



Methods of Promoting Learning and Data Quality in Citizen and Community Science

*Christothea Herodotou**, *Eileen Scanlon* and *Mike Sharples*

Institute of Educational Technology, The Open University, Milton Keynes, United Kingdom

Citizen science or community science (CS) programmes are engaging volunteers in specific stages of the scientific research, mostly data collection and processing. They are less likely to have an explicit objective to support and promote volunteers' learning. In response to that, "citizen inquiry" emphasizes citizens' learning and community education, by viewing CS as an opportunity to educate the general public in thinking and acting scientifically. In citizen inquiry, citizens can take part in all the stages of the scientific research, from setting up an inquiry of personal interest, to deciding on the methods of data collection, analysis, and reporting. To ensure data quality when non-professionals design their own or take part in existing investigations, we have designed a bespoke online technological solution, the nQuire platform (nquire.org.uk), with support from the Open University/BBC partnership. nQuire scaffolds the design of high quality scientific inquiries through an authoring functionality and a process of data quality review by experts. In this paper, we detail how nQuire can support data quality assurance and control. We present case studies of how data quality was managed in two projects: "Heatwaves: Are you coping?" and "Pollinator Watch."

Keywords: data quality, volunteers, learning, online community and citizen science, temperature data, biodiversity data

OPEN ACCESS

Edited by:

Sven Schade,
European Commission, Joint
Research Centre (JRC), Italy

Reviewed by:

Baiba Pruse,
Ca' Foscari University of Venice, Italy
Nuria Castell,
Norwegian Institute for Air
Research, Norway

*Correspondence:

Christothea Herodotou
christothea.herodotou@open.ac.uk

Specialty section:

This article was submitted to
Climate Risk Management,
a section of the journal
Frontiers in Climate

Received: 06 October 2020

Accepted: 10 May 2021

Published: 31 May 2021

Citation:

Herodotou C, Scanlon E and
Sharples M (2021) Methods of
Promoting Learning and Data Quality
in Citizen and Community Science.
Front. Clim. 3:614567.
doi: 10.3389/fclim.2021.614567

INTRODUCTION

Citizen science or Community science (CS) is a research paradigm in which members of the public or citizens, often referred to as volunteers or amateurs, take part in scientific activities initiated by scientists and/or community members depending on the CS form. CS can be initiated by scientists in contributory projects (Shirk et al., 2012) where the public primary contributes data, or be the result of a collaboration between members of the public and scientists where the former are involved in most or all aspects of research (co-created projects, see Shirk et al., 2012). These activities are described as "communal experiences" that bring the community together to examine and understand a topic of interest (Audubon Center, 2018). CS has been viewed as a distinct field of inquiry which can engage volunteers with "relevant, authentic, and constantly changing dimensions of primary research" (Jordan et al., 2015, p.211). It can support and extend research in any discipline including social, natural and physical sciences, such as helping scientists identify species (Herodotou et al., 2017). There are subtle differences between a CS activity and more traditional ways of engaging people with research such as a survey or workshop and these should be considered before naming an activity as CS (ECSA, 2020). Documents defining CS and its characteristics such as the "Ten Principles of CS" by ECSA (2015) could help with making this distinction.

The degree of volunteers' engagement with CS activities varies. For example, in some projects, volunteers decide what should be studied, while in others they contribute to specific aspects of the scientific method such as collecting or processing data (Shirk et al., 2012; Haklay, 2013). Most projects engage citizens in the processes of data collection (Hecker et al., 2018), such as making observations of biodiversity, or data processing such as transcribing specimens. Recently, there has been a shift from scientist-led approaches to CS to a more active engagement of the public in scientific activities that are not restricted to processes of data collection and analysis (König, 2017; Herodotou et al., 2018). In particular, the importance of devising personally-meaningful investigations, by having citizens set their own research agendas that match their needs and interests has been highlighted (Anastopoulou et al., 2012). Efforts are also made to expand the application of CS across disciplines, including for example social sciences (e.g., Dunn and Hedges, 2018), while ways to engage diverse demographics with CS activities such as young people are explored (Herodotou et al., 2020).

CS projects are designed with the aim to solve problems or improve science in ways that would not have been possible without the support of volunteers. Yet, an intentional integration of learning objectives for volunteers is less likely to be found in the design of CS programmes, nor the possibility to participate in all the stages of the research process and contribute to decisions and outputs (Edwards, 2015; Herodotou et al., 2020). Assessment of volunteers' learning in CS projects is showing promising outcomes such as educating themselves in scientific thinking and how science works, appreciating nature and contributing to science initiatives (Freitag and Pfeffer, 2013; Aristeidou and Herodotou, 2020). CS could democratize research by allowing citizens to "take agency in the research process" and by directing research toward solving prominent societal problems such as enabling sustainability transitions toward, for example, renewable energy sources and sustainable agriculture (Sauermaun et al., 2020). Despite significant benefits, CS programmes designed with an explicit focus on citizens and their growth and development remain scarce; such projects could involve citizens in the problem identification and framing of sustainable solutions, aligning policy agendas with the interests of the public, contribute own socio-political understanding of a topic, and help generate solutions and behavior change (Sauermaun et al., 2020). CS benefits are often a "by-product" of citizens' engagement with scientific activities. It still remains to design and assess CS projects that have explicit objectives to support citizens' learning and agency, while also produce quality datasets that can be used to inform research and policy.

Citizen or Community Inquiry

Citizen or community inquiry (CI) is an innovative approach to inquiry learning located at the intersection between "citizen or community science" and "inquiry-based learning" (e.g., Quintana et al., 2004). It refers to the distributed participation of members of the public in joining and initiating inquiry-led scientific investigations. Unlike the majority of CS projects that engage citizens in data collection activities, it aims to engage citizens in all the stages of the scientific research, from setting up personally

meaningful projects to collecting and analyzing data (Herodotou et al., 2017). Specifically,

"It fuses the creative knowledge building of inquiry learning with the mass collaborative participation exemplified by citizen science, changing the consumer relationship that most people have with research to one of active engagement" (Sharples et al., 2013).

Citizen or community inquiry emphasizes the intentional integration of inquiry-specific learning objectives into the design of community science activities. It aligns to a degree with Wiggins and Crowston (2011) categorization of education projects in which "education and outreach" are the primary goals of CS, as opposed to, for example, investigation projects the focus of which is to achieve certain scientific goals. What is distinct with "community inquiry" is the focus on a specific set of learning objectives that enable the development of inquiry skills. These learning objectives could be described using Bloom's updated taxonomy for teaching, learning, and assessment (Anderson et al., 2001) and categorized into (a) remember (recognizing/identifying, recalling), (b) understand (interpreting, exemplifying, classifying, summarizing, inferring, comparing, explaining), (c) apply (executing, implementing), (d) analyse (differentiating, organizing, attributing), (e) evaluate (checking, critiquing), and (f) create (hypothesizing, designing, producing). These learning objectives could be grouped into four major knowledge dimensions: (1) factual knowledge—knowledge of terminology, specific details and elements, (2) conceptual knowledge—knowledge of classifications, principles, theories, (3) procedural knowledge—knowledge of subject-specific skills, techniques, criteria for deciding when to use specific techniques, and (4) metacognitive knowledge- self-knowledge in relation to a subject matter.

Community inquiry could initiate with engaging volunteers in discussions about their own experiences and observations in relation to a topic of interest (remember/call). These discussions could result in brainstorming and elaborating on specific research questions that could be answered through one or more CS projects (understand). The next step is for volunteers to define the research method for collecting and analyzing data such as types of data collected, questions asked, methods of data analysis, and collect the actual data (apply). Following data collection, volunteers are involved in the process of data analysis and interpretation (analyse) during which they sort collected data, compare and contrast findings. This step should be followed by a process of evaluation of the findings (evaluate) in terms of answering the research question, relating findings to existing studies, and demonstrating new understandings. The final step is to promote impact by finding ways to share findings (reporting, publications, social media etc) with different audiences (create).

The community inquiry paradigm shifts the emphasis of scientific inquiry from scientists to the general public, by having non-professionals (of any age and level of experience) determine their own research agenda and devise their own science investigations underpinned by a model of scientific inquiry. Such an approach may sound visionary and particularly challenging, given that some volunteers may not have the necessary skills

and training to take part or initiate scientific activities. As explained by Gura (2013), it is difficult to ascertain the quality of the data when you do not know if data were collected by a botany professor or “a pure amateur with an untrained eye” looking at wildflowers. Thus, volunteers’ involvement may bias or undermine the scientific process and the production of valid and reliable outcomes. In the next section, we present an overview of how scientists could support volunteers’ participation and learning in CS in ways that can result in high quality datasets. We then show how an online technological solution, nQuire, can be combined with a review and approval process to scaffold the design of community inquiry investigations.

Data Quality in CS

There is often skepticism as to whether data generated by or collected from volunteers are accurate enough to inform future research and policy initiatives. Although considerably large in quantity, such datasets are often heterogeneous and hard to scale up to the population (Kelling et al., 2015). The accuracy of collected data should be judged project-by-project, considering several factors. In large-scale CS projects, high quality data can be achieved through a simple process of data collection and *post-hoc* data quality controls such as computational modeling, while in small-scale CS projects with complex data collection processes, expert-led training, and good quality materials can generate good quality science outcomes (Parrish et al., 2018). Kosmala et al. (2016) showed that a growing number of CS projects have managed to produce data with accuracy equal to, or in some cases, superior to that of scientists and proposed a set of strategies that when followed can enhance data accuracy. These refer to piloting and improving the design of a project, considering the level of difficulty of tasks, making systematic the tasks and data entry (e.g., selection from a predefined list of options), ensuring calibration of equipment used, recording of metadata that may influence data capture such as contextual factors or volunteers’ characteristics, standardizing when data should be captured such as place, time, duration of recording, comparing findings to those of professionals in project reporting, allowing for sufficient data quantity and diversification (e.g., covering a range of geographic areas) as per the research questions, and storing data in a concise format to allow for interpretation and easy use. Adding simple instructions that model best practices, alongside technical enhancements that can support correct recording (e.g., status indicator of GPS availability), have also been shown to improve the measurement processes and reduce errors (Budde et al., 2017).

Technology becomes increasingly important in improving data accuracy. Certain technological features or solutions can help to overcome accuracy errors by, for example, systematizing methods of data collection and providing automated ways of giving feedback to volunteers about the quality of collected data. For instance, online social trust models were shown to effectively measure the trustworthiness of citizen generated data, while filtering was used to remove unreliable data (Hunter et al., 2013). A data quality measure framework has been proposed consisting of four steps: (a) identification of data quality dimensions, (b) application of data quality measures, (c) analyzing data and

identifying errors, and (d) implementing tools and approaches that can improve data quality (Hunter et al., 2013). Scientists should consider data quality assurance as part of the process of designing a CS project, by identifying possible quality problems and how the project could be adopted to accommodate these. For instance, they should define the minimum sample size and the sampling sites from which data should be collected, that best address project objectives (Weigelhofer and Pölz, 2016). In the next sections, we present how we considered data quality issues in the design of the nQuire platform and we detail data quality and control issues in two nQuire projects related to temperature and biodiversity datasets.

nQuire

Citizen or Community Inquiry has been operationalised in a bespoke, online solution, the nQuire platform (www.nquire.org.uk). nQuire has been designed with support from the BBC Tomorrow’s World initiative to explicitly scaffold citizen-led investigations and inquiry learning processes (Herodotou et al., 2014). At the moment, it hosts 39 investigations with contributions ranging from 100 to 230K. These investigations have been initiated by mainly scientists (from across different universities), organizations such as the the Royal Meteorological Society and the Young Foundation, and individuals with an interest in science or research. The nQuire functionality allows members of the public, from lay individuals to scientists, to set up their own investigations, discuss and negotiate their ideas, and reach a consensus. In particular, nQuire facilitates inquiry learning by enabling citizens to brainstorm and collect potential research questions through, for example, contributing to investigations where data are open to anyone to read, comment, and discuss (see Bloom’s objectives of remember and understand), define research methods and collect data in the form of texts, images, and sensor data (e.g., light, sound) through the use of an authoring functionality (apply), analyse and interpret collected data in the form of graphs and narrative that can be open to anyone to access (analyse), and write and share findings in the form of an interim or final report directly with participants (via email) or anyone visiting the platform through publishing a Pdf document on nQuire (evaluate and create). The platform supports two types of investigation: confidential missions and social missions (see **Figure 1**). In confidential missions, all data are private and only accessed by the mission author. In social missions, data are open for other people to view and discuss online. nQuire could potentially support the development of all five CS project types as proposed by Shirk et al. (2012) including contractual, contributory, collaborative, co-created and collegial projects. Yet, the vision behind nQuire is to enable the design of “extreme citizen science” projects (Haklay, 2013), or co-created and collegial projects (Shirk et al., 2012).

An example of a confidential mission is the “What’s your chronotype?” mission for people to explore their sleep patterns. The mission attracted 6,700 contributions. Participants entered some personal details such as age and gender, and then responded

The screenshot shows the nQuire website interface. At the top left is the nQuire logo with the tagline 'EXPLORE YOUR WORLD'. To the right are navigation links: Home (underlined), Discover, About, Register, and Sign in. The Open University logo is in the top right corner. Below the navigation are two mission cards. The first card, 'Heatwave: Are you coping?', features a silhouette of a person looking through binoculars against a sunset background. Below the title is the question 'Tell us how warm it is and how you are feeling' and a response count of 1,290. The second card, 'Remote teaching in Africa', features a photo of three children looking at a book. Below the title is the question 'What does remote teaching in Africa look like?' and response counts of 62, 145, and 32. At the bottom of the screenshot is a large purple heading: 'Join missions to explore your world'.

FIGURE 1 | A confidential and a social CS investigation on nQuire.

to a series of questions to uncover whether they are e.g., a “morning” or an “evening” person. Each participant received instant and personalized feedback based on their responses, immediately after they submitted their answers. An example of a social mission is “Remote teaching in Africa.” Teachers from Africa were asked to share the challenges they faced due to Covid-19 and ways they overcome these. Responses were public. This enabled participants to communicate and learn from colleagues who may be facing similar issues. The mission contributed original insights about the educational situation in Africa as affected by the pandemic and produced guidelines as to how to support learners and teachers in this context.

Missions can include a rich mix of elements, such as: sounds or images as prompts; the ability of participants to upload a picture or sensor data as a response; and a variety of response types, such as slider scales, dropdown lists, and Likert scales. The platform provides an authoring tool to create new missions by setting up a “big question,” an outline of the mission, adding a variety of response types, scoring each response category, and authoring customized feedback to the participant based on their scores. The mission can be divided into sections, with separate feedback from each section. All missions are checked by the nQuire research team at The Open University before they go live, to make sure they are safe, ethical, and legal. Owners of

the mission (for a confidential mission) or any user (for a social mission) can download data in spreadsheet format. Results from a mission can be published on the platform.

DATA QUALITY IN DESIGNING CS PROJECTS ON nQuire

A significant factor that could determine the quality of data collected is how a CS project is designed, in particular, what questions volunteers are asked to address and how well these are formulated, how methods of data collection are explained to them, what tools are available to enable accurate data collection, and what benefits there are for volunteers to motivate participation and quality of data collection. CS projects work better when they have defined and clear goals that are communicated to volunteers from the start, scientists with appropriate expertise are involved, the approach can be trialed and improved, participants understand the value of the project and the benefits they get out of it, and the quality of collected data can be measured (Tweddle et al., 2012). In technology-enhanced CS projects, issues of easy and accessible interface design, use of conventional language, easy registration, real time communication functionality, visualization of data

collected, and motivational features are amongst the issues to be considered (Skarlatidou et al., 2019). In particular, what can support citizen-centered scientific investigations are mobile affordances, scaffolding the process of scientific inquiry, enabling learning by doing, and being a part of a community (Herodotou et al., 2018). Collective motives, that is the importance one assigns to the collective goals of the project, as well as reputation were found to be positively related to the quality of contribution (Nov et al., 2014), suggesting that benefits to volunteers should be considered when designing a CS project and explicitly shared with them to promote high quality data collection. In nQuire, these benefits are found in mechanisms that promote learning from participation, detailed in section The Process of Review and Approval (second paragraph).

In the case of nQuire, we set in place two mechanisms that support the design of scientifically robust and ethical CS projects: an authoring tool that scaffolds the process of design and a process of review and approval that ensures data quality and control over published investigations.

The Authoring Tool

The authoring tool is free to access after registration with nQuire. It has been designed to scaffold the process of setting up a CS investigation from start to end. It enables any individual, scientist or citizen, to set up, manage, pilot, and launch an investigation of personal interest and share findings with volunteers. It is structured around four steps (see **Figure 2**):

Start

The first step asks for details about the investigation including its name, a big question, an outline of the mission, benefits to science, benefits to citizens, time to be completed, and a mission image. It prompts authors to think of these aspects through structured forms, accompanied by explanations and examples.

Build

In the second step, the author decides whether data from the mission will be confidential or open to the public to read and comment on, whether to score responses to specific questions

and provide immediate and personalized feedback to participants based on their responses, and whether they intend to show the contributions on a map that is visible to participants. In this step, authors can set up the questions of the investigation (see **Figure 3**). Participants can contribute data in the form of text, images, geolocated and sensor data. For example, they can be asked to collect sound or light data using the sensors of their personal mobile device. A graph of the recorded data will be captured and uploaded to the investigation. Authors can enable data visualizations for specific mission questions. This feature enables the production of graphs that can be shown live on the nQuire platform while the mission is running. They can be automatically updated, the more data are collected.

Enhance

The third step enables enhancements to the investigation, including the use of a customized splash screen, social sharing details, a custom consent form, and drafting feedback for volunteers. In the feedback form, authors are asked to provide details around the importance of the investigation, some interesting facts about it, and prompts for participants to read relevant information from specific websites or completing similar investigations on nQuire.

Finish

The fourth step enables piloting of the investigation through the provision of a unique URL that can be shared with a small group of participants for testing, and a request for launch, which triggers the process of review and approval (see next section).

The Process of Review and Approval

The process of review and approval is managed by academics in the nQuire team who are responsible to review the investigations and make requests for changes if they deem appropriate. The process entails checks on aspects that include:

- **Ethics:** Appropriate ethical approvals should be in place. In cases when authors have no access to an ethics board, the process can be managed by the Open University's ethics board.

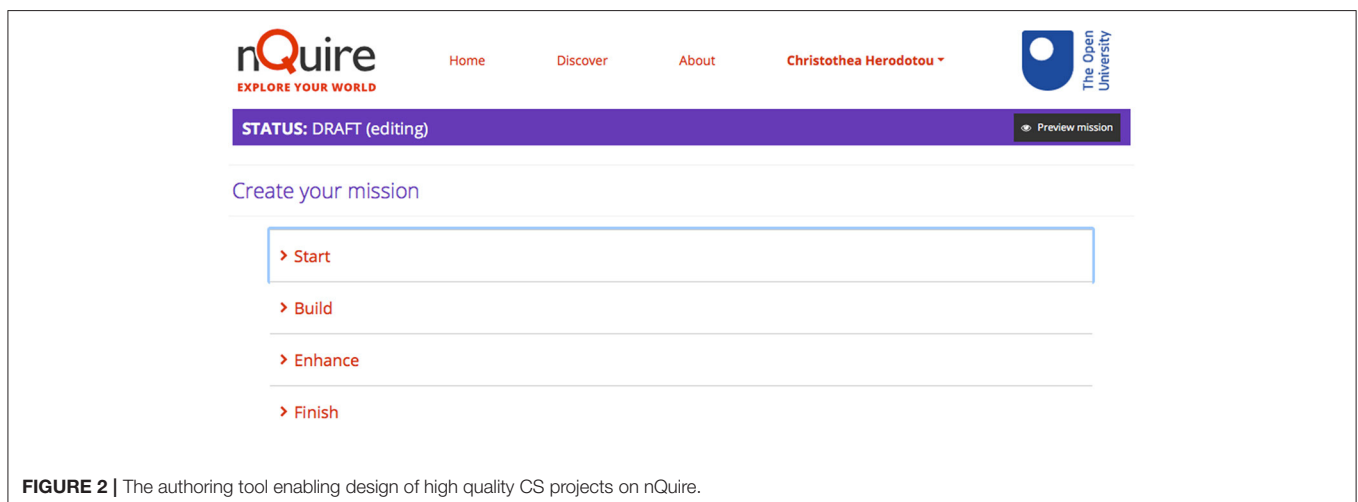
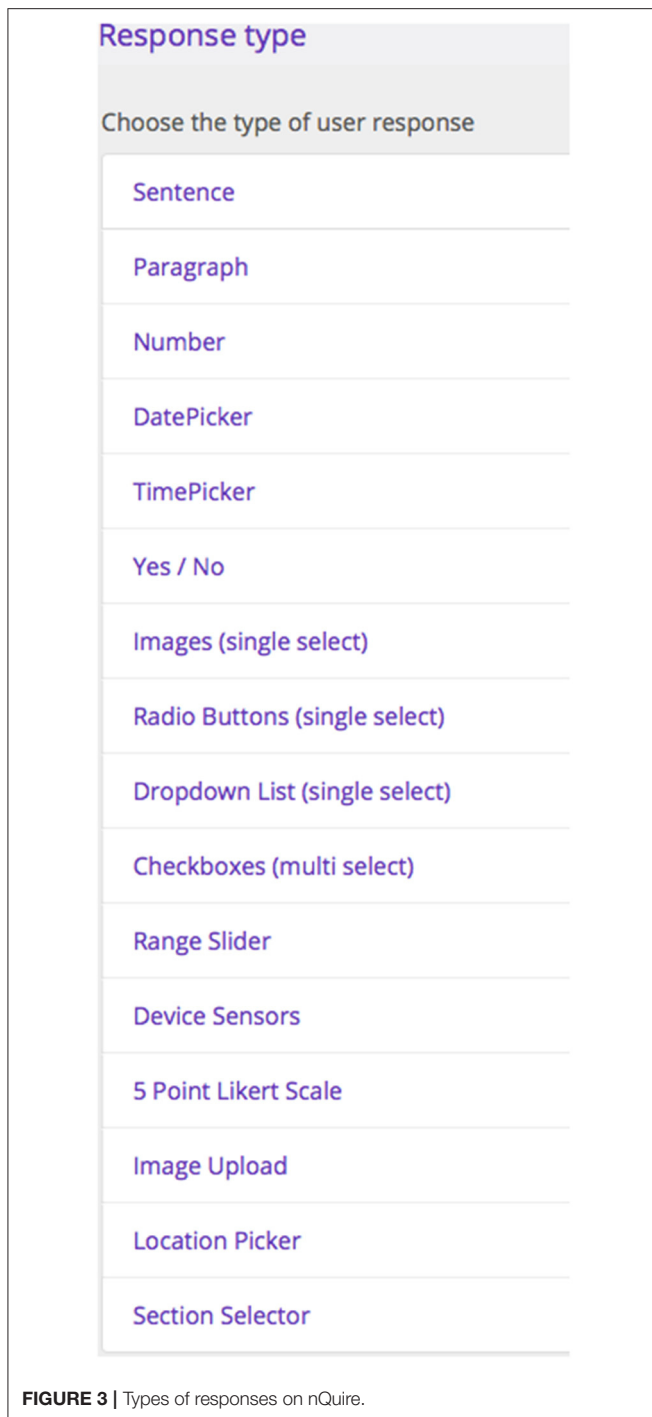


FIGURE 2 | The authoring tool enabling design of high quality CS projects on nQuire.



- Author: Details about the author/s of the investigation should be legitimate and accurate.
- Originality: The investigation should build on appropriate and relevant literature and have the potential to produce original insights contributing to authentic science. To accomplish that, we engage with scientists who have expertise in the area of investigation and who can assess its originality.

- Mission brief: The mission brief should state clearly what the investigation is about, the benefits from taking part and the value to science, what volunteers will be asked to do and for how long, and how data will be analyzed.
- Language: The language used should be accessible (understandable by most people) and not offensive in any ways.
- Questions: Any predefined responses to questions should be distinguishable and all possible options should be considered and provided.
- Copyrights: Images or any other material used should be clear of any copyright restrictions.
- Piloting: The investigation should be piloted with a few participants and feedback should be sought after prior to launching that can inform the design of questions and ensure that content and tasks are understandable and easily implemented.
- Publishing findings: When the investigation reaches a certain number of contributions, authors are asked to prepare and share an interim report with preliminary findings via nQuire. This report is free to access and ensures that citizens gain access to findings.

The authoring tool and the process of review and approval are two mechanisms that promote data quality controls, support the design of scientifically robust investigations and can foster learning for volunteers. In particular, in terms of the latter, learning is scaffolded through: (a) the provision of immediate and personalized feedback. For example, one of the nQuire investigations is asking volunteers to assess the degree to which a number of statements about Covid-19 are valid. Based on the responses, participants receive personalized feedback about the correctness of those statements and the degree to which they may be prone to misinformation in news. (b) clearly defined and explicit benefits to citizens stated in the mission brief and explained further in the feedback form, (c) for social missions, opening up, and visualizing data, allowing volunteers to comment on the data and communicate with others about the interpretation of outcomes, (d) engagement with hands-on activities for collecting data such as taking pictures of biodiversity or making temperature measurements or assessing the therapeutic impact of sounds from nature on well-being, and (e) the process of designing an investigation from scratch with support from the authoring tool and communication with the nQuire team that reviews and approves missions.

In the next sections, we detail how data quality issues were monitored in two nQuire investigations about heatwaves and pollinators. In particular, we comment on how the design of the investigations was informed by best practice in CS quality assurance, evidenced in the literature, and how we plan to analyse collected data considering for issues of data control.

QUALITY OF TEMPERATURE DATA

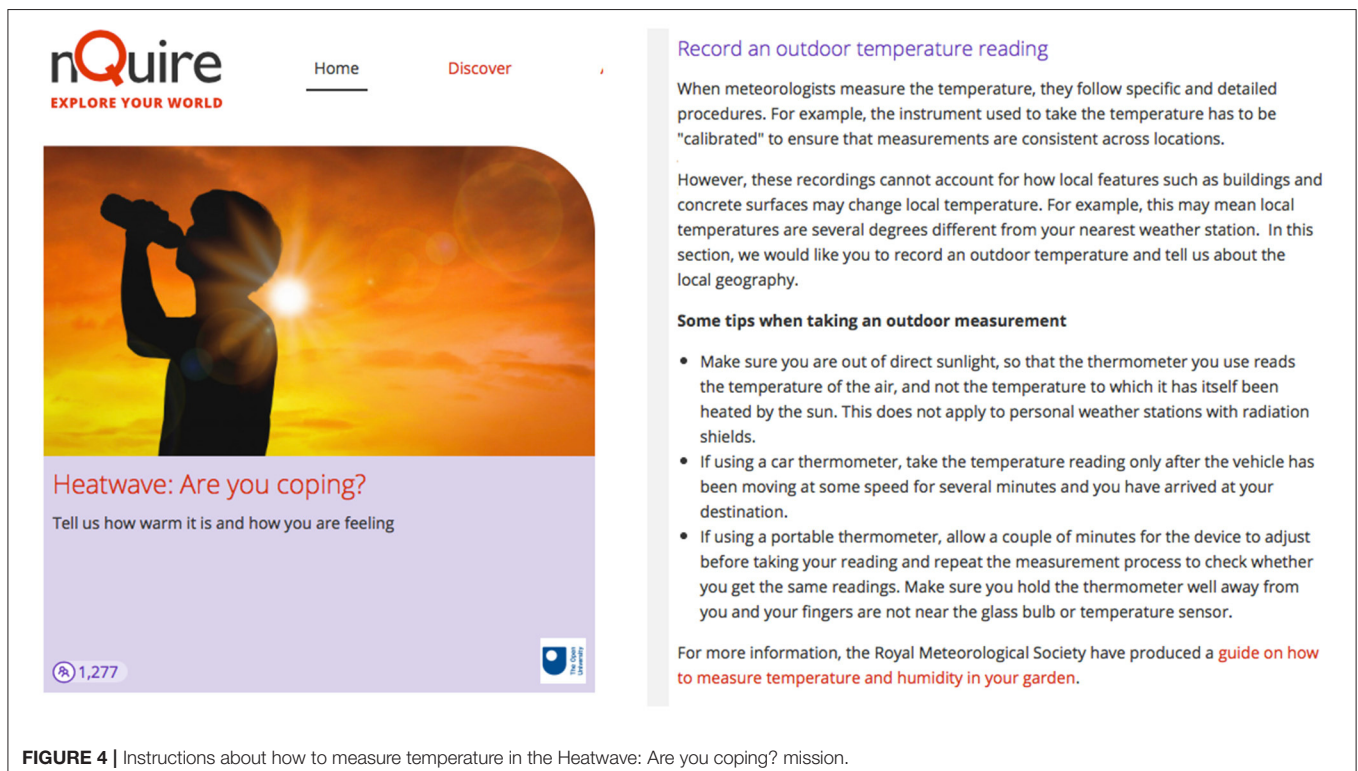
The “Heatwave: Are you coping?” investigation has been designed in collaboration with the Royal Meteorological Society and support from the BBC Weather (see <https://nquire.org.uk/>

mission/heatwave-are-you-coping/contribute). The mission was the outcome of a workshop with citizens and organizations interested in weather issues, which was organized by the Open University UK, as part of the UKRI funded project EduCS: EDUCating Citizens and organizations in Citizen Science methodologies. Workshop attendees were asked to brainstorm, vote, and rank ideas for research investigations they would like to design using nQuire. How comfortable people feel in extreme weather conditions was one of the two most popular investigations (alongside the impact of climate change). The investigation with more than 1,200 responses, was launched on the 7th August 2020, during which England experienced a heatwave and was ended in September 2020. The purpose of the mission was to explore how people's experiences of hot weather may differ depending on where they live and work, and how people are able to adapt their routines to heat. Citizens were asked to take their first temperature recording around 3–4 pm, when maximum daily temperatures are normally observed. The rationale behind the mission was to collect data about how different people are affected by extreme weather conditions and how working and living conditions could be improved. Results could, for example, help people plan for heatwaves in the future. In terms of the learning benefits for citizens, the mission was an opportunity to learn about what forecast temperatures mean in practice, how to make and record measurements, and how to increase personal comfort in a heatwave.

CS temperature measurements have the unique value of providing data about air temperature on scales smaller than those measured by the official meteorological service, and such

data could be possibly used in weather monitoring or even forecasting (Cornes et al., 2020). Yet, the quality of weather data collected is a major challenge and a source of bias, often related to possible overheating of the thermometer by, for example, not being shielded. This was an issue raised and discussed during the workshop, with weather scientists expressing concerns about the quality of data collected and whether amateur scientists could actually offer reliable recordings. To address these concerns, we first reviewed relevant literature about methods for improving weather data accuracy. Amongst the proposed approaches was to collect information about the instrument used and the conditions it is exposed to, that when considered could improve data accuracy (Meier et al., 2017) and consider for the use of a statistical correction—Generalized Additive Mixed Modeling (GAMM)—developed to improve accuracy when analyzing weather data (Cornes et al., 2020).

Considering these insights, we devised a number of mechanisms that could help improve the data accuracy of weather measurements. In particular, (a) in the mission instructions, we included top tips about how to capture the temperature, especially tackling the issue of overheating of the instruments, written in an accessible language and avoiding technical terminology (see **Figure 4**), (b) the Royal Meteorological Society developed an online guide about how to capture temperature and humidity in gardens which was attached to the mission instructions, and which provided technical details about different devices and their accuracy, (c) in the mission questions, citizens were asked to report on contextual information such as the instrument they used



nQuire
EXPLORE YOUR WORLD

Home Discover

Heatwave: Are you coping?
Tell us how warm it is and how you are feeling

1,277

Record an outdoor temperature reading

When meteorologists measure the temperature, they follow specific and detailed procedures. For example, the instrument used to take the temperature has to be "calibrated" to ensure that measurements are consistent across locations.

However, these recordings cannot account for how local features such as buildings and concrete surfaces may change local temperature. For example, this may mean local temperatures are several degrees different from your nearest weather station. In this section, we would like you to record an outdoor temperature and tell us about the local geography.

Some tips when taking an outdoor measurement

- Make sure you are out of direct sunlight, so that the thermometer you use reads the temperature of the air, and not the temperature to which it has itself been heated by the sun. This does not apply to personal weather stations with radiation shields.
- If using a car thermometer, take the temperature reading only after the vehicle has been moving at some speed for several minutes and you have arrived at your destination.
- If using a portable thermometer, allow a couple of minutes for the device to adjust before taking your reading and repeat the measurement process to check whether you get the same readings. Make sure you hold the thermometer well away from you and your fingers are not near the glass bulb or temperature sensor.

For more information, the Royal Meteorological Society have produced a [guide on how to measure temperature and humidity in your garden](#).

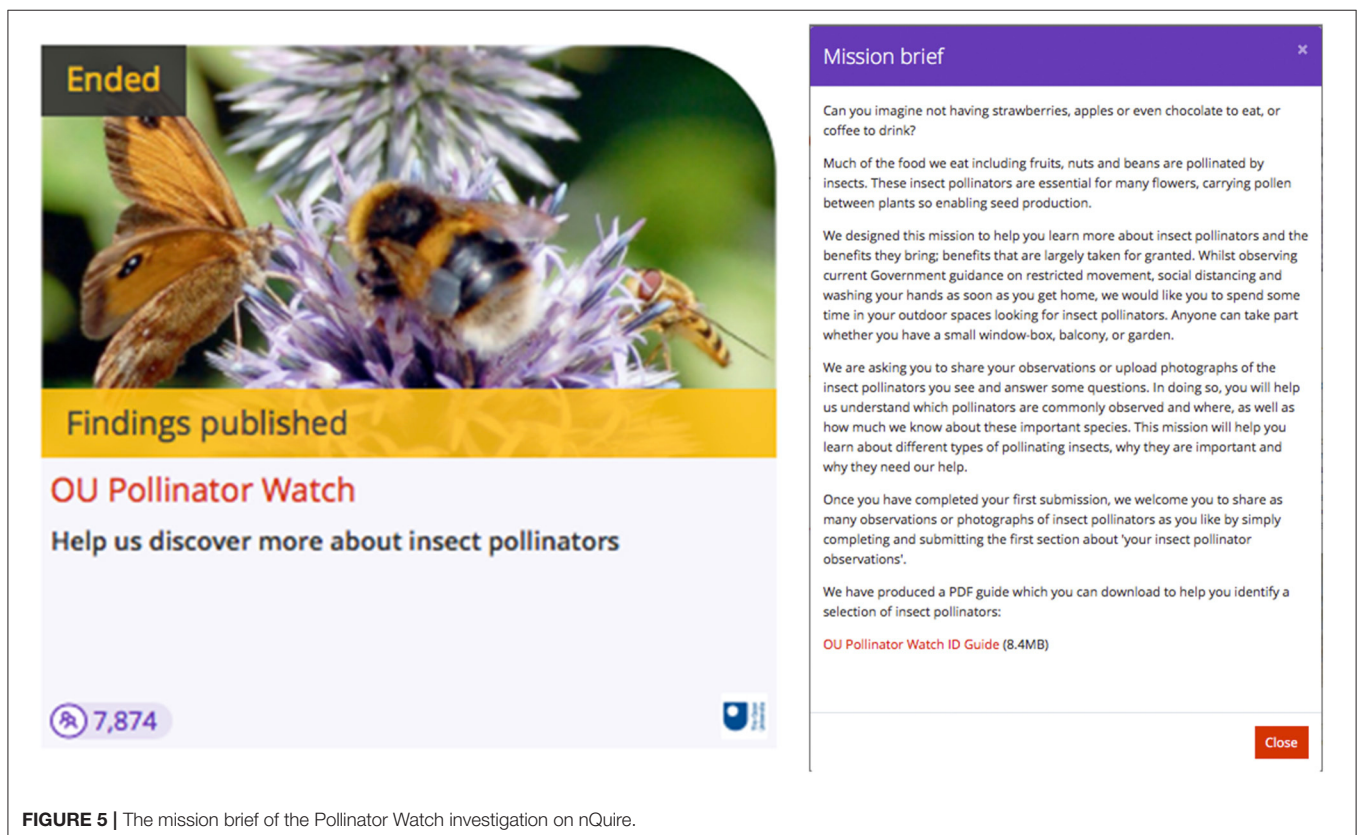
FIGURE 4 | Instructions about how to measure temperature in the Heatwave: Are you coping? mission.

to collect data, distance from nearest buildings, area density, relative humidity and their clothing. These data could inform data analysis and be considered in assessing the accuracy of the data collected. (d) During the process of data analysis we plan to undertake a number of checks to enhance the quality of collected data including data cleaning (that is removing data points that are not valid due to mistakes during data collection or reporting), and removing duplicate cases or extreme cases (“outliers”) by plotting the data and inspecting for points that are far outside the majority of contributions. We will draw from existing studies such as Li et al. (2020) to get the best possible outcome. (e) We plan to compare data points reported in the same geographic location (postcode) and identify the degree of agreement amongst them. The more accurate the data, the more likely those points will overlap. This is similar to the approach adopted in other CS platforms such as the iNaturalist where species identifications is graded based on whether community members agree and confirm the given identification, and (f) we will compare citizens’ recordings to the official weather data for a specific area and identify the degree of agreement or distance between the recordings.

QUALITY OF BIODIVERSITY DATA

The mission “Pollinator Watch” was designed by the Open University (OU) and promoted by the BBC 2 Springwatch series in 2020, attracting more than 7,800 contributions. The OU has

a long lasting collaboration with the BBC for over 40 years for the production of television and radio series. As part of it, in 2019 it designed and launched, with support from the British Trust of Ornithology, the Gardenwatch mission attracting more than 200K contributions. Following this successful collaboration, the nQuire team was asked to design a mission for the 2020 Springwatch series. To identify and set up a relevant mission, a team of scientists from the OU with expertise in biodiversity was brought together. Insect pollinators was the chosen topic as pollinators are essential for many flowers, carrying pollen between plants and enabling seed production, and are under extinction due to several threats such as the destruction of wild habitats and pesticides (e.g., <https://bit.ly/3g7ww1Z>). The mission could help citizens to learn about different types of pollinators, the benefits they bring and how they can be protected (see **Figure 5**). Citizens were asked to share their observations (upload photographs) of insect pollinators they see and answer some questions. These data would help scientists understand which pollinators are commonly observed and where, especially in the UK, as well as how much citizens know about these important species. In addition to that, scientists were interested in capturing any effects of the government’s restricted movement due to the Covid-19 lockdown in the UK, in particular whether citizens’ interactions with nature have changed. Preliminary data analysis has been published on the mission page and emailed to consented participants (see <https://nquire.org.uk/mission/oupollinatorwatch>).



Ended

Findings published

OU Pollinator Watch

Help us discover more about insect pollinators

7,874

Mission brief

Can you imagine not having strawberries, apples or even chocolate to eat, or coffee to drink?

Much of the food we eat including fruits, nuts and beans are pollinated by insects. These insect pollinators are essential for many flowers, carrying pollen between plants so enabling seed production.

We designed this mission to help you learn more about insect pollinators and the benefits they bring; benefits that are largely taken for granted. Whilst observing current Government guidance on restricted movement, social distancing and washing your hands as soon as you get home, we would like you to spend some time in your outdoor spaces looking for insect pollinators. Anyone can take part whether you have a small window-box, balcony, or garden.

We are asking you to share your observations or upload photographs of the insect pollinators you see and answer some questions. In doing so, you will help us understand which pollinators are commonly observed and where, as well as how much we know about these important species. This mission will help you learn about different types of pollinating insects, why they are important and why they need our help.

Once you have completed your first submission, we welcome you to share as many observations or photographs of insect pollinators as you like by simply completing and submitting the first section about 'your insect pollinator observations'.

We have produced a PDF guide which you can download to help you identify a selection of insect pollinators:

[OU Pollinator Watch ID Guide \(8.4MB\)](#)

Close

FIGURE 5 | The mission brief of the Pollinator Watch investigation on nQuire.

To enhance the quality of data collected, a visual guide with images and names of pollinators was designed and attached to the mission's instructions, encouraging people to use it when they are observing pollinators. This was a document with sample images from each category of pollinator and relevant information. A more dynamic guide such as an online interface with an open choice of filters, rather than directed filtering or direct visual comparison, could have improved further the accuracy of a species identification (Sharma et al., 2019). Also, citizens were asked to assess the degree to which they are confident that the identification they made is correct and state whether they normally observe and identify pollinators. These data will inform our approach to checking the correctness of the identifications; given the large size of contributions, we plan to sample a subset of them based on a set of criteria including previous experiences of identifying pollinators and confidence, general interest in environmental issues, age, and gender. Biodiversity scientists will then make their own identifications of the pollinators in these photographs, compare their identifications to those of citizens and ascertain the degree of correctness. An alternative option would be to upload or share the images with a biodiversity platform such as iNaturalist or iSpot. Such CS platforms make use of a combination of human and machine learning mechanisms for identifying species observations. For example, in iNaturalist, when uploading an observation you receive automatic recommendations as to what the species may be, generated by machine learning algorithms. In addition to that, a community grading system is in place which assesses data quality. An observation is verified when 2/3 of the community agrees on a taxon. In the case of nQuire, this was not feasible due to image copyright issues; participants were given the option to maintain or release rights from the images they took.

To enable easy access to the pollinators investigation, we allowed users to take part without registering with nQuire, yet noting that this would not allow them to access their data after submitting them. A registered user has access to their own private dashboard where they can see all the missions they took part in, their responses, as well as any missions they may have created or launched. Non-registered users raised a data control challenge as it became impossible to infer or identify duplicate responses in the dataset (and remove them) or identify whether a participant submitted more than one observation. We plan to treat non-registered users as a separate category and run the analysis considering for these limitations.

DISCUSSION AND CONCLUSIONS

In this paper, we have argued that the participation of volunteers in authentic scientific activities is a great learning opportunity that can promote development of scientific thinking skills and community inquiry. Such skills are particularly relevant to, not only those planning to follow a STEM career, but every individual, no matter what their career may be. Scientific thinking is a tool that can help with approaching and solving everyday problems. It is about how one looks at the world, questioning what others say, approaching problems in organized

and creative ways, learning to analyze why things went wrong and being open to new ideas that can change the way we think and act (The Royal Society, 2020). Community inquiry can raise awareness, drive behavior change and support transitions toward more sustainable ways of living in areas such as public health and environmental conversation (Sauermaun et al., 2020). It can enhance the sustainability of research projects and democratize science, especially when communities are invited to take part in all stages of the scientific research, projects are locally relevant and are addressing both the social and technical aspects of sustainability, and by eliminating tensions between traditional science and CS (Sauermaun et al., 2020; Froeling et al., 2021). The significance of developing community inquiry through participation in CS activities comes with a major challenge, that of producing high quality datasets that can be used to inform future research and policy initiatives (Kosmala et al., 2016; Parrish et al., 2018). Citizens are often not trained, or do not have the skills, to conduct scientific activities and thus their involvement is often faced with skepticism. To address this challenge, we detailed how nQuire, an online CS platform, and a process of review and approval by scientists can promote high quality data collection and help volunteers learn from participation in CS.

A number of mechanisms can help to achieve quality assurance and control in CS projects, while at the same time prompt learning and participation. These mechanisms should be made explicit to enable volunteers to reflect and improve their practices over time, and scientists to assess the quality of data collected. There is yet a need for transparent and accessible data management standards that can help assess CS projects and the degree to which they produce reliable results (Borda et al., 2020). Toward that direction, technology plays a major role in producing high quality datasets. In the case of volunteers who are designing their own investigations, it can support the design and management within a single environment and help to standardize processes of data collection. In particular, the authoring tool in nQuire scaffolds the process of designing, managing, piloting, improving, and launching a CS project, by structuring the project around four stages and giving guidelines as to what is required in each stage. For projects focused specifically on biodiversity and species identification, it could be used alongside other platforms that scaffold the identification process through community contributions or machine learning algorithms such as iNaturalist and iSpot. Also, it scaffolds decisions around data collection by offering participants a library of tools they can use to collect data including, for example, image, geolocation, sensor, and text data. This enables any individual with or without expertise in science to initiate a project they or their communities are personally interested in. The process of review and approval is an opportunity for volunteers to communicate and learn from scientists about how they can improve their designs and create scientifically robust investigations. It is a quality assurance process which ensures that the content and structure of the investigation are appropriate by, for example, reviewing the questions, language, ethics, originality, and others. In terms of the data collected, these can be visualized in graphs while the mission is running

and shown live on nQuire. Also, data can be downloaded at any time by mission authors in the form of a CSV file for further analysis. Platforms such as PlutoF could be used for organizing and managing databases from across different projects or for making raw data available and open access.

Reflecting on two CS projects related to temperature and biodiversity, what became evident is the importance of providing instructions and modeling participation to enable volunteers to collect data in a consistent manner, by for example creating a relevant guide (e.g., Budde et al., 2017) of how to take temperature or showcasing what pollinators look like. Also, it is important to collect contextual information related to the project such as geographic information or volunteers' demographics (Meier et al., 2017; Parrish et al., 2019) that could help scientists assess the quality of measurements, for example, details about the device used to collect temperature or expertise in identifying species. The quality of data can be further enhanced at the point of analysis by considering for data cleaning and filtering as well as more specialized approaches to statistical analysis (e.g., Parrish et al., 2018) that may be specific to a field of study. Existing literature should also be considered, as lessons learnt from other studies can inform the design and implementation of a project or how data are analyzed and reported. Finally, CS projects should make explicit (state clearly) the benefits to science alongside the benefits to volunteers, as the latter can support data quality contributions (Nov et al., 2014), while project findings and interpretations should be shared with participants (Robinson et al., 2018) and if possible, encourage feedback and communication around them.

REFERENCES

- Anastopoulou, A., Sharples, M., Ainsworth, S., Crook, C., O'Malley, C., and Wright, M. (2012). Creating personal meaning through technology-supported science learning across formal and informal settings. *Int. J. Sci. Educ.* 34, 251–273. doi: 10.1080/09500693.2011.569958
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., et al. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives (Complete Edition)*. New York, NY: Longman.
- Aristeidou, M., and Herodotou, C. (2020). Online citizen science: a systematic review of effects on learning and scientific literacy. *Citiz. Sci. Theory Pract.* 5, 1–12. doi: 10.5334/cstp.224
- Audubon Center (2018). *Why We're Changing From "Citizen Science" to "Community Science."* Available online at: <https://bit.ly/3sf28ow> (accessed April 12, 2021).
- Borda, A., Gray, K., and Fu, Y. (2020). Research data management in health and biomedical citizen science: practices and prospects. *JAMIA Open* 3, 113–125. doi: 10.1093/jamiaopen/ooz052
- Budde, M., Schankin, A., Hoffmann, J., Danz, M., Riedel, T., and Beigl, M. (2017). Participatory sensing or participatory nonsense? mitigating the effect of human error on data quality in citizen science. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 1–23. doi: 10.1145/3131900
- Cornes, R. C., Dirksen, M., and Sluiter, R. (2020). Correcting citizen-science air temperature measurements across the Netherlands for short wave radiation bias. *Meteorol. Appl.* 27:e1814. doi: 10.1002/met.1814
- Dunn, S., and Hedges, M. (2018). "From the wisdom of crowds to going viral. The creation and transmission of knowledge in the citizen humanities," in *Citizen Inquiry: Synthesizing Science and Inquiry Learning*, eds C. Herodotou, M. Sharples, and E. Scanlon (London: Routledge), 25–41.
- ECSA (2015). *The Ten Principles of Citizen Science*. Available online at: <https://osf.io/xpr2n/wiki/home/> (accessed April 12, 2021).
- ECSA (2020). *ECSA's Characteristics of Citizen Science*. Available online at: https://ecsacitizen-science.net/wp-content/uploads/2020/05/ecsac_characteristics_of_citizen_science_-_v1_final.pdf (accessed April 12, 2021).
- Edwards, R. (2015). *Enhancing Informal Learning Through Citizen Science—Background Literature*. Washington, DC: Center for Advancing Informal Science Education. Available online at: https://www.informalscience.org/sites/default/files/Enhancing_Informal_Learning_Through_Citizen_Science_Review__PDF.pdf (accessed September 28, 2020).
- Freitag, A., and Pfeffer, M. J. (2013). Process, not product: investigating recommendations for improving citizen science "success". *PLoS ONE* 8:e64079. doi: 10.1371/journal.pone.0064079
- Froeling, F., Gignac, F., Hoek, G., Vermeulen, R., Nieuwenhuijsen, M., Ficorilli, A., et al. (2021). Narrative review of citizen science in environmental epidemiology: setting the stage for co-created research projects in environmental epidemiology. *Environ. Int.* 152:106470. doi: 10.1016/j.envint.2021.106470
- Gura, T. (2013). Citizen science: amateur experts. *Nature* 496, 259–261. doi: 10.1038/nj7444-259a
- Haklay, M. (2013). "Citizen science and volunteered geographic information: overview and typology of participation," in *Crowdsourcing Geographic Knowledge*, eds Sui, D., Elwood, S., and Goodchild, M. (Dordrecht: Springer), 105–122.
- Hecker, S., Garbe, L., and Bonn, A. (2018). *The European Citizen Science Landscape—A Snapshot*. London: UCL Press.
- Herodotou, C., Aristeidou, M., Miller, G., Ballard, H., and Robinson, L. (2020). What do we know about young volunteers? an exploratory study of participation in zooniverse. *Citiz. Sci. Theory Pract.* 5:2. doi: 10.5334/cstp.248
- Herodotou, C., Aristeidou, M., Sharples, M., and Scanlon, E. (2018). Designing citizen science tools for learning: lessons learnt from the iterative

AUTHOR CONTRIBUTIONS

CH produced a first draft of the manuscript after discussions with ES and MS about quality issues that could be discussed and reported in the paper. ES and MS reviewed and revised the paper, included additional literature, and elaborated on the arguments made. All authors contributed to the article and approved the submitted version.

- development of nQuire. *Res. Pract. Technol. Enhanc. Learn.* 13, 1–23. doi: 10.1186/s41039-018-0072-1
- Herodotou, C., Sharples, M., and Scanlon, E. (Eds.). (2017). *Citizen Inquiry: Synthesising Science and Inquiry Learning*. Routledge.
- Herodotou, C., Villasclaras-Fernández, E., and Sharples, M. (2014). “Scaffolding citizen inquiry science learning through the nQuire toolkit,” in *Proceedings of EARLI SIG 20: Computer Supported Inquiry Learning, August 18–20 (Malmö)*.
- Hunter, J., Alabri, A., and van Ingen, C. (2013). Assessing the quality and trustworthiness of citizen science data. *Concurr. Comput. Pract. Exp.* 25, 454–466. doi: 10.1002/cpe.2923
- Jordan, R., Crall, A., Gray, S., Phillips, T., and Mellor, D. (2015). Citizen science as a distinct field of inquiry. *BioScience* 65, 208–211. doi: 10.1093/biosci/biu217
- Kelling, S., Fink, D., La Sorte, F. A., Johnston, A., Bruns, N. E., and Hochachka, W. M. (2015). Taking a “Big Data” approach to data quality in a citizen science project. *Ambio* 44, 601–611. doi: 10.1007/s13280-015-0710-4
- König, A. (2017). “Sustainability science as a transformative social learning process,” in *Sustainability Science: Key Issues*, eds A. König and J. Ravetz (London: Routledge), 1–26.
- Kosmala, M., Wiggins, A., Swanson, A., and Simmons, B. (2016). Assessing data quality in citizen science. *Front. Ecol. Environ.* 14, 551–560. doi: 10.1002/fee.1436
- Li, J. S., Hamann, A., and Beaubien, E. (2020). Outlier detection methods to improve the quality of citizen science data. *Int. J. Biometeorol.* 64, 1825–1833. doi: 10.1007/s00484-020-01968-z
- Meier, F., Fenner, D., Grassmann, T., Otto, M., and Scherer, D. (2017). Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Clim.* 19, 170–191. doi: 10.1016/j.uclim.2017.01.006
- Nov, O., Arazy, O., and Anderson, D. (2014). Scientists@ Home: what drives the quantity and quality of online citizen science participation? *PLoS ONE* 9:e90375. doi: 10.1371/journal.pone.0090375
- Parrish, J. K., Burgess, H., Weltzin, J. F., Fortson, L., Wiggins, A., and Simmons, B. (2018). Exposing the science in citizen science: fitness to purpose and intentional design. *Integr. Comp. Biol.* 58, 150–160. doi: 10.1093/icb/icy032
- Parrish, J. K., Jones, T., Burgess, H. K., He, Y., Fortson, L., and Cavalier, D. (2019). Hoping for optimality or designing for inclusion: persistence, learning, and the social network of citizen science. *Proc. Natl. Acad. Sci. U.S.A.* 116, 1894–1901. doi: 10.1073/pnas.1807186115
- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., et al. (2004). A scaffolding design framework for software to support science inquiry. *J. Learn. Sci.* 13, 337–386. doi: 10.1207/s15327809jls1303_4
- Robinson, L. D., Cawthray, J. L., West, S. E., Bonn, A., and Ansine, J. (2018). “Ten principles of citizen science,” in *Citizen Science: Innovation in Open Science, Society and Policy*, eds S. Hecker, M. Haklay, A. Bowser, Z. Makuch, J. Vogel, and A. Bonn (London, UCL Press), 1–23.
- Sauermann, H., Vohland, K., Antoniou, V., Balázs, B., Göbel, C., Karatzas, K., et al. (2020). Citizen science and sustainability transitions. *Res. Policy* 49:103978. doi: 10.1016/j.respol.2020.103978
- Sharma, N., Colucci-Gray, L., Siddharthan, A., Comont, R., and Van der Wal, R. (2019). Designing online species identification tools for biological recording: the impact on data quality and citizen science learning. *PeerJ Life Environ.* 6:e5965. doi: 10.7717/peerj.5965
- Sharples, M., McAndrew, P., Weller, M., Ferguson, R., FitzGerald, E., Hirst, T., et al. (2013). *Innovating Pedagogy 2013*. Milton Keynes: Open University.
- Shirk, J., Ballard, H., Wilderman, C., Phillips, T., Wiggins, A., Jordan, R., et al. (2012). Public participation in scientific research: a framework for deliberate design. *Ecol. Soc.* 17:29. doi: 10.5751/ES-04705-170229
- Skarlatidou, A., Hamilton, A., Vitos, M., and Haklay, M. (2019). What do volunteers want from citizen science technologies? a systematic literature review and best practice guidelines. *JCOM J. Sci. Commun.* 18:A02. doi: 10.22323/2.18010202
- The Royal Society (2020). *Why science is for me?* Available online at: <https://royalsociety.org/topics-policy/education-skills/teacher-resources-and-opportunities/resources-for-teachers/resources-why-science-is-for-me/> (accessed September 29, 2020).
- Tweddle, J. C., Robinson, L. D., Pocock, M. J. O., and Roy, H. E. (2012). *Guide to Citizen Science: Developing, Implementing, and Evaluating Citizen Science to Study Biodiversity and the Environment in the UK*. Lancaster: NERC/Centre for Ecology and Hydrology. Available online at: <http://nora.nerc.ac.uk/id/eprint/20678/> (accessed October 4, 2020).
- Weigelhofer, G., and Pölz, E. M. (2016). “Data quality in citizen science projects: challenges and solutions,” in *Front. Environ. Sci. Conference Abstract: Austrian Citizen Science Conference (Lunz am See)*, 1–4.
- Wiggins, A., and Crowston, K. (2011). “From conservation to crowdsourcing: a typology of citizen science,” in *2011 44th Hawaii International Conference on System Sciences (Washington, DC: IEEE)*, 1–10.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Herodotou, Scanlon and Sharples. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.