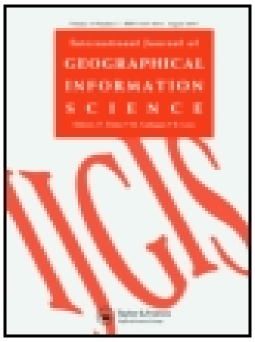

Taylor & Francis
Taylor & Francis Group

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: http://www.tandfonline.com/loi/tgis20

Usability evaluation for geographic information systems: a systematic literature review

René Unrau & Christian Kray

To cite this article: René Unrau & Christian Kray (2018): Usability evaluation for geographic information systems: a systematic literature review, International Journal of Geographical Information Science

To link to this article: https://doi.org/10.1080/13658816.2018.1554813

Published online: 19 Dec 2018.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2018.1554813

REVIEW ARTICLE

Usability evaluation for geographic information systems: a systematic literature review

René Unrau and Christian Kray

Institute for Geoinformatics, Westfälische-Wilhelms-Universität, Münster, Germany

ABSTRACT
The growing dissemination of geographic information systems (GIS) in recent years has increased the availability and use of geospatial tools. Despite this trend in GIS, usability is often not considered, and studies on usability are relatively rare. However, it has been shown that the effectiveness, efficiency and user satisfaction of these tools when used for geospatial tasks may be impacted by their design and terminology. In this article, we present a comprehensive review of the relevant literature. We analyze and compare publications of GIS usability evaluations, the methods that were used, and the findings reported. We thus produce a more detailed picture of GIS usability evaluations. Our contributions are (1) an overview of methods applied in case studies for evaluating the usability of GIS; (2) the identification of gaps and resulting weaknesses of previous studies; and (3) challenges and opportunities for future GIS usability studies. Our findings reveal a lack of usability guidelines for GIS, a lack of methods for analyzing complex workflows, and a lack of usage studies over longer periods of time. These contributions can help GIS developers and designers to improve the usability of their systems.

ARTICLE HISTORY
Received 6 April 2018
Accepted 28 November 2018

KEYWORDS
WebGIS; usability; evaluation methods; literature review; user interface

1. Introduction

Already 10 years ago, Haklay and Zafiri (2008) stated that 'GIS [geographic information systems] have become more commonplace at work and home'; and they are no longer solely used by a few professionals. Public authorities of all levels are building geospatial web services in order to improve information dissemination, offer public services or integrate citizens into planning processes. At the same time, modern web browser technologies facilitate the integration of business processes related to geospatial data into GIS. The increased use of GIS, in particular by novice users with little GIS experience, means that more profound usability evaluations of user interfaces (UIs) are needed. The term *usability*, as defined by the International Standardization Organization (ISO) as part of their standard on ergonomics of human-system interaction, means: 'The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use' (ISO, SFSEN 1998). Even before this, Nielsen (1994) proposed an engineering principle that approaches usability with five measurable attributes: learnability, efficiency, memorability, errors and satisfaction. Thus, the usability of an application or system depends heavily on the user – their abilities to use it easily and efficiently and to be satisfied with it.

CONTACT René Unrau r.unrau@uni-muenster.de

GIS novices struggle with typical GIS elements, for example, layer controls and map icons, which make it hard for them to exploit all tools and may even prevent them from using the system at all (Skarlatidou and Haklay 2006, Hoover *et al*. 2014, Poplin 2015). This may affect the usability of the overall application. These users have varying levels of expertise, backgrounds and previous experience, and they might not be able to acquire the required technical knowledge easily or at all (Haklay and Tobón 2003). This knowledge about GIS users is crucial for application designers to understand their users' expectations, perceptions and needs. Such knowledge supports the identification of usability issues prior to use and hence can promote the design of an efficient and pleasant experience.

Given the growing trend of applying GIS in new domains with new users while also offering complex tools and UIs for geospatial tasks, it is unclear if such systems offer suitable usability for their diverse users. To examine this, we carried out a literature review of usability evaluations of cartographic interfaces with GIS functionality, and we provide an overview of the applied evaluation methods and investigated systems. We analyzed different system types (e.g. simple web mapping systems) which were built for a wide range of users. All case studies that fulfilled the inclusion criteria were assessed regarding a defined set of parameters, which included the system's type, target user group, evaluation findings, data collection approaches, the study's context and the number of participants. Our contributions are as follows: (1) we give an overview of methods applied in case studies for evaluating the usability of GIS; (2) we identify gaps and resulting weaknesses of previous studies; and (3) we highlight challenges and opportunities for future usability studies regarding GIS.

The remainder of this article is structured as follows: Section 2 describes in detail the method we used for searching and selecting publications. Here, we also present the rationale for the chosen sources, criteria and parameters. Section 3 gives an overview of the results, our defined categories and the distributions among related usability evaluation parameters. In Section 4, we discuss our method as well as the results and we formulate research challenges and opportunities. The paper ends with a conclusion of our contributions in Section 5.

2. Review method

To provide insights into various aspects of GIS user evaluation studies – for example, underrepresented target user groups, the distribution and balance of data collection approaches, or the concentration of findings for a particular aspect – we assessed and filtered a wide range of publications based on a well-defined process (see Figure 1). Although our approach is not a meta-analysis in the strictest sense (Gurevitch *et al*. 2018), the study presented here is a meta-review of published usability evaluation studies of GIS.

For the identification step, we queried major databases in the field of computer science (ACM Digital Library, IEEE Xplore, dblp), one search engine (Google Scholar), and the proceedings of two conferences (ACM CHI Conference on Human Factors in Computing Systems and SIGSPATIAL). We also reviewed selected journals in the domains of cartography, GIS and human–computer interaction (HCI). Our choice of journals was based on the GIScience journal ranking (Caron *et al*. 2008). We screened the aims, scope and recent publications of the ranked journals for the topic of usability and selected five journals for our review (see Table 1).

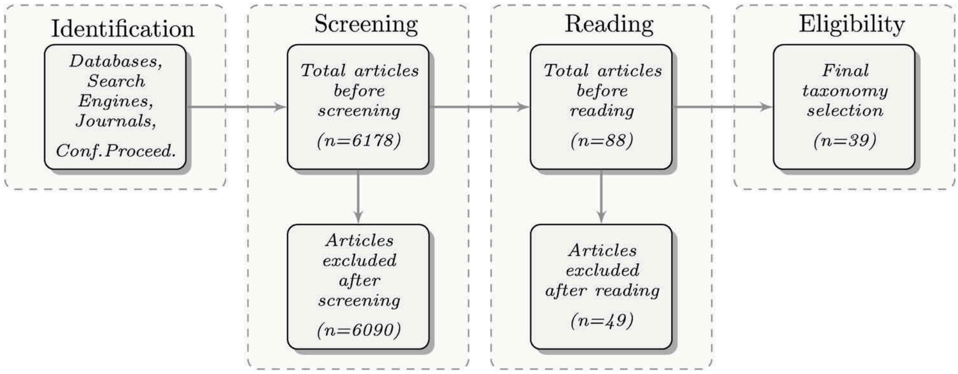

Figure 1. Peer-reviewed publications were selected from various literature sources (databases, search engines, journals, conferences) and filtered based on a well-defined process.

Table 1. Selected journals in the domain of GIS and cartography were screened for review process by considering articles published between 2008 and 2017.

Journal title
International Journal of Geographical Information Science (IJGIS)
Cartography and Geographic Information Science (CaGIS)
Geoinformatica
Transactions in GIS
Cartographic Journal

To narrow down the list of results to relevant fields for the databases and search engine, we made use of a Boolean search containing the following terms: ('map' OR 'GIS') AND 'Usability'. For conference proceedings and journal articles, the selection was limited to items published in the period between January 2008 and December 2017. The overall number of identified publications was 6178. Several publications were discovered in multiple places, such as on databases and Google Scholar and were included only once in the count.

Next, we screened all publications based on their titles and abstracts. Publications that were not peer reviewed or could be clearly defined as not relevant to our scope were excluded here. These were publications that did not evaluate the usability of one or more particular systems but instead provided, for example, an overview of the system's functionality, architecture or deployment. Furthermore, publications that discussed geovisualization aspects or usability guidelines without referring to a particular system were removed as well. We also ensured that all publications were peer reviewed by checking the editor's or publisher's review process on their website.

In the third step (reading), the remaining set of 88 publications was fully read by at least one of the authors, and each publication was checked to determine if it met the following inclusion criteria: (1) We acknowledged that the study evaluated usability aspects if the applied methods could be assigned to at least one of the defined categories (testing, inquiry, inspection, see Section 2.1.2). (2) As this review focuses on GIS functionality, we ensured that the evaluated components addressed general aspects of (cartographic) UIs. For instance, many usability issues detected with desktop GIS (e.g. Esri's ArcMap) are related to the available screen size (Haklay and Jones 2008, Haklay and Zafiri 2008), the hints provided by the system (Hossain and Masud 2009), the displayed error messages (Davies and Medyckyj-Scott 1996), the number of available tools (Edsall *et al*. 2001) and the quality of the provided documentation (Davies and Medyckyj-Scott 1994); thus such studies were included. (3) Likewise, publications on web mapping systems (e.g. an interactive map found on a news website) were considered to be relevant if their findings were also applicable to GIS, for example, if they were related to well-known map tools (Nivala *et al*. 2008) or addressed aspects of interactive maps (You *et al*. 2007, Manson *et al*. 2012, May and Gamble 2014, Kiefer *et al*. 2017) or interface complexity (Roth and Harrower 2008). (4) We excluded usability evaluations of mobile cartographic interfaces because – although GIS has become available for mobile devices in recent years – the context of use, display size and user interactions are substantially different from other systems (Reichenbacher 2003). Comparing these mobile systems with stationary systems would demand additional parameters and was therefore beyond the scope of this article. In the end, 39 papers were rated as eligible and included in our literature review (see, Table A1).

2.1. *Usability evaluation dimensions*

In order to thoroughly analyze the included publications, we defined a set of usability evaluation parameters. In this process, the key goals were to provide an overview and allow comparison of commonly applied strategies for usability evaluation in the GIS domain. In the following, we introduce and motivate the parameters that were used in our literature review.

2.1.1. *Application type and target users*

Applications were categorized into desktop GIS, web mapping systems and hybrid approaches that provide limited GIS functionality and are operated in a web browser (WebGIS). Additionally, we distinguished applications based on the role of target users and assigned each case study to one of the following groups:

(1) *Casual*: systems that are used by citizens or customers without profound GIS expertise and that are typically provided by either public administration agencies or commercial providers.
(2) *Domain*: systems that are used by people with specific domain knowledge, such as in health, environment, planning, forestry or fisheries fields.

We differentiated between these user groups to assess users' characteristic differences, for example their motivations and goals.

2.1.2. *Evaluation methods, context and participants*

While including actual users during usability evaluations represents an important aspect of a study (Nielsen 1994), an alternative is to consult usability or domain experts for feedback on an interface. We, therefore, identified the following categories to compare evaluation methods on a higher level:

(1) *Testing* is conducted by observing users as they interact with the interface of the evaluated system and then collecting data, i.e. as users finish the task.
(2) *Inquiry* is done by asking users to evaluate a system and provide feedback after interacting with the interface; this feedback is typically achieved via interviews, surveys or ratings.
(3) *Inspection* is conducted by users or experts who analyze the system based on a set of heuristics or criteria.

These methods cover summative and formative objectives and, thereby, help us to achieve a general assessment of GIS usability as well as uncover particular problems. Although user and expert inspections differ in their designs, they share the purpose of identifying the potential in an interface. However, we will report on the differences between these two groups in our results.

These evaluation methods were further differentiated by considering additional parameters. Some parameters that might affect the relevancy of results are the context of the study (lab vs. field) as well as the selected participants. Lab studies ensure a tightly controlled and reproducible environment but do not consider the real context of use (Lallemand and Koenig 2017). Field studies constitute a more naturalistic approach but are sometimes difficult to conduct, and their added value is controversial (Nielsen *et al*. 2006). We, therefore, assessed the context of each study and assigned each publication as either being lab- or field-based. Additionally, we assessed the number of study participants (sample size) as well as their expertise and relation to the system being evaluated.

2.1.3. *Usability findings*

Evaluation findings in the reviewed publications covered a wide range of usability aspects with different levels of detail. Frequently, the investigation of usability issues through the chosen evaluation method was part of a user-centered design (UCD) process with preceding and subsequent steps. For example, Roth *et al*. (2015) present an iterative UCD process with a needs assessment study, an alpha release, a usability study and a revised interface concept. We concentrated our analysis on the findings that emerge from these studies related to usability aspects pertaining to GIS and excluded contributions that went beyond this topic. Two approaches were established, each with a set of categories to examine these findings from different angles.

We established these categories after reading the selected publications. In the first cycle, we pre-coded relevant passages that included contributions throughout the publications. We then coded actual findings into the addressed components of the system and its accompanying material (documentation, map, legend, etc.). In the second cycle, we rearranged our codes and grouped similar findings that shared characteristics into our final categories. If a finding could be assigned to more than one category, it was coded with all matching categories.

The first approach to assess the usability evaluation studies was based on the ISO, SESEN (1998) definition of usability that represents well-defined and abstract measures:

(1) *Effectiveness*: the accuracy and completeness with which specified users can achieve specified goals in particular environments.
(2) *Efficiency*: the resources expended in relation to the accuracy and completeness of goals achieved.
(3) *Satisfaction*: the comfort and acceptability of the system as perceived by its users and other people affected by its use.

The second approach was focused on the main purpose that each particular finding addresses.

(1) *User guidance*: findings that cover general aspects of the entire system that are related but not limited to the documentation, error messages and recovery of the systems, as well as conclusions about users' orientation.
(2) *Strategy*: findings that address elaborations of users' interactions during multi-step tasks or entire sessions, such as task scores or users' evaluations of the selected workflow.
(3) *Tool use*: findings that include results from inquiry, inspection or testing which address observations of specific GIS elements (e.g. search, selection, overview map, measurement or further analysis tools) as well as the absence of expected elements, missing or unused functionality, and so on.
(4) *Map interaction*: findings that focus on the essential feature of every cartographic application and include, for example, a comparison of different techniques (pan, zoom) as well as remarks on map rendering aspects or the initial extent.
(5) *Icon representation*: findings that are related to observations of the used symbology (e.g. tool icons, map symbols, etc.).

2.1.4. *Data collection approaches*

After the evaluation method has been categorized, we analyzed the data collection approach used in the different studies. Broadly speaking, there are quantitative as well as qualitative data collection approaches. Inspections usually include the application of heuristic evaluations, expert comments or interviews, while user testing and inquiry approaches are more diverse. After reading the selected publications, the following range of quantitative approaches was identified:

(1) *Task scores*: user testing metrics such as task completion time or error rate.
(2) *Subjective ratings*: ratings by users or experts that are based on numeric scales, thereby making them quantitative.
(3) *Interaction logging*: intermediate interactions or even fine granular data like mouse tracks are incorporated to extend task scores.
(4) *Eye tracking*: measures the time users spent watching certain UI elements and other metrics, such as the number of glances, and so on.
(5) *Heuristic evaluation*: typically scale ratings based on UI standards performed by experts (e.g. identification of violations of dialogue principles).

Additionally, the following qualitative approaches were identified:

(1) *Think aloud*: the user is required to speak while interacting with the interface to help evaluators understand the user's thoughts.
(2) *Videotaping*: allow evaluators to analyze sessions after the actual test, where evaluators can, for example, discuss user behavior with experts or the users themselves.
(3) *Questionnaire*: a set of (open-ended) questions used to assess, for example, users' ages, backgrounds or satisfaction with the evaluated system (may also be quantitative). There is a set of standardized questionnaires that facilitate comparison across studies, for example, NASA TLX (Hart 2006) or SUS (Brooke 1996)
(4) *Interview*: similar to a questionnaire but allows evaluators to ask follow-up questions or change the order of questions. In addition, this is usually carried out in a synchronous way.
(5) *User comments*: comments can be collected during or after testing, and they might not be in response to a specific question, such as in interviews or questionnaires.
(6) *Expert comments*: comments usually provided after the inspection of a system in order to extend heuristic evaluations.
(7) *User observations*: evaluators take notes as they watch users interacting with the interface.

3. Results

3.1. *System types and target users*

Of all case studies that fulfilled the inclusion criteria (see Section 2), as many as 26 evaluated systems were built for casual users, whereas the remaining 13 systems targeted domain users (see Figure 2).

Although a number of publications in the domain users category originated from the same authors (Komarkova *et al*. 2007, 2009a, 2009b, 2010, 2017), they evaluated different eGovernment WebGIS (that the authors did not develop themselves) with different

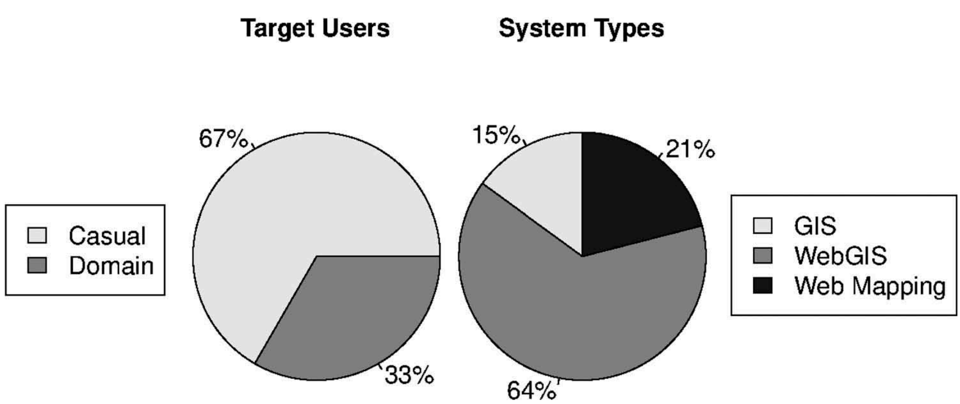

Figure 2. Percentages of (sample size $n = 39$) target users types (casual vs. domain) and system type (GIS vs. WebGIS vs. web mapping) show that most evaluated systems are built for casual users and WebGIS are the most often evaluated type of system.

methods, data collection approaches and participants. Systems with little GIS functionality were limited to analyzing data such as crime or poverty (Çöltekin *et al*. 2009). Selected public participation WebGIS went beyond this rather passive use by enabling user contributions of data or comments on, for example, planning processes (Poplin 2015). More sophisticated applications supported complex geoprocessing tasks such as computing driving distance polygons with respect to water hauling areas (Hoover *et al*. 2014).

All considered desktop GIS were built exclusively for domain users (six systems, 15%). WebGIS were addressed by the majority of selected publications (25 systems, 64%) and represent the only type built for domain users and casual users. The remaining case studies evaluated web mapping systems that were exclusively built for casual users.

The evaluated systems for domain users either focused on a single field of application, such as health (Edsall *et al*. 2001, Cinnamon *et al*. 2009), crime (Roth *et al*. 2015) and hydrologic (Tsou and Curran 2008) data analysis, or compared multiple systems from different fields (Haklay and Zafiri 2008).

3.2. *Methods and parameters*

The selected case studies used one or more usability evaluation methods. User testing represents the most commonly applied method among evaluators and was applied in 82.1%, closely followed by inquiry (64.1%) and inspections (23.1%). All nine inspected systems were reviewed by experts; none of the selected publications had users analyze interface heuristics or similar criteria.

As already mentioned, usability evaluations were often part of an overall UCD process which extended stand-alone evaluations of usability by additional stages that focused on, for example, utility (Tsou and Curran 2008). In the usability evaluations, the testing and inquiry methods were typically used in tandem to generate more insights from different perspectives. Ingensand and Golay (2011) conducted a remote evaluation that considered detailed interaction logs for measuring the performance of predefined tasks (testing) and involved user satisfaction ratings based on a scale from 1 to 5 (inquiry). Fechner *et al*. (2015) evaluated an approach for real-time collaborative editing of geo-data and provided an additional combination by conducting a user study (testing) and expert interviews (inquiry). Furthermore, some of the UCD approaches repeated user testing or expert inspections to integrate and verify feedback or to investigate different aspects of previous evaluations. Manson *et al*. (2012) asked two groups of participants to perform the same tasks. While in the first phase, mouse tracking was applied to test the usability of map navigation, the second phase was extended to eye-tracking. Similarly, Roth *et al*. (2015) prepared two identical online surveys in order to compare usability scores from the beta and full release of the system. Although UCD supports the development phase of GIS by involving users in many stages, none of the studies included in our literature review evaluated the use of the same users during *and* beyond the development phase.

The mean and median number of study participants were very similar between evaluations using testing and inquiry approaches (see Table 2). However, the median values greatly deviated from the means and represent less than half of the value in both cases.

Table 2. Average and median number of study participants for each method. Two of the usability evaluations that applied inspection as a ethod did not mention the number of evaluators and were therefore excluded from this table.

	Mean	Median
User *testing*	40.53	18.5
User *inquiry*	41.55	18
Expert *inspection*	5.7	4

As the median is more sensitive to outliers, we can attribute the higher mean to usability evaluations that included high numbers of participants in testing and inquiry approaches. Taking a closer look at those outliers, two studies conducted remote evaluations (Unrau *et al*. (2017) with 152 participants, and Ingensand and Golay (2011) with 331 participants), one utilized a postal survey (Davies and Medyckyj-Scott (1994), 159 participants) and one was conducted with college students (Komarkova *et al*. (2009a), 165 participants).

Regarding the distribution of target users among the three evaluation methods, more than half of the evaluations which applied inspection as a method investigated systems for domain users and, thereby, represent the highest proportion for this user group among all methods (see Figure 3). GIS built for casual user scenarios represent the majority of systems that were evaluated via testing (23 systems) or inquiry (18 systems).

We further analyzed the context in which testing and inquiry usability evaluations were conducted. In general, lab-based studies were performed more often than field-based studies (see Figure 4). However, the data expose a higher number of field-based studies for inquiries (28% field-based vs. 16% lab-based).

The quality of the sample selection constitutes an additional indicator for the significance of testing and inquiry evaluations. We assigned each usability testing and inquiry study to one of the following two groups: The first group contained evaluations where the

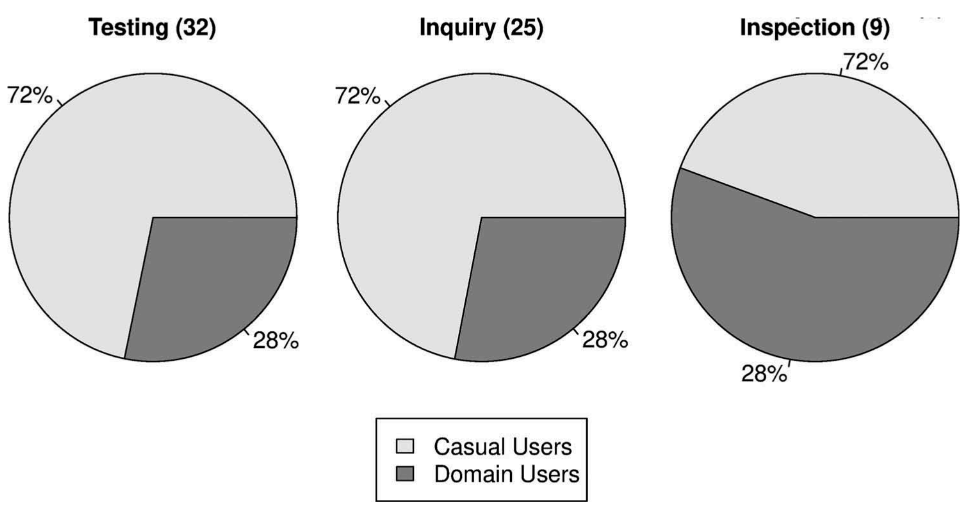

Figure 3. The evaluated systems were assigned to be either accessed by casual or domain users. The pie charts show percentages of these system categories among usability evaluation methods. Absolute numbers indicate the number of evaluated systems for each method.

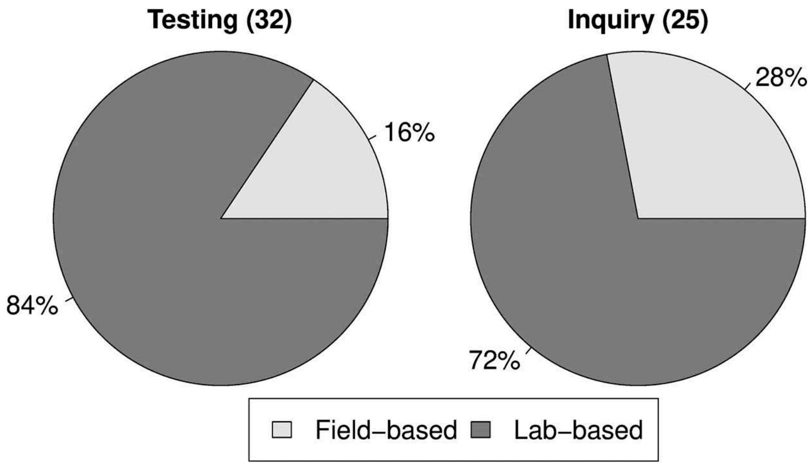

Figure 4. Percentages of the study context (lab based vs. field based) among testing and inquiry evaluations. Absolute numbers indicate the amount of evaluated systems for both methods.

participants did not represent the actual target users, meaning that they might not have had the appropriate level of GIS knowledge, or where the study authors did not comment on the selection of participants at all. Oftentimes, GIS such as eGovernment applications for the public were tested using participants with an academic background or college students (You *et al*. 2007, Ingensand and Golay 2011). If these students had a background in GI-science, we rated them as not being representative of actual target users of eGovernment systems built for the masses. The second group contained evaluations with participants that represented actual users of the GIS (Ingensand and Golay 2010, Unrau *et al*. 2017).

In total, the number of evaluations with actual target users and without actual target users was the same. However, just 33% of GIS usability evaluations for casual users was conducted with participants that represented the actual target user group (see Figure 5). In contrast, 84% of GIS usability evaluations for domain users was conducted with actual target users as participants.

3.3. *Usability findings*

We categorized findings by general usability measures and GIS-related usability aspects. Using our measures, the most reported general usability measures of all investigated case studies were about the efficiency (48.7%) followed by user satisfaction (38.4%) and then effectiveness (20.5%, see Figure 6). Findings that are related to efficiency were reported most often in applications built for casual users.

For the GIS-specific usability aspects, the most frequently reported findings involved the usability potential for guidance and documentation (53.8%). Stated problems were, for example, 'error reports that do not explain clearly why they have occurred or how they should be corrected' (Davies and Medyckyj-Scott 1996, Komarkova *et al*. 2010) and the 'lack of feedback after the user had issued a command to indicate whether the system was performing any operations' (Haklay and Tobón 2003). Researchers (Hossain and Masud 2009) also discovered that hints to indicate data retrieval by users were missing (the interface did not reflect

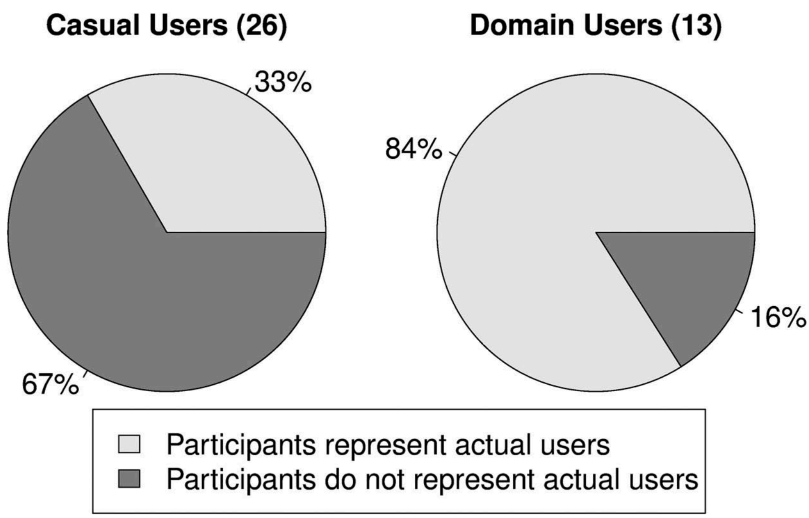

Figure 5. Percentage of studies that were conducted with actual target users versus those which were not among the target user group. Absolute numbers indicate the number of evaluated systems for both types of target users.

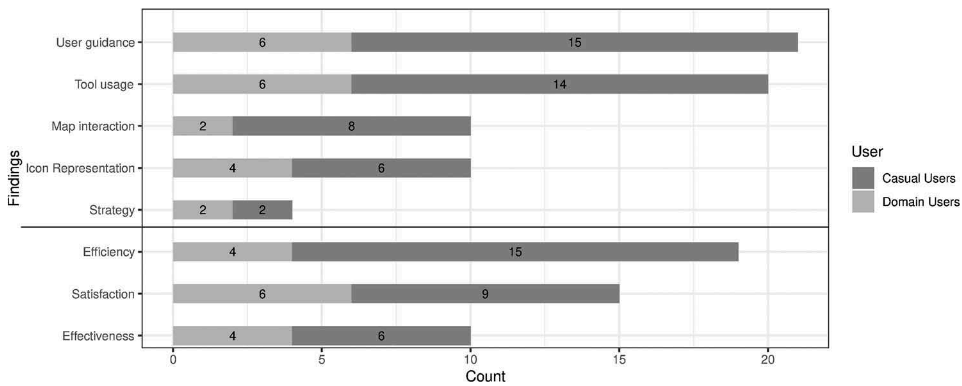

Figure 6. Usability findings assigned to ISO usability measures (effectiveness, satisfaction, efficiency) and GIS related aspects reveal high values for efficiency, tool use and user guidance but are similarly distributed among target users.

the loading state) or 'no visible user interface element which enables users to add nodes to an existing line or area' (Weber and Jones 2011). Furthermore, some of these evaluations revealed that users struggle with basic tasks, such as 'adding layers to the map' (Newman *et al*. 2010), finding a specific plan (Komarkova *et al*. 2009b) or locating a help button (Çöltekin *et al*. 2009). In addition, expert inspectors rated the system's ability to describe the content poorly (Steinmann *et al*. 2004) and discovered a 'lack of visual guidelines on the interface to indicate how tools were operated' (Haklay and Tobón 2003). Another GIS-specific finding that was often reported dealt with tool use (51.3%). Usability evaluations revealed problems with, for example, the size of the search bar and how results were displayed (Nivala *et al*. 2008, Poplin 2015, Komarkova *et al*. 2017), how distance measurements results were displayed (Ingensand

2006), or the location of overview map tools (Komarkova *et al.* 2009a). The case studies also included findings related to map interactions and icon representations (both 25.6%). Issues with icon representations included, for example, users expecting different functionality behind a button (Unrau *et al.* 2017), 'icons used for map function buttons [that] were not salient to non-GIS participants' (Newman *et al.* 2010) or icons that were not sufficiently expressive (Komarkova *et al.* 2010). Similar problems were uncovered for buttons (Ingensand 2006, Çöltekin *et al.* 2009), tools (Komarkova *et al.* 2009a, Düren and Bartoschek 2013) and pop-ups (Alaçam and Dalcı 2009). Finally, the case studies also revealed that evaluations of user strategies within the application were identified in just 10.3% of all selected publications. These findings addressed the comparison of actual workflows with those expected by GIS maintainers (Unrau *et al.* 2017) and among different study participants (Fechner *et al.* 2015).

3.4. *Data collection approaches*

Approaches for collecting data in usability evaluations are manifold, and most studies applied more than one approach. Many approaches are designed to support the investigation of one or more particular usability metrics. For instance, task scores are helpful to estimate the efficiency and effectiveness of an application. Evaluators often merged results from multiple approaches before drawing conclusions. Fechner *et al.* (2015) collected user comments from participants but also logged interactions during use. Due to this reason, comparing the application of data collection approaches with actual findings was not meaningful. We did, however, distinguish between quantitative and qualitative approaches to infer the tendency of an evaluation (see Table 3). In total, 90% of all considered publications made use of quantitative approaches, whereas 72% applied qualitative approaches.

Among the selected case studies, capturing task scores represented the most frequently applied approach (35.9%). Studies that applied task scores evaluated, for example, completion time (Çöltekin *et al.* 2009) and rate (Weber and Jones 2011). The next most common types of data collection approaches were subjective ratings by experts or users followed by a set of qualitative measures, such as user comments that were collected on expected functionality (Newman *et al.* 2010), think-aloud protocols that, for example, 'gave the impression of a better organized interface' (Skarlatidou and Haklay 2006), and videotaping to identify at which point in the process users struggled (Weber and Jones 2011). Each of the top six data collection approaches was applied in at least one-quarter of the selected

Table 3. Data collection approaches ranked by their number of applications in the selected evaluations.

Data collection approach	Quant.	Qual.	Frequency (%)
Task scores	✓		35.9
Subjective ratings	✓		30.8
User comments		✓	30.8
Think aloud		✓	28.2
Videotaping		✓	25.6
Questionnaire	✓	✓	25.6
Interaction logging	✓		20.5
Eye tracking	✓		15.4
Heuristic evaluation	✓		10.3
Expert comments		✓	7.7
Interview		✓	7.7
User observations		✓	5.1

evaluations, with only subtle differences in use frequency (25.6–35.9%). However, they were fundamentally different in their execution and include information gathered from users and experts as well as quantitative and qualitative data. Similarly, the remaining approaches were also quite distinct, including rather novel ideas such as eye-tracking for evaluating how frequently and how long users fixated on something (Manson *et al.* 2012) as well as well-established concepts like interviews to identify user preferences (Kiefer *et al.* 2017).

With the exception of eye tracking and user observations, all approaches for harvesting data within usability evaluations were applied for at least one GIS for domain user, and one for casual users (see Figure 7). The approaches of task scores and user comments were most often applied to casual user systems, whereas subjective ratings were the most applied approach for getting feedback on domain user systems.

4. Discussion

By selecting and analyzing publications that contain usability evaluations of GIS, we aimed to provide an overview of investigated GIS applications, applied data collection approaches and usability findings. The outcomes provide a detailed view of case studies that were reported in the literature and commonly identified usability issues in this domain. Overall, the data suggest that there are too few in-depth investigations of user sessions and interactions, as can be seen from the exceptionally low amount of user strategy analyses that were performed for both target users groups (see Figure 6). Further, these studies rarely determined the underlying reasons for low efficiency scores or why user guidance was lacking. Regarding target users, we found that 50% of testing and inquiry evaluations were conducted with participants that did not represent actual target users of the investigated systems. In the following sections, we discuss our review method as well as the main areas of our results: target users, study participants, usability findings and data collection approaches. We conclude by highlighting certain challenges and opportunities for usability evaluations of GIS.

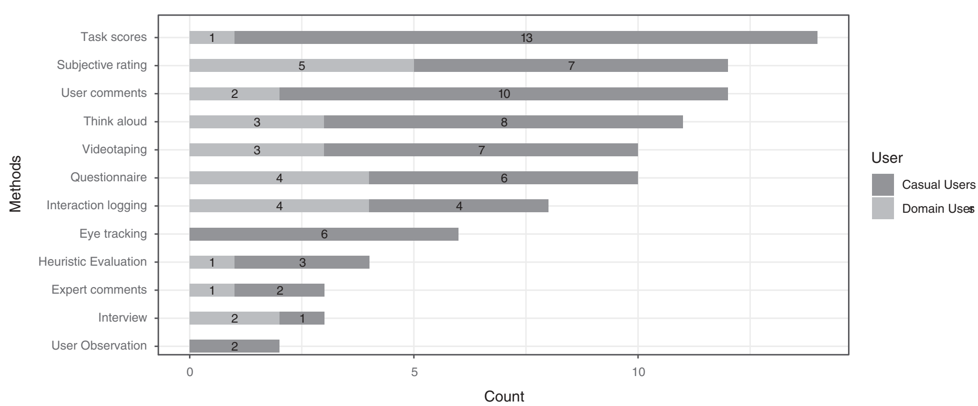

Figure 7. Ranking of data collection approaches and distribution of target users among these reveal outliers on both ends of the spectrum.

4.1. *Review method*

The literature review consisted of an extensive examination of relevant conferences, journals, database libraries and one search engine. Nevertheless, our review may not have been exhaustive, as we may have missed some publication that would have fulfilled the defined inclusion criteria. The reasons for this limitation are two-fold. First, the term *geographic information system* is not used to describe all applications that provide tools for collecting, displaying, manipulating and analyzing spatial data from different sources. Often, broader terms are given for these systems, such as *interactive maps* (Roth *et al*. 2015), *web mapping applications* (Tsou and Curran 2008) or *geovisual analytics applications* (Roth and MacEachren 2016). Additional terms for these systems might exist but were not covered by our search queries. Second, additional publications might not have been identified if they were published in journals or at conferences outside the domain of cartography and GIS. The ACM Conference on Human Factors in Computing Systems (CHI) was the only conference for HCI publications that we considered, and there are likely different cultures in the different disciplines: publications from CHI frequently have a higher impact than those in HCI journals. Nonetheless, the large number of screened CHI publications (4059), combined with keyword queries for additional case studies in the selected sources produced a broad and diverse set of related evaluations.

GIS has undergone rapid development both technically and regarding its use in different areas. The borders of system types are disappearing as WebGIS become more powerful and attract new users. Accordingly, the evaluated applications and applied methods are subject to continuous change. The publication years of selected case studies may represent an important selection criterion in order to provide a current snapshot for GIS usability evaluation. However, the investigation of usability within GIS depicts an underrepresented field in research and case studies are relatively rare. The median publication year of the 39 selected items is 2009 ($\mu = 2009.45$), meaning that half the case studies analyzed are older that 9 years and half are more recent. This underlines the need for more scientific publications of up-to-date GIS usability evaluations.

4.2. *Target user*

The distribution among systems built for casual and domain user revealed a high proportion of systems for casual use (66.7%). This emphasis could be explained through three factors. First, this domain represents one of the most quickly growing application areas for providing GIS functionality. Driven forward by European Union legislation[1] and national strategies,[2] many public administrations and authorities are setting up data portals and carrying out evaluations of them. Second, typical Internet users and, thereby, GIS novices all represent potential users, which might promote demands for usability analyses. Third, usability might be perceived as less important for domain users, as they are more willing to invest resources into learning new UIs. In this context, Unwin (2005) emphasizes the 'danger that existing tools will be mapped uncritically and without modification' to users even though users' education and backgrounds vary in the growing market of business to business GIS. Therefore, the challenge will be to transfer expert GIS functionality to novice applications while facilitating use by consumers that do not have extensive GIS knowledge. This is not an easy task, as such systems vary greatly in terms of target domain and provided functionality.

4.3. *Study participants*

Many usability evaluations considered in this review were conducted with participants that do not represent actual users of the evaluated systems (Figure 5). The underlying reasons for this vary. First, the targeted research question (e.g. fundamental perception) might not require participants to have a specific background or domain knowledge. Second, the acquisition of actual target users is often complicated and expensive. For example, systems used in business processes might be used by only a small number of individuals who cannot be consulted for usability testing. Third, companies which operate usability labs are not always willing to publish their findings, as this information might constitute intellectual property. Even if the results are published, these might not appear in peer-reviewed journals but instead on company websites (e.g. https://www.microsoft.com/en-us/research/). However, it is important to test UIs with the actual target users of the application. Nielsen (1994) emphasizes the need for user testing by persons who are not involved in the development of the software (i.e. not developers, designers or managers) in order to gain new insights into the usability of a UI. Nielsen goes on and even de-emphasizes the importance of the actual technique or the number of participants, as some usability testing is always better than none.

4.4. *Usability findings*

Every second evaluated application had problems with user guidance. Frequently, this leads to reduced use, as casual users are not inclined to invest much effort into, for example, reading the documentation that goes with infrequently accessed software (Nielsen 1994). Instead, major web mapping systems (such as GoogleMaps) may have already influenced users' expectations of GIS usability, so that users expect other systems to behave similarly. In contrast, user strategy-related findings (see Figure 6) were reported rarely and might indicate that there are currently not enough in-depth workflow analyses. Research in this area might be limited due to the technical expertise required to set up such logging mechanisms as well as the lack of appropriate data analysis for huge session data sets (Lazar *et al*. 2017). Nonetheless, a thorough understanding of common workflows can also benefit user guidance and interface design in general (e.g. via wizards leading users through a workflow).

4.5. *Data collection approaches*

The most often applied data collection approach among the selected case studies is based on quantitative data (see Table 3). This might be because such methods frequently require less effort which is desired when carrying out a snapshot study (Haklay and Zafiri 2008) or a pilot experiment (Haklay and Jones 2008). In contrast, traditional user observations were used only in 5.1% of all case studies. This might be partially due to the complex interfaces and functionality of GIS that make real-time observations of user behavior more difficult, compared to, for example, when users are visiting news websites. Eye-tracking represents an additional quantitative data collection approach and was used only rarely as well. However, this approach might be in particularly interesting for applications that focus on exploratory use such as WebGIS, which are provided by authorities for exploring open geospatial data. As eye-tracking is typically used for measuring users' focus points and duration, it was either applied for investigating map interactions (Manson *et al*. 2012, May and Gamble 2014, Kiefer *et al*. 2017) or users' exploration of buttons and icons (Alaçam and Dalc 2009, Çöltekin *et al*. 2009, Weber and Jones 2011). In the future, modern data analysis techniques, such as the combination of data processing and human domain knowledge through visual analytics (Keim *et al*. 2008, Unrau *et al*. 2017), might facilitate the investigation of the underlying reasons by combining quantitative data processing with qualitative knowledge from domain experts.

4.6. *Challenges and opportunities of usability evaluations for GIS*

Based on our analysis and discussion of GIS evaluations, we discovered three main challenges and opportunities for future usability studies.

4.6.1. *Usability guidelines for GIS*

Usability issues related to user guidance were reported most often (53.8% of all selected case studies). Concrete problems were detected with respect to, for example, discoverability (Weber and Jones 2011), affordances (Roth and Harrower 2008) and error recovery (Davies and Medyckyj-Scott 1996). These issues could not be attributed to a specific group of systems of the same type or with the same target users. Instead, reviews diagnosed a general lack of user guidance for many GIS.

Further investigation of the underlying reasons is required to develop appropriate solutions for GIS that can be applied as common heuristics. Besides appropriate terminology and error messages (Nielsen 1994), there is potential for more GIS-specific principles, such as for interactions that require users to switch between traditional UI elements and the map element. As GIS is also being applied more in mobile device and application areas, measures such as the context of use and the limited screen size must be considered as well. In the end, concepts must be developed to avoid what makes GIS hard to use (Traynor and Williams 1995).

4.6.2. *Methods for analyzing GIS interactions and workflows*

Usability findings related to user strategies were detected in 15.4% of all selected case studies. This type of holistic investigation of user sessions revealed more in-depth findings (Roth and MacEachren 2016) compared to, for example, the analysis of single map interactions. For instance, the analysis of users' reactions to error messages could contribute to improved error recovery. Harvesting of fine granular interaction data might also be beneficial as some issues of GIS cannot be easily detected via user observations. Frequently, GIS are complex, and functions are placed on top of the geographic content to allow for dynamic search, selection and manipulation of geospatial data. Using these functions often requires a combination of map interaction and a traditional UI. For example, mouse track data could facilitate the understanding of how GIS users manipulate geospatial data.

4.6.3. *GIS use over longer periods*

Although some of the selected case studies applied UCD processes, usability evaluations were exclusively conducted at a single point in the lifecycle. Snapshots of usability were either taken during development (Roth *et al*. 2015) or in production (Unrau *et al*. 2017). Nielsen (1994) emphasizes that 'Using the system changes the users, and as they change they will use the system in new ways'. This includes the learning process of novice users as well as the requirements of an expert user (e.g. shortcuts). Therefore, it is essential to observe and evaluate use over a longer period of time. Automated data collection approaches may facilitate this endeavor (Lazar *et al*. 2017) because extensive 'manual' usability evaluations can consume many resources.

5. Conclusion

This article presents a thorough literature review of existing work in the field of usability evaluations for GIS. Following a rigorous and replicable approach, we collected more then 6000 publications from academic databases and key outlets in GIS and HCI research. Using a set of well-defined inclusion criteria, we extracted a set of 39 papers that we analyzed in depth along a number of dimensions. The results provide a detailed review of the current state of the art in usability evaluation of GIS.

The main contributions of our work are three-fold. First, we provided an overview of usability case studies, of what methods were applied when evaluating GIS, and of common findings that were reported in the literature. Key findings in this context are that multiple evaluation studies reported a general lack of user guidance in GIS; also, results showed overall low efficiency for using geospatial tools. In terms of target users, GIS for casual use were more often evaluated than systems for domain use. For the evaluation themselves, lab-based studies with user tests were identified most often, whereas inspections or field-based studies were relatively rare. The evaluation of user strategies in GIS applications was reported only rarely. Furthermore, the majority of studies was conducted in labs and with participants that did not represent actual target users.

From these results, we discussed and extracted key challenges as well as opportunities for future usability evaluations in this domain. We recommend the development of tailored heuristics and usability guidelines for GIS, as complex tasks often require a complex UI that should be aligned to web design standards. In addition, the methods for evaluating GIS use should be capable of capturing data for these interactions and workflows. Finally, conducting longitudinal studies for determining learnability aspects and requirements of expert users could represent an opportunity for additional evaluations.

Notes

1. https://ec.europa.eu/digital-single-market/en/legislative-measures (accessed 28 November 2018).
2. https://www.gov.uk/government/publications/government-transformation-strategy-2017-to-2020 (accessed 28 November 2018).

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

René Unrau is a PhD student at the Institute of Geoinformatics at the University of Münster, Germany. He received his Masters in Geoinformatics from the University of Münster in 2016 and works as a software engineer for web-based geographic information systems (WebGIS) at con terra in Münster. His research interests are usability evaluation methods for WebGIS and Visual Analytics. He is especially interested in remote and asynchronous evaluation approaches as well as tools for visually investigating map interaction data.

Christian Kray is a professor in Geoinformatics at the Institute of Geoinformatics (ifgi) in Münster, Germany. He leads the Situated Computing Lab (sitcom) at ifgi, which focuses on research into location-based and situated systems, evaluation methods for mobile systems and innovative interaction with spatial concepts. Before coming to Münster, Chris was a senior lecturer at the School of Computing Science at Newcastle University, UK, and a postdoctoral researcher at InfoLab 21, Lancaster, UK. He holds a PhD in Computer Science from Saarland University, Germany.

ORCID

René Unrau http://orcid.org/0000-0003-2536-9747
Christian Kray http://orcid.org/0000-0002-4199-8976

References

Alaçam, Ö. and Dalc, M., 2009. A usability study of webmaps with eye tracking tool: the effects of iconic representation of information. *International Conference on Human-Computer Interaction*. Berlin: Springer.

Brooke, J., 1996. Sus-a quick and dirty usability scale. *Usability Evaluation in Industry*, 189 (194), 4–7.

Caron, C., *et al*., 2008. Giscience journals ranking and evaluation: an international delphi study. *Transactions in GIS*, 12 (3), 293–321. doi:10.1111/tgis.2008.12.issue-3

Cinnamon, J., *et al*., 2009. Evaluating web-based static, animated and interactive maps for injury prevention. *Geospatial Health*, 4 (1), 3–16. doi:10.4081/gh.2009.206

Çöltekin, A., *et al*., 2009. Evaluating the effectiveness of interactive map interface designs: a case study integrating usability metrics with eye-movement analysis. *Cartography and Geographic Information Science*, 36 (1), 5–17. doi:10.1559/152304009787340197

Davies, C. and Medyckyj-Scott, D., 1996. Gis users observed. *International Journal of Geographical Information Systems*, 10 (4), 363–384. doi:10.1080/02693799608902085

Davies, C. and Medyckyj-Scott, D., 1994. Gis usability: recommendations based on the user's view. *International Journal of Geographical Information Science*, 8 (2), 175–189. doi:10.1080/02693799408901993

Düren, M. and Bartoschek, T., 2013. Usability engineering for educational web GIS. *In*: *The 16th AGILE international conference on geographic information science,* Leuven, Belgium.

Edsall, R.M., MacEachren, A.M., and Pickle, L., 2001. Case study: design and assessment of an enhanced geographic information system for exploration of multivariate health statistics. *Infovis*, 1, 159.

Fechner, T., Wilhelm, D., and Kray, C., 2015. Ethermap: real-time collaborative map editing. *In*: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, Seoul, Korea ACM, 3583–3592.

Gurevitch, J., *et al*., 2018. Meta-analysis and the science of research synthesis. *Nature*, 555 (7695), 175. doi:10.1038/nature25753

Haklay, M. and Jones, C.E., 2008. Why your boss should buy you a larger monitor?. *GIS Professional*.

Haklay, M. and Tobón, C., 2003. Usability evaluation and ppgis: towards a user-centred design approach. *International Journal of Geographical Information Science*, 17 (6), 577–592. doi:10.1080/1365881031000114107

Haklay, M. and Zafiri, A., 2008. Usability engineering for gis: learning from a screenshot. *The Cartographic Journal*, 45 (2), 87–97. doi:10.1179/174327708X305085

Hart, S.G., 2006. Nasa-task load index (nasa-tlx); 20 years later. *In*: *Proceedings of the human factors and ergonomics society annual meeting*. Los Angeles, CA: Sage Publications Sage CA, Vol. 50, 904–908. https://doi.org/10.1177%2F154193120605000909

Hoover, J.H., *et al*., 2014. Designing and evaluating a groundwater quality internet gis. *Applied Geography*, 53, 55–65. doi:10.1016/j.apgeog.2014.06.005

Hossain, D. and Masud, M.M., and others, 2009. Evaluating software usability of geographic information system. *International Journal of the Computer, the Internet and Management*, 17, 37–54.

Ingensand, J., 2006. Developing web-gis applications according to hci guidelines: the viti-vaud project. *In*: *Geographic hypermedia*. Berlin: Springer, 409–420.

Ingensand, J. and Golay, F., 2010. Task-oriented usability evaluation of a WebGIS for a real-world community. *URISA Journal*, 22 (1), 41.

Ingensand, J. and Golay, F., 2011. Remote-evaluation of user interaction with WebGIS. *In*: *Web and wireless geographical information systems*, Kyoto, Japan, 188–202.

ISO, SFSEN., 1998. 9241-11 (1998). *Ergonomic requirements for office work with visual display terminals (VDTs)–Part II guidance on usability*.

Keim, D., *et al*., 2008. Visual analytics: definition, process, and challenges. *In*: *Information visualization*. Berlin: Springer, 154–175.

Kiefer, P., *et al*., 2017. Controllability matters: the user experience of adaptive maps. *GeoInformatica*, 21 (3), 619–641. doi:10.1007/s10707-016-0282-x

Komarkova, J., Jakoubek, K., and Hub, M., 2009a. Usability evaluation of web-based gis: case study. *In*: *Proceedings of the 11th international conference on information integration and web-based applications & services*, Kuala Lumpur, Malaysia, ACM, 557–561.

Komarkova, J., Jedlicka, M., and Hub, M., 2009b Web-based geographic information systems and their usability. *In*: *Proceedings of the 9th WSEAS international conference on applied computer science*. World Scientific and Engineering Academy and Society (WSEAS), Genova, Italy, 97–102.

Komarkova, J., *et al*., 2007. Usability of geoweb sites: case study of czech regional authorities web sites. *BIS'07 Proceedings of the 10th international conference on Business information systems,* Poznan, Poland, 411–423.

Komarkova, J., *et al*., 2017. Usability evaluation of web-based gis by means of a model. *In*: *Information and digital technologies (IDT), 2017 international conference on*. IEEE, : Zilina, Slovakia, 191–197.

Komarkova, J., *et al*., 2010. Problems in usability of web-based gis. *Information Systems Journal*, 9 (2), 419–425.

Kong, N., Zhang, T., and Stonebraker, I., 2014. Common metrics for web-based mapping applications in academic libraries. *Online Information Review*, 38 (7), 918–935. doi:10.1108/OIR-06-2014-0140

Lallemand, C. and Koenig, V., 2017. Lab testing beyond usability: challenges and recommendations for assessing user experiences. *Journal of Usability Studies*, 12 (3), 133–154.

Lazar, J., Feng, J.H., and Hochheiser, H., 2017. *Research methods in human-computer interaction*. Hoboken, NJ: John Wiley & Sons Ltd.

Lobo, M.-J., Pietriga, E., and Appert, C., 2015. An evaluation of interactive map comparison techniques. *In*: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, Seoul, Korea, ACM, 3573–3582.

Manson, S.M., *et al*., 2012. Using eye-tracking and mouse metrics to test usability of web mapping navigation. *Cartography and Geographic Information Science*, 39 (1), 48–60. doi:10.1559/1523040639148

May, J. and Gamble, T., 2014. Collocating interface objects: zooming into maps. *In*: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, Toronto, Ontario, ACM, 2085–2094.

Newman, G., *et al*., 2010. User-friendly web mapping: lessons from a citizen science website. *International Journal of Geographical Information Science*, 24 (12), 1851–1869. doi:10.1080/13658816.2010.490532

Nielsen, C.M., *et al*., 2006. It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. *In*: *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, Oslo, Norway, ACM, 272–280.
Nielsen, J., 1994. *Usability engineering*. San Francisco, CA: Elsevier.
Nivala, A.-M., Brewster, S., and Sarjakoski, T.L., 2008. Usability evaluation of web mapping sites. *The Cartographic Journal*, 45 (2), 129–138. doi:10.1179/174327708X305120
Poplin, A., 2015. How user-friendly are online interactive maps? Survey based on experiments with heterogeneous users. *Cartography and Geographic Information Science*, 42 (4), 358–376. doi:10.1080/15230406.2014.991427
Reichenbacher, T., 2003. Adaptive methods for mobile cartography. *In*: *Proceedings of the 21st international cartographic conference ICC: cartographic renaissance*, Durban, South Africa, 1311–1321.
Roth, R.E. and Harrower, M., 2008. Addressing map interface usability: learning from the lakeshore nature preserve interactive map. *Cartographic Perspectives*, 60, 46–66. doi:10.14714/CP60.231
Roth, R.E. and MacEachren, A.M., 2016. Geovisual analytics and the science of interaction: an empirical interaction study. *Cartography and Geographic Information Science*, 43 (1), 30–54. doi:10.1080/15230406.2015.1021714
Roth, R.E., Ross, K.S., and MacEachren, A.M., 2015. User-centered design for interactive maps: a case study in crime analysis. *ISPRS International Journal of Geo-Information*, 4 (1), 262–301. doi:10.3390/ijgi4010262
Schimiguel, J., *et al*., 2004. Inspecting user interface quality in web gis applications. *Conference Proceedings: VII Brazilian Symposium on Geoinformatics, 2004, Campos do Jordão, São Paulo, Brazil*.
Skarlatidou, A. and Haklay, M., 2006. Public web mapping: preliminary usability evaluation. *Proceedings: GIS Research UK,* Nottingham.
Steinmann, R., Krek, A., and Blaschke, T., 2004. Analysis of online public participatory gis applications with respect to the differences between the US and Europe. *In*: *Proceedings of the urban data management symposium04,* Chioggia-Venice, Italy.
Traynor, C. and Williams, M.G., 1995. Why are geographic information systems hard to use? *In*: *Conference companion on human factors in computing systems*, Denver, Colorado, ACM, 288–289.
Tsou, M.-H. and Curran, J.M., 2008. User-centered design approaches for web mapping applications: a case study with usgs hydrological data in the united states. *In*: *International perspectives on maps and the internet*. Berlin: Springer, 301–321.
Unrau, R., Ostkamp, M., and Kray, C., 2017. An approach for harvesting, visualizing, and analyzing WebGIS sessions to identify usability issues. In: *EICS '17: proceedings of the ACM SIGCHI symposium on engineering interactive computing systems*, Lisbon, Portugal, 33–38.
Unwin, D.J., 2005. Fiddling on a different planet? *Geoforum*, 36 (6), 681–684. doi:10.1016/j.geoforum.2005.04.003
Weber, P. and Jones, C.E., 2011. Usability of editors: what to improve?. *In*: Schmidt M and Gartner G, eds. *Proceedings of the 1st European state of the map conference,*Wien, 14–33.
You, M., *et al*., 2007. A usability evaluation of web map zoom and pan functions. *International Journal of Design*, 1 (1): 15–25.

Appendices

Table A1. Selected usability evaluation studies. The last column refers to testing and/or inquiry participants. Systems without a value in this column were inspected by experts.

Author and year	Sys. type	Sys. users	Study type	#Part.
Alaçam Dalcı (2009)	Web Mapping	Casual	Lab	26
Cinnamon *et al.* (2009)	WebGIS	Domain	Field	8
Çöltekin *et al.* (2009)	WebGIS	Casual	Lab	30
Davies and Medyckyj-Scott (1994)	GIS	Domain	Field	159
Davies and Medyckyj-Scott (1996)	GIS	Domain	Field	21
Düren and Bartoschek (2013)	WebGIS	Casual	Lab	9
Edsall *et al.* (2001)	GIS	Domain	Lab	37
Fechner *et al.* (2015)	WebGIS	Casual	Lab	39
Haklay and Tobon (2003)	WebGIS	Casual	Lab	9
Haklay and Zafiri (2008)	GIS	Domain	Field	35
Haklay and Jones (2008)	GIS	Domain	Lab	2
Hoover *et al.* (2014)	WebGIS	Domain	Lab	30
Hossain and Masud (2009)	GIS	Domain	Lab	4
Ingensand (2006)	WebGIS	Casual	Lab	4
Ingensand and Golay (2010)	WebGIS	Casual	Lab	20
Ingensand and Golay (2011)	WebGIS	Casual	Field	331
Kiefer *et al.* (2017)	Web Mapping	Casual	Lab	24
Komarkova *et al.* (2007)	WebGIS	Casual	Lab	1
Komarkova *et al.* (2009b)	WebGIS	Casual	Lab	10
Komarkova *et al.* (2009a)	WebGIS	Casual	Lab	165
Komarkova *et al.* (2010)	WebGIS	Casual	Lab	
Komarkova *et al.* (2017)	WebGIS	Casual	Lab	12
Kong *et al.* (2014)	WebGIS	Casual	Lab	17
Lobo *et al.* (2015)	WebGIS	Casual	Lab	15
Manson *et al.* (2012)	Web Mapping	Casual	Lab	64
May and Gamble (2014)	Web Mapping	Casual	Lab	63
Newman *et al.* (2010)	WebGIS	Casual	Lab	16
Nivala *et al.* (2008)	Web Mapping	Casual	Lab	8
Poplin (2015)	WebGIS	Casual	Lab	26
Roth and Harrower (2008)	Web Mapping	Casual	Field	8
Roth *et al.* (2015)	WebGIS	Domain	Field	20
Roth and MacEachren (2016)	WebGIS	Domain	Lab	10
Schimiguel *et al.* (2004)	WebGIS	Domain	Lab	
Skarlatidou and Haklay (2006)	Web Mapping	Casual	Lab	30
Steinmann *et al.* (2004)	WebGIS	Casual	Lab	
Tsou and Curran (2008)	WebGIS	Domain	Field	8
Unrau *et al.* (2017)	WebGIS	Domain	Field	152
Weber and Jones (2011)	WebGIS	Casual	Lab	10
You *et al.* (2007)	Web Mapping	Casual	Lab	96