. Zhao, X. Wang and G. Sukthankar
Recognizing household activities from human motion data using active learning and feature selection 17

M. Beetz, M. Tenorth, D. Jain and J. Bandouch
Towards automated models of activities of daily life 27

B. French, D.P. Siewiorek, A. Smallagic and T. Kamarck
Lessons learned designing multi-modal Ecological Momentary Assessment tools 41

H.-Y. Liu, R. Cooper, R. Cooper, A. Smallagic, D. Siewiorek, D. Ding and F.-C. Chuang
Seating virtual coach: A smart reminder for power seat function usage 53

K. Bunning, B. Heath and A. Minnion
Interaction between teachers and students with intellectual disability during computer-based activities: The role of human mediation 61

A. Steinfeld, R.D. Aziz, L. Von Dehisen, S.Y. Park, J.I. Maisel and E. Steinfeld
The value and acceptance of citizen science to promote transit accessibility 73

M. Sciogaj and R.A. Behr
Lessons for the United States from countries adapting to the consequences of aging populations 83

VISIT OUR WEBSITE ON WWW.IOSPRESS.NL