Usability and usefulness of internet mapping platforms in participatory spatial planning

Michal Rzeszewski*, Jacek Kotus
Faculty of Geographical and Geological Sciences, Adam Mickiewicz University, Poznan, Poland

ARTICLE INFO

Keywords:
PPGIS
Participatory mapping
Human computer interaction
Urban planning
Eye tracking

ABSTRACT

Internet mapping platforms are widely used in participatory planning. They carry the promise of accessibility, a wider audience, and more precise spatial data. However, they are critiqued for issues of trust, digital exclusion, and low usability. While the first two problems can be dependent on the local context, the last one is connected with the introduction of map editing tools. They change browser experience and require accessing spatial knowledge, which is a cognitively demanding task, especially for older people. This can negatively influence both quality and quantity of the data being produced. In this paper, we ask questions about the perception of online forms of participation, the accuracy and reliability of the data being collected in participatory mapping, differences between participants age groups, and design considerations for creating successful and inclusive participatory mapping tools? We conducted usability testing experiment with 30 participants. They were presented with various map editing tasks that were to be completed using a mock internet participatory mapping platform that was designed to mimic real-life applications. The sessions were recorded using eye-tracking equipment.

We have observed a considerable enthusiasm for using and interacting with web mapping tools, but this was not reflected in the quality of the data which was very low overall. Known map features were added in wrong places even in simpler tasks. Although there were considerable differences between the oldest and youngest people in the map editing behavior, the quality and quantity of the data were similar. We conclude that mapping platforms can be useful for participatory planning and age of the participants is not a barrier for their inclusion when the interface is properly designed. But in the same time care is needed to provide robust quality control. We also propose a research agenda composed from three avenues of research: data quality, usability and user experience and utility of the map.

1. Introduction

In the field of spatial planning, the issue of public participation is increasingly important. In many countries ensuring a way for people to take active part in shaping space of their place of living is required by law. But even when it is not required, participation is often initiated by grassroots movements and is welcomed by citizens (Gstach, Sinning, Mönchgesang, Kotus, & Sowada, 2016; Kotus & Sowada, 2017). This trend crystallizes in different forms of public engagement, which are realized on many spatial scales and are supported by various tools and techniques. The intersection of social processes and spatial technology that emerges from this process is termed as the field of interest of public participation geographic information systems (PPGIS) - which were aimed at empowering the less privileged groups in society and supporting public participation (Brown et al., 2014; Sieber, 2006). The PPGIS as a term was created in the 90s and was the result of the critique that the geographic information system (GIS) received from the social sciences (Curry, 1995; Sheppard, 1995). As Brown and Pullar (2012) noted, the definition of what exactly counts as PPGIS is nebulous at best. The first definition is attributed to Schroeder (1996) and states that PPGIS is “a variety of approaches to make GIS and other spatial decision-making tools available and accessible to all those with a stake in official decisions”. Later, many more definitions have emerged that focused on issues like the struggle for spatial justice (Aberley & Sieber, 2002), the emancipatory potential of grassroots movements and public participation (Sieber, 2006), or even just the combination of planning practices and GIS (Ramasubramanian, 2010). When PPGIS tools and principles are being applied to development and land-related planning, they can also be described as participatory mapping—a term that refers to the process of communicating spatial knowledge by using maps.
An extensive review of the PPGIS and participatory mapping applications is beyond the scope of this paper; interested readers are encouraged to seek more information within other publications (e.g., Brown, 2012a; Sieber, 2006). What is important for our line of inquiry is the observation that in recent years, an increasing number of these solutions are web-based (Brown & Kyttä, 2014; Kingston, Carver, Evans, & Turton, 2000). Applications range from urban planning (Bugs, Granell, Fonts, Huerta, & Painho, 2010; Meng & Malczewski, 2010), spatial decision support systems (Carver, Evans, Kingston, & Turton, 2001; Kessler, Rinner, & Raumbl, 2005; Rinner & Bird, 2009), mapping landscape values (Carver et al., 2009), assessments of park visitation (Norman & Pickering, 2017), strategic planning of wind farms (Simao, Densham, & Haklay, 2009), and other various mapping applications and geopatterns that are focused mainly on urban spaces (Brown & Kyttä, 2014; Czekpiewicz et al., 2016; Evans & Waters, 2007; Hasanzadeh, Broberg, & Kyttä, 2017). In all these instances, using web interfaces in PPGIS endeavors seems all but natural. It provides creators of mapping tools with a rich software environment with numerous advanced web frameworks. Many of them are also shared according to Free and Open Source Software principles—for example, Leaflet and OpenLayers, which plays well with the emancipatory aims of PPGIS by providing access to collaborative mechanisms of control.

While in practice this possibility is limited by the level of technical expertise of the collaborators, there are two important reasons why these applications are built and implemented in a web environment. One is to increase the level of participation by including additional people. The assumption here is that certain groups of people, who would not participate in traditional real-life meetings and workshops, would be willing to participate from the comfort and anonymity of their homes; additionally, they would also be drawn by the allure of the new technology. Online tools are less expensive, they enable research on a wider scale, and, most importantly, they allow sustainable implementations that can be scaled in time and space (Jankowski, Czekpiewicz, Miodkowski, Zwoliński, & Wójcicki, 2017). The second incentive for using online mapping tools is that they offer significant potential for data collection. This is not just the kind of data that is readily available through paper-based questionnaires, but information that can be enriched by a spatial component at the source. Participants can add a text comment or a photo to a specific area. They can mark a point or draw a polygon on a map or a satellite image. They can even discuss the issues with geotagged comments. The resulting database can be then used for spatial analysis and as a basis for making informed decision on the future of urban spaces. The process of creating a map can be in itself viewed as empowering for a local community.

However, there are serious limitations in using PPGIS mapping methods, which cast some doubt over potential benefits. Firstly, the inclusive potential of the online tools can be as well negated by exclusion issues. Any tool that is based on internet penetration is bound to raise serious questions about the presence of a digital divide and digital exclusion in the local community. While online tools do increase the overall number of people that participate (Czekpiewicz et al., 2016; Jankowski et al., 2017), it has also been shown that the demographic of these participants is highly skewed toward younger, more educated residents (Stern, Gudes, & Svoray, 2009). This problem can be especially pronounced in many places wherein the digital divide (Brabham, 2009; Czekpiewicz et al., 2016) is vividly visible and the feeling of collective control extremely small. This also highlights the issue of trust for online maps, both on the participant side (Skarlatidou, Haklay, & Cheng, 2011) and urban planner side (Brown, 2015). The second limitation is connected to the process of gathering spatial knowledge from local residents. Participants may not always be as familiar with a map-based interface; even if they use web-based maps like Google Maps or Bing Maps, they use them just for browsing and navigation and not to create spatial content. This poses a serious challenge to the usability design. Map editing functions, which include creating spatial object like points and polygons, are most likely unknown to participants; moreover, these introduce specialized GIS functions that complicate the conventional browser experience (Zhao & Coleman, 2007).

Q1: Can PPGIS mapping tools be successful in a post-socialist context or similar contexts marked by distrust and digital divide?
Q2: How do users perceive online forms of public participation?
Q3: What are the differences in using PPGIS mapping tools between various age groups?
Q4: What is the accuracy and reliability of data that is collected using PPGIS mapping platforms?
Q5: What are the design considerations for building inclusive web mapping platforms for public participation?

In this paper, we outline a research agenda that can be used to...
explore the general issues in the usability of PPGIS mapping practices and user experiences that shape these practices. We also place our research within the post-socialist narrative since we think that it has unique characteristics that set it apart in the context of participation in urban planning (Andrusz, Harloe, & Szelenyi, 2011) and these participatory initiatives first need to break through many existing layers of distrust. One of the possible sources of distrust are the maps themselves. Maps in Poland were, for many years, a tool of power that was kept secret. Moreover, readily available maps were full of errors, deliberately introduced by mapmakers. That being said, all of the limitations mentioned above are present and often more vividly, in indigenous and remote area communities of developed countries (Corbett, 2013). The influence of the digital divide and issues of trust has been shown in numerous studies to be important in the adoption of participatory mapping solutions in those contexts (Smith, Ibáñez, & Herrera, 2017; Maloney-Krichmar, & Preece, 2004, pp. 445–456) principles—starting from paper prototyping to two tests with 5 users each. In each iteration, changes were made to assure the greatest usability of the system for a wide range of age groups. However, we placed some limits on the changes that were made—our survey of mapping platforms gave us a set of functions and user interface characteristics that were shared among almost all surveyed examples. We wanted to include these characteristics despite the fact that they were sometimes detrimental to the general usability. For example, during our test it became apparent that users were often confused or uncomfortable with a choice between various options of generating input content, for example adding points versus adding lines. While this could indicate that we should limit the number of options, we opted to include at least two in our interface because a majority of the surveyed platforms offer a similar choice. We decided on points and polygon features as originating from different ends of the complexity spectrum. The final web application was based upon Mapbox JS, Leaflet libraries, and PostGIS database to collect the data with PHP scripts for communication. We used Open Street Map as the base map, but a radio button was available for switching to Google satellite view. Every user was given a unique, randomly selected login and password. At every sub-page of the platform, both context and text help was provided.

2.2. Participants

We recruited 30 participants from the inhabitants of the Jezycy district in Poznan, Poland. For the recruitment procedure, we had posted an open call for participation and each volunteer was asked to fill a short personal information questionnaire. Only volunteers who fulfilled our requirements were selected. We required that they would have to have been living in the area for more than 5 years and that they would originate from various backgrounds and age groups to gather the widest range of possible behaviors and views. It was especially important for us to recruit older (60+) people, as they are commonly underrepresented in participation initiatives that use online tools (Gottwald, Laatikainen, & Kyttå, 2016). Due to the ethical considerations, we did not include minors in the study. In the final sample there were 15 men and 15 women. Of these, 10 of the people did not have previous experience with web maps and none participated in eye tracking usability testing or participatory mapping before. The education level of the sample was higher than average with 17 people having some kind of tertiary qualification—56.6% in comparison with the average of 20.7% (2016) for this region of Poland. There were four students, two unemployed, and six retired persons. Two people perceived their IT skills as poor, 10 as average, eight as good, and 10 as very good. In the analysis, we divided the users into four age groups: 18–24 (n = 6), 25–39 (n = 8), 40–59 (n = 9), and > 60 (n = 7). All participants were fully informed about the nature of our experiment and were told that they could leave at any point without consequences. One person left after the tasks were described and was replaced. Each person was compensated for their participation. No personal data was collected—both recordings and surveys were anonymized.

2.3. Experiment design

The task presented to the participants was separated into 4 distinct stages. In stages 1–3, there were two questions that could be answered by adding a point or polygon, with or without text comments in the popup, to a specific location on the map. In stage 4, only the polygon tool was available, with the possibility of adding a comment. All the questions, content, and tools available for users are described in Table 1.

The questions and stages were structured in such a way as to require increasing level of attention as well as to help to familiarize the users with the map and questionnaire interface. At Stage 1, we asked two very simple questions that we assumed were not mentally challenging to answer verbally; by observing the increased stress and mental load during the experiment, we could also observe the influence of forcing the participants to use map tools. Stage 2 consisted of a similarly simple task, but we also tested the level of spatial knowledge about the Jezycy district by asking to locate two very recognizable landmarks—one of them being an areal feature (Jezycy Market). Successes in Stages 1–2 were used as a proxy for measuring accuracy of the data. Stage 3 required users to fully utilize their spatial knowledge of Poznan. By asking questions that concerned the whole city area, we tried to encourage participants to fully engage in communicating their opinions and to observe various behaviors regarding navigating the map and using the editing tools. Finally, at Stage 4 there was only a single task that required participants to make comments on the spatial development of their neighborhood. We assumed that at this stage, users would be familiar with the map interface and the limited spatial scope of the task would reduce the level of mental effort in more stressed individuals. At each stage, there was a set of tools that could be used to mark and add features (Table 1). We allowed participants to make choice between marking places with polygons or points. In this way, we gained insight into the way people perceive different places—do they
differentiate between areas and points on a map and in real space? In Stage 4, we removed the option to use point symbols for making features to force users to use polygons instead. In most questions, they also had the option to add a text comment and to name the newly added feature. Together with a Concurrent Thinking Aloud protocol (see below), we hoped that this would allow us to better understand the motivations of the users, especially in Stages 3 and 4.

The measure of success was different in each stage (Table 2), but in general, the criteria were relaxed and slightly biased toward success—only one question was needed to be answered correctly and when some of the people added more than one point in Stages 1 and 2, we took into account the more accurate points.

### 2.4. Measures of performance and cognitive load

To assess participants' ability to perform spatial tasks, we counted the number of successfully finished tasks (see above). To describe their behavior, we recorded the time spent in each stage, time to add the first feature, the zoom level used to add each feature, the maximum zoom level used, the number of created points and polygons, the amount of text in the comments, and the spatial clustering of features in Stage 3. In every stage, apart from Stage 4, the initial zoom was set at 13. In Stage 4, we concentrated on the neighborhood area and the map scale was deliberately larger—therefore, we exclude this stage from the analysis of zoom levels. Sessions were recorded with stationary eye tracking equipment—Tobi X-60, which has a 60 Hz sampling rate and twin-eye measurement mode. We measured gaze coordinates, pupil dilation, and eye tracking (TTFF) and fixations for each of the 8 separate parts of the screen (map, left panel with task description and feature tools, logo, title, zoom controls, layer control, logout button, and help button). Eye tracking recorded were visualized with heatmaps that highlighted areas of the of stronger visual attention. Since heatmap visualization hides changes in the extent of the map due to zooming and panning, we have employed an additional measure to gauge visual attention clustering of the participants—Nearest Neighbor Index (NNI) (Levine & others, 2004)—in which lower values indicate more clustering. During sessions, participants were encouraged to use the Concurrent Thinking Aloud (CTA) protocol, which is widely used in usability testing (McDonald, Edwards, & Zhao, 2012), to help identify the problems they faced with tasks. In this method participants are encouraged to verbalize their thoughts during task executions, with only slight and unstructured encouragements from the person supervising the test (“Please, keep talking”). All statement were recorded together with screen image.

Although cognitive load can be approximated with pupil dilation (Kahneman & Beatty, 1966), there are many external factors that can influence this measure. Therefore, we measured cognitive load with a declarative method—by asking questions (see Appendix A) to separately gauge mental load (ML) and mental effort (ME), both of which are highly correlated components of cognitive load (CL) (Krell, 2015). Mental load is related to the complexity of the task and is defined as the amount of cognitive capacity required to process it. Mental effort is related to the subject, and it is the amount of cognitive capacity the participant invests while solving the task. ML was measured by questions Q1–Q6 on the CL Scale and ME by questions Q7–Q12. Scores were added together but some of the questions (Q2, Q5, Q6, Q8, Q10, and Q11) were “reversed,” which means that a value of “5” on the survey contributed to “1” in the overall sum. To further gauge problems that can be directly related to the user interface, we also asked questions using the System Usability Scale (Brooke, 1996) that was adapted for Polish users. Participants answered questions on the Likert scale (see Appendix A), where a summary score of less than 68 was taken to indicate that the usability was perceived as subpar. This scale measures in a quick and dirty way the system’s ease of use and overall satisfaction of its users. While it is technology independent, the average score changed with the different interfaces used for testing—a value of 68 is reported by many studies as an average for web interfaces (Bangor, Kortum, & Miller, 2008).

During analysis of the results, we used statistical methods to compare values for different age groups both within a single stage and between stages. Whenever in the paper a significant difference is stated, it means that one-way ANOVA was used and the difference between groups was considered significant if $p < 0.05$. To determine in which pair of groups the difference was most significant, the Tukey Honest Significant Differences (TukeyHSD) method was used. Both methods were implemented using software functions from the R “stats” package (R Core Team, 2016).

### 3. Results

The overall results of the experiment indicate that our tasks were not as simple as we imagined. Only one-third of the people were able successfully complete all four stages and 4 (13.3%) participants were unable to finish even a single task (Fig. 1). Differences were seen among age groups, with younger people performing slightly better—all members of this group always completed at least two stages. When we examine the stages separately, no correlation was seen between age and number of successes, with the single exception that members of the oldest group were less likely to succeed in Stage 4. This was however also because they more often resigned at this point as they were too tired to continue. Therefore, we cannot conclude whether this can be

### Table 1

<table>
<thead>
<tr>
<th>Stage</th>
<th>No.</th>
<th>Content of the question</th>
<th>Tools and features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Please mark on the map with the point or polygon the place where you live</td>
<td>point/polygon</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>Please mark on the map the location of the Jezycki Market</td>
<td>point/polygon</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Please mark on the map the location of the New Theater</td>
<td>point/polygon</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>Please mark on the map places that are the most attractive for tourists visiting Poznan (please add at least short comment to explain your choice, add 1–10 places)</td>
<td>point/polygon/comment other markers</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>Please mark on the map places that you like spending time the most (please add at least short comment to explain your choice, add 1–10 places)</td>
<td>point/polygon/comment</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>Please mark on the map places in Jezyce district in which you would like to change the current spatial development</td>
<td>polygon/comment/other markers</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Stage</th>
<th>Success criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Feature added within 250 m or within the same street block as the real location of the participant’s place of living</td>
</tr>
<tr>
<td>2</td>
<td>Feature added within 250 m or within the same street block as the real location of Jezyce Square or New Theater (only one was needed)</td>
</tr>
<tr>
<td>3</td>
<td>A single feature added and commented</td>
</tr>
<tr>
<td>4</td>
<td>A single polygon feature added and commented</td>
</tr>
</tbody>
</table>
attributed to the difficulty of the task itself. Fatigue and drop-out is also the reason why Stage 4 is the only part in which the oldest group spent the least time (Fig. 3).

What is surprising is that only 18 (60%) participants successfully completed Stage 1 success and only 15 (50%) succeeded in Stage 2 (Fig. 2). Analysis of the recording and CTA revealed that only 4 (13,3%) participants, from the two oldest age groups, were consciously reporting problems with finding required places and 3 (10%) of them subsequently opted for leaving the map empty. The remaining 8 (26,7%) participants marked wrong places seemingly unaware of their error. The fact that Stage 3 was the easiest for all age groups can be explained by the fact that we did not check the accuracy of the features; the participants had a lot of freedom, which we believe helped to alleviate stress and allowed seemingly free exploration of the map.

While at first evaluation, the participant age determined the amount of time spent on each set of tasks (Fig. 3)—with older people requiring more time—this was found to be significant only in the first two stages and when we compared the oldest and youngest groups. It seems that although map interface acts as a deterrent for older people, this effect quickly disappears. The two older groups spent significantly more time adding features in the first two tasks, but they also added many more features in Stage 3. Therefore, overall, there was no significant disparity between the number of points and polygons they contributed per minute. Surprisingly, people who stated that they had no previous contact with web maps did not display greater problems in using our mapping platform. No significant differences in the number of successes and time required for completing each stage were found.

The users perceived usability of our map interface was measured using System Usability Scale (SUS) (Fig. 4), on which scores greater than 68 represent good usability (Brooke, 1996). There is almost a linear relationship between age and SUS score, with older people perceiving a web page interface as much more complicated and demanding, compared to younger users. Using short questionnaire surveys, we also estimated the cognitive load that was required from the users to understand (mental load) and complete (mental effort) all tasks (Fig. 4). Our assumption was that the older participants would perceive the tasks as more difficult and strenuous, but the results did not confirm this hypothesis. Significant discrepancies were seen between individuals, which were much larger than the discrepancies between groups. One thing that was clearly visible was that in the oldest group, no participant marked the tasks as easy on CL scale (Appendix A).

To further investigate the process of creating the data, we examined the zoom levels that were used by the participants during map edits. A significant difference was seen between the youngest users and the remaining participants. The older groups were adding points using different map scales—zooming only when they were unable to find a feature they were looking for. Some participants did not use the zoom function at all, probably due to the lack of familiarity with web map interfaces. The behavior was visibly different in the 18–24 group, in which zooming was often the first thing the participants done when they started resolving the task. No feature was added at a level lower than 15; the most frequent level at which features were added was 18. This situation is also reflected in the spatial extent of the added features (see Figs. 6 and 7).
discrepancies in the above measures between age groups cannot be solely attributed to the problems with map interface. Initially, this may be the case—as suggested by more time required by older users in the first two stages—but with more time spent in the interface, these difficulties became less evident, while the differences in behavior toward the map and the space it represented became more pronounced. We investigated this by analyzing heatmaps from eye-tracking together with session video recordings. We selected two pairs of heatmaps that showed the differences in map use between younger and older participants. The first example (Fig. 6) illustrate how the older participant (P1) scanned the interface thoroughly at each stage before acting and the younger participant (P2) grasped the rules of adding features very early and, apart from Stage 1, paid very little attention to the interface itself. Both spent a comparable amount of time gazing at the map, but of the additional 15 min of screen time that P1 required, almost 80% was utilized to scan the interface. While P2 completed every task, P1 did not succeed at all tasks. What is also important is the fact that P1 was not aware of the location errors that were made (Stages

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>18-24</th>
<th>25-39</th>
<th>40-59</th>
<th>&gt;60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage 2</th>
<th>18-24</th>
<th>25-39</th>
<th>40-59</th>
<th>&gt;60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes</td>
<td>10.0</td>
<td>7.5</td>
<td>5.0</td>
<td>2.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage 3</th>
<th>18-24</th>
<th>25-39</th>
<th>40-59</th>
<th>&gt;60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage 4</th>
<th>18-24</th>
<th>25-39</th>
<th>40-59</th>
<th>&gt;60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Fig. 3. Time spent by participants at each experiment stage. Violin plots represent distribution for every age group with median and quantile ranges shown with standard box plots.
1–2) and that some of the point and polygons were not added at all (Stages 3–4). Possibly because of this, P1 reported levels of the cognition load measures that were below the average. While differences in this comparison are mainly visible in the way the map controls are used, they are also noticeable on the map itself. Visual inspection of the heatmaps revealed that P1’s gaze seemed to be less concentrated on a single spot and covered slightly larger areas of the map in Stages 2 and 3. While in Stages 1 and 4 it seemed that the opposite happened, NNI index was still lower in both cases for P2 (0.12 and 0.17 respectively) than for P1 (0.26 and 0.22). This can be even more easily seen in the second case (Fig. 7) that also compares two participants—older (P3) and younger (P4). Both participants were successful in completing the tasks in each stage and spent a comparable amount of time on each task. There was however, a difference in the map navigating behavior. P4 concentrated their gaze on very few spots, used very high levels of zoom, and added very little features in carefully chosen places (NNI: 0.08–0.016). P3 scanned larger areas and added more features on the lower zoom levels (NNI: 0.21–0.37). This difference in the way the map is used is evident between oldest and youngest participants with varied intensity but in the same direction—this is also visible in zoom level measurements (Fig. 5).

After the eye-tracking session and completion of the CL and SUS questionnaires, users were also asked to complete another short questionnaire survey (Appendix A) that was aimed at gathering people's attitudes (Q1) and perception (Q2) of the web mapping platforms. The general outcome of the exercise in terms of attitude was optimistic—(90%) of the 30 people answered that it changed their attitude toward participation to a more positive one. They also displayed a willingness to participate in public consultation; this change in attitude was more motivated by the ability to use internet tools rather than to deal with traditional meetings (Fig. 8). Answers to the questions on the possible improvement to this kind of participation tool (Fig. 9) revealed that only 27% of the participants wanted to work with a paper map and 57% would rather work with a mobile app. Interestingly, 40% of the people would welcome the option to communicate without using a map at all.

4. Conclusions & discussion

In this study, we have attempted to answer several questions. The first question (Q1) was concerned with the possibility of the successful use of PPGIS mapping tools within populations marked by distrust and a digital divide—in our case post-socialist Poland. Our participants have not displayed any resentment toward the maps and technology that were used, and in all age groups, there was an equal readiness to share spatial knowledge. It seems therefore that the distrust is not connected to the mapping technology but is solely the result of a negative attitude toward authorities (Rose-Ackerman, 2004). The perception of online mapping tools among participants of our experiment (Q2) was a very positive one. Very significant is also the fact that the presence of an online tool is a positive factor in the general willingness to participate—especially in view of the fact that for one-third of our participants, this was a first-time experience with web maps in any

Fig. 4. Cognitive load and usability measures measured by post-experiment questionnaire surveys. Violin plots represent distribution for every age group with median and quantile ranges shown with standard box plots.

Fig. 5. Zoom levels used by different age groups during map edits in Stages 1–3 of the experiment—data points recorded for each point added on the map. Violin plots represent distribution for every age group with median and quantile ranges shown with standard box plots.
form. This confirms that internet platforms can help to reach a wider audience (Jankowski et al., 2017) and that there is great potential in engaging people with such web mapping interfaces. We did not identify any major obstacles for introducing PPGIS mapping tools in regards to the way online maps are received by the public.

Our experiment also showed considerable differences in the behavior toward online mapping tools between the oldest and youngest participants (Q3). Older people found using the map interface much more difficult and demanding—a finding that was also reported in previous studies (Gottwald et al., 2016). They required much more time to accomplish each task and while they failed more frequently than younger participants, the differences were not significant. It seemed that the limiting factor for all participants was finding and adding features to the map. Surprisingly, all age groups failed often in the

Fig. 6. Heatmaps from eye-tracking recordings for participants P1 and P2. Warmer colors represent areas of stronger visual attention. Background images represent initial extent of the map window for each stage, which later changed according to user actions (zoom and pan). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
relatively simple task of finding well-known landmarks. Poplin (2015) found that more than 50% of the people perceived the task of finding a location as difficult; our study confirms that. This means that the quality of the data gathered using web map interfaces can be very low (Q4). What is even more important is that this quality is not dependent on the previous familiarity with similar tools or demographic characteristics of the participants—which is a significant finding for the participatory mapping agenda. Older people simply took more time to complete the tasks; their performance suggested that they even contributed more quality content due to their greater familiarity with the area. There is also a very visible need for some kind of a uniform rudimentary quality control. Recording zooming behavior or using simple navigational tasks can be employed as a proxy measure to estimate the locational accuracy of the added features. Results of the experiment can

Fig. 7. Heatmaps from eye-tracking recordings for participants P3 and P4. Warmer colors represent areas of stronger visual attention. Background images represent initial extent of the map window for each stage, which later changed according to user actions (zoom and pan). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
also be used to propose tips for designing inclusive and usable web map platforms (Q5). Resolving even simple tasks required a large amount of cognitive effort, and from the heatmap analysis, we can draw a conclusion that much of this effort, especially in the older age group, was dedicated to using the interface and deciphering the instructions. What we suggest as new in our study is that although enthusiastic about web maps, our participants would also welcome non-map forms of input (40%) and mobile apps (57%) that would perhaps help users with locational accuracy and navigational task. For example - dedicated mobile app could be used to pinpoint and describe locations simply by traveling to them with GPS enabled. This would potentially remove the need for the map as a data gathering tool.

In view of the above findings, it seems that the beginnings of a new research agenda can be sketched out that will focus on the practical applications of participatory mapping in urban planning and other related domains in which PPGIS tools are widely used. One of the reasons we are placing the focus specifically on urban planning is related to the fact that according to a study by Babelon et al. (2017), the rate of practical implementations in this area outpaces academic inquiry. We also fully agree with Radil and Anderson (2018) that PPGIS has become de-politicized and that it operates within rather than against existing power structures, which we think is particularly visible in urban planning where participation is often initiated by the authorities. Here, we also propose an attempt to re-invigorate critical engagement with the tools that are used in participatory mapping. This research agenda could be composed of three components: data quality, usability and user experience, and the utility of a map. It must be noted that apart from the usability, our three proposals differ from the key issues for mapping tools and technology that were identified during workshops organized at the 2017 Participatory Mapping/GIS Conference (Brown & Kyttä, 2018) and below we give our reason for this separate conceptualization.

The first component—data quality, is surprisingly often overlooked in participatory mapping. One of the reasons may be the fact that there are no established guidelines for controlling precision and accuracy of spatial data and that there are several variables involved that are difficult to control (Brown, 2012b; Czepkiewicz, Jankowski, & Zwoliński, 2018) or even that the location of the created spatial objects are fuzzy and undefined by design (Brown & Pullar, 2012; Czepkiewicz et al., 2018). Nevertheless, some methods for controlling accuracy were proposed that were based on using participant mapping as a proxy measure of accuracy (Brown, 2012b). Methods that we propose in this study are also based on proxy measures (zoom level and time) but also on accuracy in real locations (testing the ability to pinpoint real, well-known locations). To be practical, these methods should be further tested but we think that developing guidelines for such quality control in participatory mapping should be an important part of the future PPGIS research agenda. The second component—the usability and user experience—is perceived by PPGIS literature as very important both from the public engagement and successful participation standpoints (Brown & Kyttä, 2018; Gottwald et al., 2016; Poplin, 2015; Meng & Malczewski, 2010; Nivala, Brewster, & Sarjakoski, 2008). A poorly designed
interface can lead to lower participation rates and influence the amount of data that is being produced as well as its quality—which connects it to the first component. This issue is further complicated by the presence of diverse user groups such as those with older adults that only recently come to the attention of PPGIS research (Gottwald et al., 2016). Research on this topic is also insufficient because the technological landscape is changing constantly and recent implementations of Virtual and Augmented reality technologies (Ball, Capanni, & Watt, 2007; Olszewski, Gnat, Trojanowska, Turek, & Wieladek, 2017) need to be addressed within the PPGIS framework as they are increasingly implemented in planning practices. Finally, the third component deals with the utility of the map. By this, we mean that we should expand critical thinking beyond the questions of how and with what tools we implement mapping technology into questioning whether the map itself is needed in a particular context. Communicating spatial knowledge through a map is a cognitively demanding task that is burdened with uncertainties. Web mapping and PPGIS rely on tools that require technical knowledge and are therefore excluding certain groups. And most importantly, maps are power (Harley, 1988) that is always brought into the participation process through nonparticipatory selection of mapping and browsing tools, background maps, cartographic projections, symbology, and measures of location and attribute accuracy. For example, when participatory mapping is by default used in the planning process, it loses its critical edge and becomes yet another tool in the arsenal of an urban planner. We still think that the maps, whether digital, web, or paper, have an essential role to play in participation. But, at the same time, we see the need to rethink their often taken-for-granted place in the participatory process. We propose that it would be worth investigating whether other non-map forms of input would perhaps be more inclusive in a wide array of contexts, and whether the map in some cases could be better placed in later, analytical stages of the process.

Findings of this study can certainly be improved. The weight of our conclusions is limited by the size of the test group. This experiment showed, however, a few avenues for investigation—such as map interaction and perception—that could be pursued with larger, more representative samples. Mixed method approaches with the addition of focus groups and in-depth interviews could also help to gain insight into the behavior of map users. These kinds of tests should also be carried out within contexts other than urban planning. By proposing this research agenda, we first and foremost wanted to highlight the fact that PPGIS and participatory mapping pose a potential area for future research.

5. Study limitation

It is necessary to highlight the limitations of our research design and methods that were used in this case, as they can be significant. Since the empirical part of our study was mainly concerned with the usability of web maps, we decided to adopt an approach described by Nielsen (1994) as the “Guerrilla HCI” in which tests are carried out with a relatively small number of participants. This approach has also been used before in the PPGIS context by Haklay and Tobón (2003). It is generally perceived that groups between five and eight people are sufficient to detect approximately 85% of the usability issues (Zhao & Coleman, 2007), and we assumed that with our sample of 30 people, we were likely to detect an even larger percentage of potential issues with our platform. However, our sample is divided among age groups and the sample size in each group is too small to show statistically significant differences. At the same time, observed participant behaviors were significant as they indicated the potential problem areas and signalled underlying problems in using maps as a form of communication with public participation. We hope that our results will ignite (or re-ignite) research agenda for the future in which technology will play an increasingly pervasive role. Yet another limitation of our study was the fact that we used a mock PPGIS platform, which could be designed in another way that will differently influence the use of the map by our participants. We made every effort to minimize this but it still needs to be taken into consideration in the evaluation of the results.

Acknowledgements

Funding: This work was supported by the National Science Center, Poland [grant number UMO-2014/15/B/HS4/00839]. Authors would also like to thank anonymous reviewers, who were extremely helpful and kind.

Appendix A. Post-experiment questionnaire survey (translated from Polish)

The survey consisted of three parts – SUS Scale, CL Scale and Questionnaire survey.

SUS Scale

The purpose of this survey is to gather Your experiences and opinions about using this web portal. We would like You to answer the following questions according to Your personal experiences during the time You were using the web portal. Please mark one number beside each question where:

1 – definitely disagree
2 – rather disagree
3 – hard to tell
4 – rather agree
5 – definitely agree

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I think I would like to use this platform often</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>2</td>
<td>I perceive this platform as needlessly complicated</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>3</td>
<td>I think this platform is easy to use</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>4</td>
<td>I think I will need a help from an expert to use this platform</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>5</td>
<td>I think various platform functions were well integrated</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>6</td>
<td>I think there are many internal contradictions in the platform</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

2 Conference was held at California Polytechnic State University, San Luis Obispo, California, USA, Jul. 31-Aug. 3, 2017.
I imagine that many people will have no problem in learning to use the platform

Platform is very hard to use

I feel confident while using this platform

I had to learn many things before I could start working on this platform

**CL Scale**

The purpose of this survey is to gather Your experiences and opinions about using this web portal. We would like You to answer the following questions according to Your personal experiences during the time You were using the web portal. Please mark one number beside each question where:

1. definitely disagree
2. rather disagree
3. hard to tell
4. rather agree
5. definitely agree

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>It was hard to find right answers</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>2</td>
<td>I didn't do well in solving tasks</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>3</td>
<td>The content was complicated</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>4</td>
<td>Task were difficult</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>5</td>
<td>I’ve no difficulty in finding the right answers</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>6</td>
<td>Tasks were easy</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>7</td>
<td>I tried very hard to resolve tasks</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>8</td>
<td>I did not have to concentrate very much</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>9</td>
<td>Tasks were easy to solve</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>10</td>
<td>Tasks required mental effort from me</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>11</td>
<td>Tasks were easy to understand</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>12</td>
<td>I've put a great effort in resolving task</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

**Questionnaire Survey**

1. Place of work: ........................................................................................................
2. Place where You live: .............................................................................................
3. What functions would You like to add to this map?
   a) Nothing
   b) Commenting on addresses instead of features
   c) Drawing lines
   d) Drawing irregular shapes
   e) Editing comments
   f) Deleting comments
   g) Adding photo
   h) Searching for streets names
   i) Sharing in social media
   j) Other
4. What was the greatest difficulty in using this map?
   a) Zooming and panning
   b) Searching for familiar places
   c) Low precision
   d) Other users comments
   e) Unknown map
   f) Using computer
   g) Other
5. How to improve this kind of participation tool?
   a) Add paper map
   b) Create mobile app
   c) Add non-map forms of commenting and let experts put them on map
   d) I have no idea
   e) Other
6. Would You participate \n in public consultations in Your neighbourhood?
   a) I will regardless of the form
   b) I will if there would be a meeting with people
   c) I will if there would be an Internet tool
   d) I will not participate
7. Have today’s experience changed Your attitude toward participation?