International Journal of Science Education, Part B
Communication and Public Engagement

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rsed20

Science literacy in action: understanding scientific data presented in a citizen science platform by non-expert adults

Yaela N. Golumbic, Barak Fishbain & Ayelet Baram-Tsabari

To cite this article: Yaela N. Golumbic, Barak Fishbain & Ayelet Baram-Tsabari (2020) Science literacy in action: understanding scientific data presented in a citizen science platform by non-expert adults, International Journal of Science Education, Part B, 10:3, 232-247, DOI: 10.1080/21548455.2020.1769877

To link to this article: https://doi.org/10.1080/21548455.2020.1769877

Published online: 30 May 2020.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, PART B
2020, VOL. 10, NO. 3, 232–247
https://doi.org/10.1080/21548455.2020.1769877

Science literacy in action: understanding scientific data presented in a citizen science platform by non-expert adults

Yaela N. Golumbic [a,b], Barak Fishbain[c] and Ayelet Baram-Tsabari [a]

[a]Faculty of Education in Science and Technology, Technion- Israel Institute of Technology, Haifa, Israel; [b]Faculty of Science, Sydney University, Sydney, Australia; [c]Faculty of Civil and Environmental Engineering, Technion- Israel Institute of Technology, Haifa, Israel

ABSTRACT

Citizen science is transforming the ways scientific knowledge is created, in that citizens participate in active scientific research, and large scientific databases can be accessed online. However, data availability does not guarantee public use or the relevance of these resources. This paper addresses the ways in which non-expert adults involved in a citizen science initiative, perceive, understand and use its scientific information.

Participants responded to an online questionnaire presenting air quality data from 'Sensing the Air' citizen science platform, followed by interpretation questions ($n = 123$). The results showed that 70% of participants were able to interpret the data presented in various visual representations. No differences were found between gender, age or education level. However, respondents with tertiary scientific education obtained higher average scores. Among users who had previous experience with the project, overall scores were higher, and differences based on respondents scientific education were fewer. This may suggest that while scientific education is important in providing skills for data interpretation, it is not the only way to acquire these skills. This study highlights the ability of non-experts to understand and apply scientific data in daily situations and the potential of citizen science to develop scientific skills, competencies and public understanding of science.

ARTICLE HISTORY
Received 31 December 2019
Accepted 12 May 2020

KEYWORDS
Public understanding of science; science literacy; citizen science; air quality

Introduction

One of the main goals of science education in schools is to develop science literacy and provide students with opportunities to develop the knowledge and skills they need to live in a scientifically and technologically enhanced society (National Academies of Sciences Engineering and Medicine, 2016; Ryder, 2001). The idea underlying this notion is that science and technology are prevalent in our day-to-day lives, today more than ever. Navigating this world and making informed decisions on topics such as medical treatment and nutrition is complex, and requires some level of understanding of scientific concepts and practices.

Feinstein (2011) coined the term 'competent outsider' to describe nonscientists who are capable of accessing and making sense of science relevant to their lives. Competent outsiders should have the capability to access and interpret science related to specific practical problems they may have. This is especially crucial in a post- truth era, characterized by a rise in misinformation circulated via social

CONTACT Yaela N. Golumbic yaelago123@gmail.com Faculty of Education in Science and Technology, Technion- Israel Institute of Technology, Haifa, Israel; Faculty of Science, Sydney University, Australia

media which threatens to undermine our ability to recognize truth (Lewandowsky et al., 2017; Peters, 2018).

Science literacy in this context demands both understanding and the implementation of scientific data (Murcia, 2009). Understanding scientific information without the will or ability to implement its conclusions, would not meet the expectations of being a competent outsider. This is because the purpose of developing scientific competences is precisely for their application: solving problems by applying newly acquired knowledge, facts, methods and rules (National Academies of Sciences Engineering and Medicine, 2016).

However, making a decision based on evidence is often far from straightforward. It is particularly difficult when trying to interpret large databases which include many variables, factors and conditions. While thoughtful visualization of the data can assist this process, some scientific skills are still necessary to fully understand scientific data and make them useful to the individual. These skills include reading graphs, formulizing research questions and reasoning with evidence (National Research Council, 2014). It is not clear, however, whether non-expert adults have the appropriate competences to do so. This question was explored here in the context of citizen science.

In recent years, ordinary citizens have had the opportunity to be exposed and contribute to the creation of new scientific knowledge through participation in citizen science projects. Citizen science is a collaborative effort of citizens and scientists, where members of the public are actively engaged in scientific research projects (Bonney, Ballard, et al., 2009). Participants engage in data collection, classification and analysis, in addition to opportunities to access large databases and engage in dialogue with experts (Mahr et al., 2018). Citizen science provides opportunities for people to directly examine, understand and use real time scientific information, without the need of intermediaries.

When discussing scientific data produced for and by the public, such as in citizen science projects, the importance of clear data visualization is enhanced. A fair requirement in citizen science is that data collected and produced by the public be transparent and accessible to the public (Albagli et al., 2015). Many citizen science projects do so by opening their data online and allowing their resources to be reused by interested parties (Golumbic, Baram-Tsabari, et al., 2019). However, making the raw data available does not guarantee its accessibility, public relevance or the use of these resources. Additional efforts are needed to make the data understandable and create visual displays that people can make sense of.

In our previous work, described in detail in Golumbic, Fishbain, et al. (2019), we describe the design of an online data presentation platform for the 'Sensing the Air' citizen science initiative. This design utilized a user centered design approach for identifying participants' needs and requirements, and developing the platform towards those needs.

Here, we continue this work by exploring the public's understanding of the air quality data presented in the platform. We examine how Sensing the Air platform was used in practice and how adults with different educational backgrounds perceive, understand and use scientific information presented in different visualization styles, to better determine how non-experts understand and apply complex scientific air quality data. This study has important implications in understanding the ability of citizen science and scientific data visualizations, in raising scientific literacy within the public domain.

Two research questions guided this work:

(1) How do people understand and apply scientific information presented in a range of visual displays within Sensing the Air platform?
(2) What is the association between scientific education and experience engaging with the platform, and participants' understanding and use of the platform visual displays?

Related literature

The term 'public understanding of science' is used in many contexts to describe public's scientific knowledge, comprehension or beliefs. In many cases it is used interchangeably with the term 'science literacy', thus underscoring its importance for informed decision making. While there is no one definition of public understanding of science, it is fair to assume it entails more than simply knowing scientific facts but instead having a holistic view of them (Huxster et al., 2018). For the purpose of this study, we define science literacy as the ability of an individual to access scientific information relevant to one's life, make sense of the information and use it to make informed decisions. We consider public understanding of science as a subset of science literacy and define it as the ability to make sense of scientific information and apply it in various contexts.

On the societal level, science literacy is often assessed through large public surveys of science opinions and knowledge. Many of these surveys are longitudinal and provide an ongoing picture of average performances over time and across countries (National Academies of Sciences Engineering and Medicine, 2016). The Science and Engineering Indicators (National Science Board, 2018) have been used to measure factual scientific knowledge in the US since 1979 and more recently include reasoning and understanding the scientific process. The findings indicate there is little change in adults' average scientific knowledge and understanding of scientific processes over time. However, scores were found to be closely related to levels of formal education. For example, whereas participants who had not completed high school had an average score of 43% on the science and engineering indicators, individuals with a bachelor's degree scored 74% on average (National Science Board, 2018). Similar results have been found in two Pew Research reports addressing scientific knowledge and interests and participation in scientific activities (Funk & Goo, 2015; Funk et al., 2017). In both studies, individuals with higher levels of education showed greater scientific knowledge, interest in science and an increased likelihood of taking part in science activities.

Studies on scientific literacy in schools illustrates the difficulties involved in teaching for this purpose (e.g. Archer-Bradshaw, 2017; Linder et al., 2007). Classroom science often does not reflect the actual practice of science as a way to develop explanations for natural phenomena using evidence and logic (Crawford, 2013). Trying to develop an understanding of real, often incomplete and uncertain science in a classroom is a complex process which often fosters miscomprehension (Allchin, 2014).

This problem is not restricted to school environments. Research has shown that topics such as uncertainty and probability are often not understood by many members of the public (e.g. Dunwoody, 2016). Studies of public engagement with science in both online and offline environments indicate that scientific data and concepts are used selectively to reinforce personal opinions; good examples are vaccinations (Larson, 2018; Orr & Baram-Tsabari, 2018), genetically modified foods (Landrum et al., 2019) and climate change (Kahan, 2017; Whitmarsh, 2011). Ultimately, public attitudes are often made up of many types of knowledge, personal judgment, values and trust (Wynne, 1991).

Furthermore, personal decisions on scientific topics are often influenced by media, context in which science information is presented, text comprehensibility and readers' comments on social media (Bromme & Goldman, 2014). Scharrer et al. (2017), describe the 'easiness effect of science popularization' in which oversimplification of scientific content causes individuals to rely on their personal capabilities when making judgments about scientific claims rather than recognize their inabilities. They therefore recommend science communicators and educators highlight topic complexity and controversy when informing laypeople about science in a comprehensible manner.

Another approach to strengthening public understanding of science is through active participation in scientific activities (Bonney et al., 2016). By actively engaging lay audiences in science it is assumed that the public will become more aware, informed and in the long run develop important scientific capabilities (National Academies of Sciences Engineering and Medicine, 2017). Several citizen science projects have reported increased learning and participatory opportunities. Participants have been reported to create a scientific identity, learn scientific content and skills, and develop personal responsibility (Jordan et al., 2015). However, the ways to best design and plan for these outcomes is not well understood. Furthermore, while citizen science often provides open access to databases, availability of information does not mean it is accessible or comprehensible by general audiences. This may require additional facilitation and design geared specifically to helping the public understand scientific data.

Data visualization

One way to facilitate the process of accessing and interpreting scientific data is through data visualization. Visualization involves grouping and developing organizational frameworks (Börner & Polley, 2014). There are many ways to group data through visualization depending on the questions raised. These include charts, such as pie chart and word clouds, tables which are simple and effective ways to convey data, graphs which are the most common form of visualization, geospatial maps and network graphs (Börner & Polley, 2014). Deciding which type of visualization to use can be challenging and may have implications for data interpretation (McCrudden & Rapp, 2017). Ultimately, the presentations of evidence for scientists and citizen scientists alike should be consistent with the analytic task at hand, which usually involves understanding causality, making multivariate comparisons, examining relevant evidence and assessing its credibility (Tufte, 2006).

How individuals interpret and learn from data visualizations is essential to understanding people's broader understanding of science. Making sense of unfamiliar visualizations is a multiple stage cognitive process, that begins with exposure to the visualization, framing, exploring, and questioning it (Lee et al., 2016). Visual displays tap such cognitive processes as selection, organization and integration of data and hence efficiency of interpretation (McCrudden & Rapp, 2017). Once familiar with a visualization style, sense making becomes easier and more straightforward.

Method

Research setting

This study was conducted as part of the citizen science project 'Sensing the Air', initiated in 2015 by the Technion, the Israel Institute of Technology with the aim of facilitating air quality research through active involvement of volunteers (Golumbic, Fishbain et al., 2019). Sensing the Air utilizes new Micro Sensing Technologies (MSUs), distributed in collaboration with project participants to continuously monitor air quality in the local environment. Measurements were transmitted daily from sensors to a central database for two purposes: (1) For participants to access air quality data for personal use in their day-to-day life, and (2) For scientists to use for modeling air quality and examining the validity of the network of sensors.

Access to the air quality data was provided using an open online platform (see Figure 1, and http://sensair.net/map.php). The platform presents both sensor data (collected through Sensing the Air) and data collected by the government and municipalities (data which is publicly available online but hard to locate and difficult for lay audiences to understand). Air quality information is presented in three formats: a general map, a pollutant-specific display per location, and graphs displaying pollutant concentrations over time. These formats respectively provide spatial, pollutant-specific and temporal information of air quality.

Study design and sample

This study was designed to examine the public's understanding of air quality data presented on the Sensing the Air platform. For this purpose, we constructed a questionnaire presenting authentic snapshots from the platform (see *Research tools and analysis* for more details), and presented them to non-experts with various science education backgrounds to determine how they understood and applied the data. Data collection spanned 6 months from December 2016 to June 2017, and resulted in 123 fully completed questionnaires.

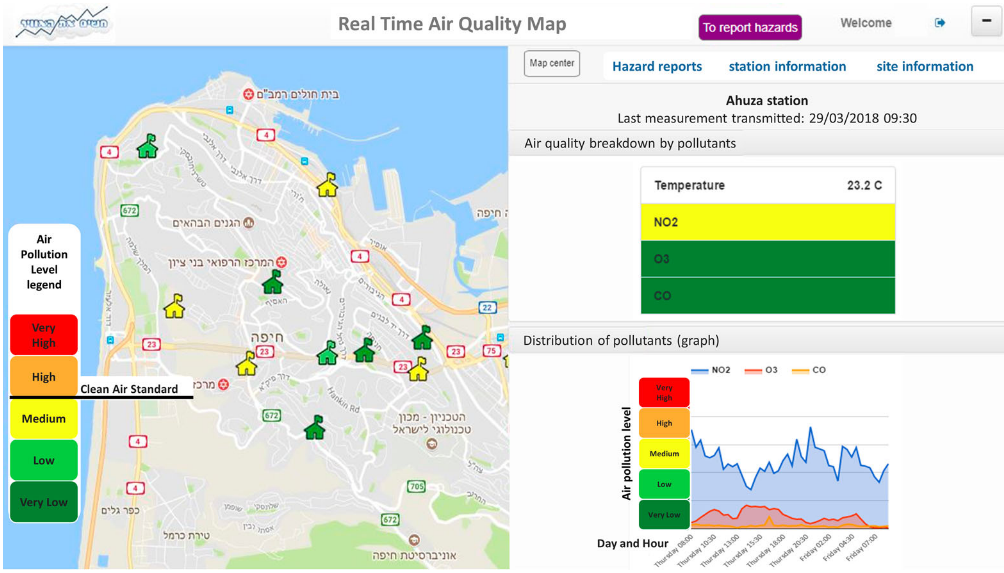


Figure 1. Visualizations on the Sensing the air platform, presenting air quality data. The figure was translated from Hebrew.

The research population was composed of Sensing the Air project participants, who were defined as adult non-experts in air quality, who volunteered for the Sensing the Air project, registered as users on the project platform, or attended a project public activity or lecture (these activities did not include explanations or demonstration of the project platform). The questionnaire was distributed to all project participants online via the Sensing the Air website, e-mail listings, and its Facebook page. In addition, participants in Sensing the Air activities and/or lectures were asked to fill out the questionnaire during the activity they attended.

Since project participants (based on the above definition) are quite diverse in terms of time commitment to the project and motivations to participate, the sample of respondents was divided into two subgroups for purposes of this study:

- Authentic users who engaged with the project of their own volition and have registered as users on the project website or volunteered to host a sensor in their home. Respondents completed the questionnaire online at their convenience, with a 22% response rate and a total of $n = 32$.
- One-time participants who attended a project talk as part of a course or teacher training session. Respondents completed the questionnaire at the end of the activity, with approximately 70% response rate (the exact number is not available) and with a total of $n = 91$ respondents.

Sample demographics (Table 1) were similar within the two participant subgroups and were consistent with other citizen science projects, which tend to engage educated participants (Soleri et al., 2016). However, our sample was more diverse than usually reported in terms of age and gender.

Research tools and analysis

Questionnaire design and structure

The main research tool used for this study was an online questionnaire. The questionnaire was built around the existing displays of air quality information on the Sensing the Air platform and aimed to examine how participants understood the air quality data presented, while determining platform accessibility and usability. The questionnaire had four sections, each based on one visualization type: map, table, graph and a combination of the three. Each visualization type presented different

Table 1. Demographics of respondents to the online questionnaire.

Demographic parameter					
Gender		**Male**	**Female**		
	One-time	43%	57%		
	Authentic	44%	56%		
	Total	43%	57%		
Age		**18–24**	**25–30**	**31–50**	**51–70**
	One-time	4%	23%	58%	14%
	Authentic	10%	29%	45%	16%
	Total	6%	24%	55%	15%
Tertiary education		**None**	**BA/BSc/BEd**	**MA/MSc/MEd**	**PhD**
	One-time	5%	45%	44%	5%
	Authentic	6%	44%	47%	3%
	Total	5%	45%	45%	5%
Highest level of science education		**Middle school**	**High school**	**Tertiary degree**	
	One-time	18%	22%	60%	
	Authentic	16%	25%	59%	
	Total	17%	23%	60%	

information and was designed to answer a different question about the data; namely: Where are the data from? What do the data present? When was the data collected? (Börner & Polley, 2014).

The questions (summarized in Table 2) were structured around screen shots taken from the data presentation platform, and used a multiple choice format (see example in Figure 2). The questions were structured to identify the respondents' ability to recognize the differences between what each visualization type presents, and what can be learned from it. For example, Figure 2 presents a map of the neighborhood with sensor locations, each indicating the air pollution measured on a 1–5 color-coded scale (green- very low to red- very high). The figure was followed by the question: Can you use this map to tell what area of the neighborhood is more polluted? Since this question refers to spatial data and aims to answer the 'where' question, the correct answer is 'yes'.

In addition to the close-ended questions, two open-ended question were included in the questionnaire and which assisted in understanding respondents reasoning process. The first asked respondents to explain their choice of answer for item 4 (which has additional complexity as described in the results section). The second open question, was located at the end of the questionnaire and was used as an integration question, dealing with the application of the air quality information in practice. This question asked respondents to recommend a walking route between two points at a specific hour of the day, based on the information provided in the platform. The average completion time for the questionnaire was 20–30 min.

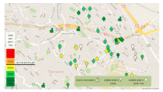
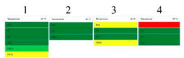
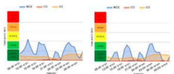
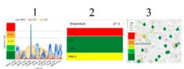
Validity

Content validity was assessed by asking two air quality experts to review the questionnaire and indicate the correct scientific answers. In addition, two project participants were administered the questionnaire during an interview and were asked to express their opinions on the questionnaire structure and phrasing. The revised questionnaire was further reviewed by five independent science education researchers who checked the questionnaire phrasing, provided insights into possible misunderstandings and suggested rephrasing options. In addition, they assessed the difficulty of questions for each visualization and ensured the questions were of similar difficulty level.

Ethical considerations

IRB approval was obtained from the Technion institutional committee (approval: Nov. 2014). All participants expressed their full consent for the academic use of the data.

Table 2. Summary of multiple-choice questions, correct answers and finding.

Question number	Visualization type	Visualization	Visualization question	Statement	Correct answer	% correct answers	Averaged percent (±std)
1	map		Can you use this map to tell:	What area of the neighborhood is more polluted?	yes	88.3	77.5 (±10.9)
2				What the most common pollutants in the neighborhood are?	no	64.7	
3				What area of the neighborhood has good air quality?	yes	84.8	
4				Why is there pollution in the neighborhood?	no	72.4	
5	table		Match the findings in the figures to the following conclusions	PM10 is the pollutant with the highest concentration	1	71.4	76.2 (±7.8)
6				One pollutant is the cause of very bad air quality	4	81.0	
7				Air quality is very good	2	84.5	
8				A number of pollutants cause the air quality to be medium	3	67.8	
9	graph		The two graphs below provide information from co-located sensors. What can be learned from them?	NO2 is the pollutant whose concentration varies the most throughout the day	yes	95.6	73.0 (±15.8)
10				The pollutant most influencing air quality is NO2	yes	58.2	
11				CO is not dangerous	no	63.4	
12				The similarity between the graphs suggest they are showing the same air phenomenon	yes	83.4	
13				Air quality in this location is generally good	yes	64.3	
14	Combined visualization		Which of the displays answers the following research questions?	When was the highest concentration of NO2 recorded?	1	75.4	72.2 (±5.2)
15				What is the air quality in Hanita St.?	3	75.2	
16				What air pollutant is causing the bad air quality now?	2	63.5	
17				What is the CO concentration over time?	1	76.0	
18				Which sensors measure medium air quality?	3	71.1	
total						74.55 (±10.1)	

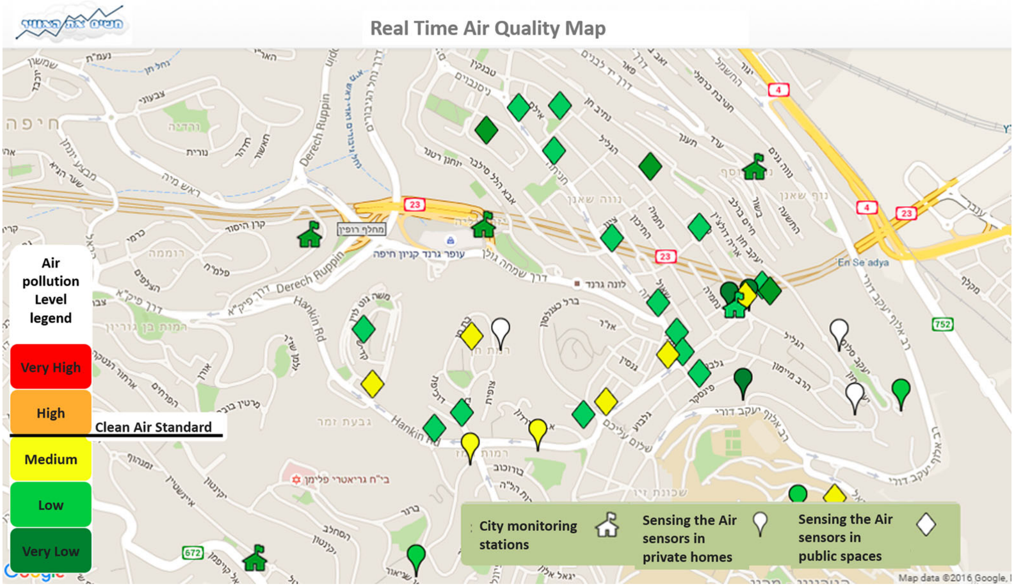


Figure 2. Print-screen from the data presentation platform used in the questionnaire (modified and translated from Hebrew). The figure is followed by the question: can you use this map to tell what area of the neighborhood is more polluted? (yes).

Data analysis

Analysis of the multiple-choice responses applied SPSS software. Two-way ANOVAs were used to compare group scores, followed by Scheffe post-hoc tests when significant F values were found. Chi-Square tests were used with categorical variables. Answering 'I don't know' was considered as an incorrect answer for the purpose of calculating the number of correct answers. In the case a respondent did not answer a specific question, the average was calculated without it. Open questions were analyzed thematically, by identifying emerging themes in the responses (Guest et al., 2011). These were subsequently used to create clusters of recurring issues and for coding the responses categorically (Braun & Clarke, 2006). Interrater reliability for the coding of the open-ended questions on 15% of the items was above 90%. Responses for the open-ended questions ranged from 49% to 63% replies per question.

Results

Our overarching goal was to explore the ways in which non-experts understand and apply complex scientific air quality data. This was examined in a range of visual displays across respondents with differing scientific backgrounds and experience with the platform. We presented participants with screen shots from Sensing the Air platform followed by a series of questions. A summary of the close-ended questions and the percentage of correct answers appears in Table 2.

Understanding and using scientific information

Overall scores for the multiple-choice section of the questionnaire were relatively high, with an average of 12.6 correct answers out of 18 (70%). The distribution of scores was negatively skewed (most scores were on the higher side of the scale with very few low scores) with over 50% of the respondents ($n = 69$) answering 13–17 correct answers. No differences were found between participants as a function of gender, age group or level of education.

The percentage of correct responses ranged from 58.2% to 95.6% between questions, with an average of 74.5 and a SD = 10.1 (Table 2). Grouping the responses according to presentation type (map, table, graph and combined), all had similar averages ranging from 72.2% to 77.5%. However, differences were found within each visualization style. The questions associated with the lowest percentage of correct answers (over one standard deviation below the mean), hence indicating the difficulty of interpreting the information in them, were items 10, 11, 13 and 16 from the graph and combined visualizations. The questions associated with the highest percentage of correct answers (over one standard deviation above the mean), hence indicating the relative clarity of the information in them, were items 1, 3 and 9 from the map and graph visualizations.

These results provide some insights into the clarity of the different visualization styles. Interestingly, the graph presentation had both the highest and lowest average scored items. This would suggest the graph presentation had some clear and some more difficult information to interpret. Closer inspection shows that the graph presentations entailed different types of data, including temporal information on the X axis, and air quality levels from a number of air pollutants on the Y axis. As such, much can be learned and understood from the graph. Some of the data may be easier to interpret than others, explaining the large differences found within this group.

The map visualization had some items with the highest percentage of correct answers, but no items with a low percentage of correct answers. This may suggest the map visualization is clearer to understand than the other visualization types.

A deeper look at the map visualization examined the reasoning behind respondents' answers to question number 4: 'Can you use this map to tell why is there pollution in the neighborhood?' (no). This was done using an open-ended question shedding light on how people understood the details in the map and what can and cannot be learned from it. Question number 4 may be confusing since assumptions could be made to explain why pollution was present in the neighborhood. However in practice, the data to prove these assumptions could not be obtained from the map; therefore the correct answer was 'no'. Of the 84 respondents who correctly answered item 4 (that the question cannot be answered), 58 answered the open-ended question providing three levels of explanation (Table 3): (A) A statement that the questions cannot be answered, or that there is no data to answer the question, with no additional explanation. For example: '*there is no way of knowing why by looking at the map*'. This level of response categorized 10% of the respondents. (B) An explanation stating that the map contained no information about types of pollutants and/or pollution sources. For example: '*There is no indication of pollution sources, only pollution levels*'. This level of explanation was coded as matching 55% of the respondents. (C) Expressing speculations as to the causes of pollution, but noting they cannot be deduced from information on the map alone. For example: '*Air pollution can emanate from a number of sources such as industries and transportation. It is difficult to determine the source from the level of pollution stated on the map*'. This level corresponded to answers provided by 14% of the respondents. In addition, 21% of the respondents in this category provided other explanations such as outlining the pollution sources in the area or stating there was pollution (Table 3).

Of the respondents who answered item 4 incorrectly, the vast majority (8 out of 9 comments) gave a reasoned albeit wrong answer. This included explanations that sources of pollution can be interpreted from the map (an incorrect assumption), for example: '*You can see that there are yellow marks on the main street, this is linked to traffic congestion on the road*'. Although this explanation makes a valid point, and may indeed be the reason for pollution in the area, it cannot be deduced from the map and therefore is incorrect. Similar explanations were given by four respondents in this group. Four other respondents speculated about the causes of pollution without indicating whether and how the map supported their claim. An example is '*Due to industry emissions*' (Table 3).

Finally, among the respondents who replied they did not know the answer, only three out of ten provided a reasoned answer explaining there are too many unknown factors. For example: '*There seems to be a relationship between the road and pollution. However, this is very simplistic and ignores many variables. Therefore, it is not possible to determine whether the map provides an answer to the*

Table 3. Summary of responses to the open-ended question related to item 4: why did you answer item 4 the way you did? The responses are categorized according to types and levels of explanations.

	Answered correctly (*n*, %)		Answered incorrectly (*n*, %)		'I don't know' (*n*, %)	
Level of explanation	***N***	%	***N***	%	***N***	%
Simple statement	6	10%	1	11%	5	50%
Provide explanation	32	55%	4	44%	3	30%
Additional speculation	8	14%	4	44%	2	20%
Other explanations	12	21%				

question above'. This explanation is very detailed and correct, but should have led the respondent to conclude that the question cannot be answered. The remainder of respondents either did not provide an explanation (e.g. '*not clear*') or cited simple speculations (e.g. '*mostly because of roads*') (Table 3).

The final questionnaire item was an open-ended question examining the application of the air quality information in a practical context. This question aimed to measure respondents' consolidation of the information presented in the platform and their ability to turn the information into practical recommendations. The responses were classified based on two variables: A. which representation of the platform the respondents said they would use to answer the question (map, graph or both). B. whether they provided an explanation of the use they would make of the representation. A full answer was defined as including a combination of information from the map and graph and an explanation that the map could provide a general indication of the route, based on locations with green sensors, whereas the graph could be used to check these sensors at specific hours of the day.

This yielded four types of responses: 1. A full answer, which combined information from the map and graph with relevant explanations. 2. A partial answer with information from both the map and graph, but with no explanation. 3. A partial answer, with either information from the map or the graph with relevant explanations. 4. A partial answer with no explanation. The majority of the respondents (40%) gave a partial answer, followed by an explanation that the map provided spatial information about pollution (Table 4). For example, '*I would use the map to find out which route had the least air pollution*'. An additional 32% of the respondents answered they would use the map, but did not provide an explanation or a use case. Altogether 72% of the respondents said they would use the map alone. Seventeen percent of the respondents provided a full answer with a full explanation, and 7% gave a full answer without an explanation. Only 4% of the respondents did not include the map information at all in their answer (Table 4). This further strengthens the previous findings that the map was clear and useful for participants.

These responses provide some insights into the way the participants made use of the information presented. Explanations were provided by 46 respondents (61%), which demonstrates the extent to which they understood and were able to use the information presented in Sensing the Air platform. Interestingly, respondents who provided explanations had higher average scores on the

Table 4. Level of explanations provided by all respondents to the walking route recommendation question.

Level of explanation	*N*	%	Example
Full answer: map + graph + explanation	13	17%	*Using the map and graph, I could suggest a route with the lowest pollution at a specific hour (one-time participant, engineer, 30–50)*
Partial answer: map + graph, NO explanation	5	7%	*Based on the map and the distribution of pollutants throughout the day (one-time participant, programmer, 25–30)*
Partial answer: map + explanation	30	40%	*I would use the map to find out which route had the least air pollution (one-time participant, programmer, 30–50)*
Partial answer: graph + explanation	3	4%	*I would look at the history of measurements in that location (one-time participant, engineering student, 18–24)*
Partial answer: map NO explanation	24	32%	*Based on the map (one-time participant, programmer, 30–50)*

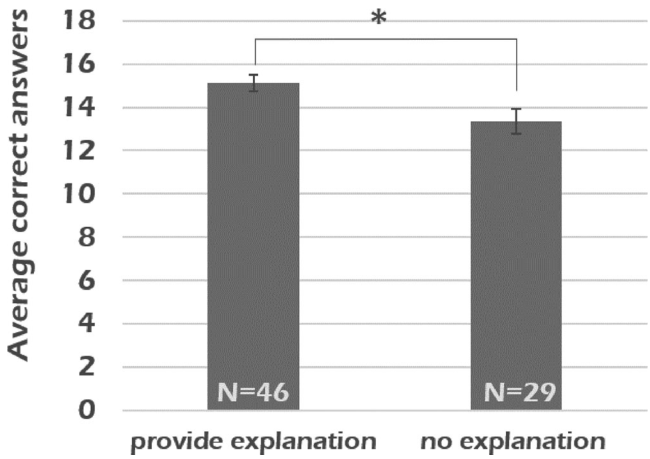


Figure 3. Average scores for respondents who provided and did not provide explanations for the recommendation question. Number of respondents are indicated at the bottom of columns. The difference was statistically significant ($p < 0.05$).

questionnaire than respondents who did not provide explanations ($t = 2.6$, $p < 0.05$) (Figure 3), indicating that a better understanding of the data could result in increased use.

Scientific education and experience with the platform

The second research question was designed to examine the association between scientific education and the ability to interpret the visualizations, comparing responses of one-time with those of authentic users. Analysis of the multiple-choice questions based on the respondents' level of scientific education revealed higher scores among respondents with a tertiary scientific degree (mean = 13.2) relative to respondents who last studied science in middle school (mean = 10.3, $F = 3.8$ $p < 0.05$) (Figure 4). Furthermore, authentic users of the platform obtained higher scores on average (mean = 14.5) than one-time users (mean = 12, $F = 9.6$ $p < 0.05$) (and relative to the overall average, mean = 12.6). Interestingly, within this (admittedly small) group, no differences were found between respondents with different scientific education backgrounds (Figure 4).

Given the differences described above in the scores obtained by respondents with different scientific education backgrounds, we tested for interactions within the different visualizations styles on the questionnaire (map, table, graph and combined). No differences were found in terms of average

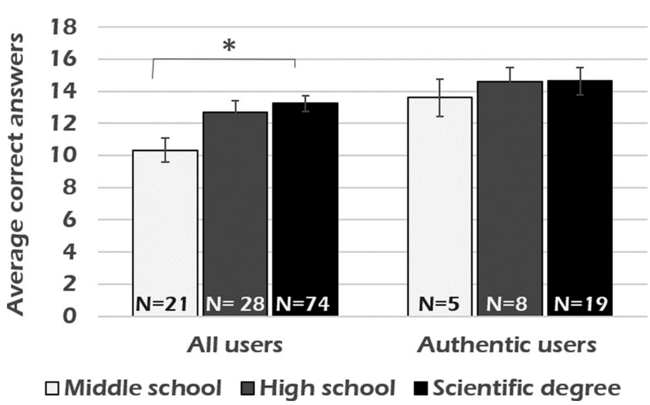


Figure 4. Average number of correct answers among respondents with different levels of scientific education. Number of respondents are indicated in the bottom of the columns. The asterisk indicates statistically significant differences ($p < 0.05$).

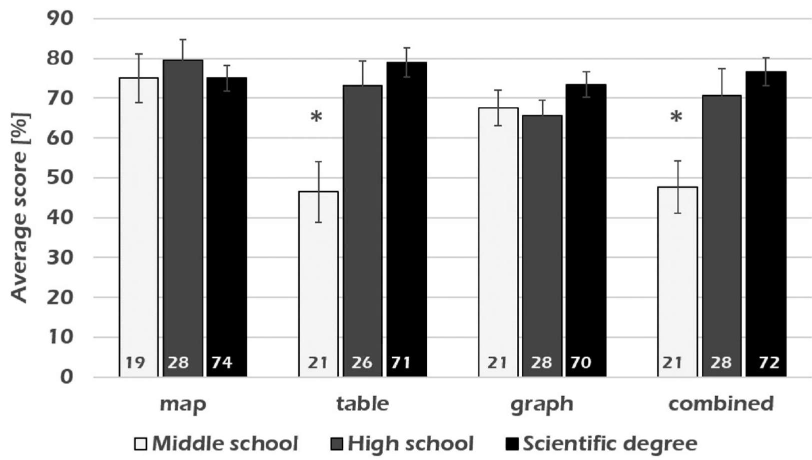


Figure 5. Average scores for the four visualization styles, according to respondents' scientific education. Number of respondents are indicated in the bottom of the columns. Asterisks indicate statistically significant differences ($p < 0.05$).

scores on the map visualization questions between respondents with different scientific backgrounds (Figure 5). Similarly, no differences were found for the graph visualization questions, which had both the highest and lowest scored items (Table 2). However, respondents who last studied science in middle school had significantly lower scores on the table and combined visualizations ($F = 8.4$ and 17.1 respectively, $p < 0.05$).

Finally, we examined the explanations provided based on the respondents' scientific education background. Although no significant differences were found in terms of the distribution of answers (full/partial answer), providing explanations was found to be associated with level of scientific education ($\chi^2 = 6.8$ $p < 0.05$). Respondents who had a tertiary scientific degree tended to provide explanations whereas respondents who last studied science in middle school did not provide such explanations (Figure 6(A)). Furthermore, authentic users tended to use more explanations than one-time users ($\chi^2 = 11.7$ $p < 0.05$) (Figure 6(B)) regardless of their scientific education.

Discussion

This study examined how citizens involved in the Sensing the Air initiative understand and apply the scientific information made available on the site. Since air quality data are often complex, confusing and hard to understand, this article aimed to assess how participants grasp such data when presented in a range of visual displays, and how scientific education and experience with the platform are associated with participants' understanding and ability to use the platform.

Overall, our findings suggest that access to air quality data provided through Sensing the Air platform is convenient and understandable to users. The average score across all the close-ended questions on the questionnaire was 70%, indicating that most participants were able to understand the air quality data presented well across presentation types. In comparison, a similar survey spanning five Zooniverse projects, with close to 2000 respondents, reported an average score of just under 50% in both general science knowledge and project specific science knowledge quizzes (Masters et al., 2016). However, the Zooniverse and Sensing the Air surveys may not be comparable, since they examine different levels of knowledge and understanding. The questions formulated in this study reflect practical use of the platform rather than examining knowledge or analysis skills. While these questions may have been easy for some participants, it may suggest that the Sensing the Air platform is easier to understand and more intuitive.

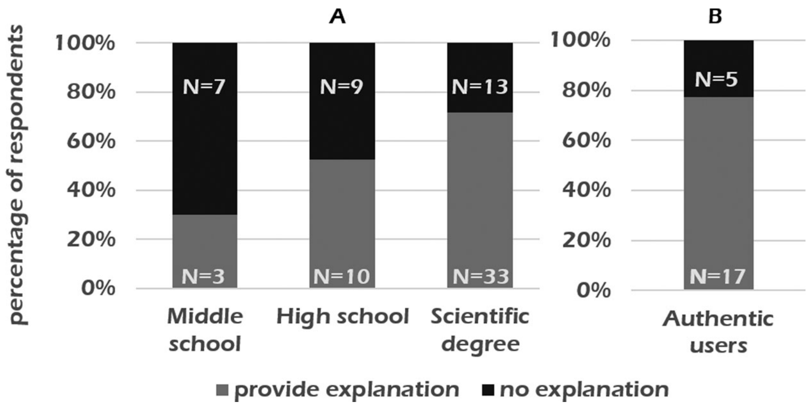


Figure 6. Respondents who provided and did not provide explanations for the recommendation question, as a function of their level of scientific education (A) and among authentic users (B). Number of respondents are indicated at the bottom of columns.

In fact, the participants in Sensing the Air stated that the platform interface is convenient and clear, and provides necessary practical information (Sensing the Air, non-published data). This contrasts with the many comments on the information available today from the Ministry of Environmental Protection (MoEP). According to participants, the MoEP information (presented on their website) is not sufficiently available, not transparent enough and therefore is not accessible.

One of the strengths of the Sensing the Air platform, as described previously (Golumbic, Fishbain et al., 2019) is its multiple level data presentations. This display enables each user to examine information at the level and type they find useful. The findings here suggest that while all presentation types were clear to participants, the map presentation appeared to be best understood across groups. Two of the map visualization items had the highest percentage of correct answers. In addition, almost all responses to the final integration question (96%) indicated that the participants would use that map presentation to select a pedestrian route as compared to 27% who indicated they would use temporal information from the graph. This again may underscore the clarity of the information presented on the map. Together these results suggest that the map visualization was clearer to understand and more useful than the other visualization styles used.

Maps are pictorial displays of object locations that are often presented in relation to other objects through static images (McCrudden & Rapp, 2017), making the interpretation process simpler (Börner & Polley, 2014). This could help account for the findings and provide insights into future visualization designs. Numerous citizen science projects use map displays to present their spatial data. Examples can be found in projects such as eBird[1], CoCoRaHS[2], OPAL[3] and many others (Golumbic, 2015).

Science education was found to be associated with the interpretation of the data presented in the platform as seen in the higher scores obtained by respondents with a scientific degree compared to participants who last studied science in middle school. This trend also emerged for respondents who provided explanations for the integration open-ended question, since they tended to be participants with higher scientific education levels. While this result presents an association between scientific education and higher scores, it does not inform us whether or not the reason for this result was due to the higher education. However, it is consistent with previous studies reporting a correlation between education levels and scientific knowledge, interest and participation (Funk et al., 2017; National Science Board, 2018).

However, among authentic users of the platform, there were fewer differences between scientific education backgrounds, and overall scores were higher. This may suggest that although scientific education is important in providing skills for interpreting scientific data, it is not the only way to acquire them. Participating in citizen science projects, being exposed to scientific data and assisting in its collection can help develop scientific skills and competencies, as suggested by Bonney, Cooper, et al. (2009). This process is intensified when the project and its goals are related to the participants' daily lives (Ballard et al., 2017), as it enhances motivation and interest. Similarly, in a study examining scientific knowledge of parents of hearing impaired children, scientific knowledge was found to be a predictor of contextual knowledge but not of advocacy knowledge and attitudes. Rather, parents were found to rely on heuristics from their daily experiences and acquire specific content knowledge only if needed (Shauli & Baram-Tsabari, 2018).

This study has certain limitations that may have introduced bias and therefore must be considered. First, its small and highly educated sample does not reflect the population. This may make it difficult to generalize the findings to diverse populations with less schooling. Second, the questionnaire took 20–30 min to complete, which may be longer than people wish to spend on a questionnaire. This could have discouraged potential individuals from participating. Furthermore, while most of the authentic users responded to the online questionnaire in their free time, the non-authentic users responded as part of a lecture, thus making their response rate higher. This could have introduced a bias in favor of the authentic users since it is possible that only confident participants completed the questionnaire.

Overall, given the literature on the development of science literacy and its goal to provide opportunities to develop scientific skills for daily life (National Academies of Sciences Engineering and Medicine, 2016; Ryder, 2001), citizen science clearly emerges as a useful tool to consider. As shown by the Sensing the Air platform, users of the platform exhibited a good understanding of the scientific information and provided cogent examples of the use of such data. This was facilitated by the user-friendly presentation of relevant real time air quality information. The findings thus underscore the importance of transparent and accessible information and the participants' varying ability to understand and use complex scientific data. The public's involvement in these issues can contribute to the management of local problems through the promotion of creative and useful ideas, and lead to greater public activism.

This paper provides an important contribution to science education and science communication, elucidating the potential of citizen science to develop scientific skills, competencies and public understanding of science. It highlights the ability of non-experts to understand and apply scientific data in daily situations and the importance of facilitating scientific information for public use. These finding also have implications for future design of citizen science websites and platforms, and underscore the importance in providing simple and clear visualizations for platform users and citizen scientists.

Notes

1. https://ebird.org.
2. https://www.cocorahs.org.
3. https://www.opalexplorenature.org.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by FP7 Environment [grant number 308524]; Israel Science Foundation [grant number 1716/12].

ORCID

Yaela N. Golumbic http://orcid.org/0000-0002-8523-5153
Ayelet Baram-Tsabari http://orcid.org/0000-0002-8123-5519

References

Albagli, S., Maciel, M. L., & Abdo, A. H. (2015). *Open science, open issues*. IBICT.

Allchin, D. (2014). From science studies to scientific literacy: A view from the classroom. *Science and Education*, *23*(9), 1911–1932. https://doi.org/10.1007/s11191-013-9672-8

Archer-Bradshaw, R. E. (2017). Teaching for scientific literacy? An examination of instructional practices in secondary schools in Barbados. *Research in Science Education*, *47*(1), 67–93. https://doi.org/10.1007/s11165-015-9490-x

Ballard, H. L., Dixon, C. G. H., & Harris, E. M. (2017). Youth-focused citizen science: Examining the role of environmental science learning and agency for conservation. *Biological Conservation*, *208*, 65–75. https://doi.org/10.1016/j.biocon.2016.05.024

Bonney, R., Ballard, H., Jordan, R., McCallie, E., Phillips, T., Shirk, J., & Wilderman, C. C. (2009, July). *Public participation in scientific research: Defining the field and assessing its potential for informal science education*. A CAISE Inquiry Group Report (pp. 1–58).

Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, *59*(11), 977–984. https://doi.org/10.1525/bio.2009.59.11.9

Bonney, R., Phillips, T. B., Ballard, H. L., & Enck, J. W. (2016). Can citizen science enhance public understanding of science? *Public Understanding of Science*, *25*(1), 2–16. https://doi.org/10.1177/0963662515607406

Börner, K., & Polley, D. E. (2014). *Visualization insight- a practical guide to making sense of data*. MIT Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bromme, R., & Goldman, S. R. (2014). The public's Bounded understanding of science. *Educational Psychologist*, *49*(2), 59–69. https://doi.org/10.1080/00461520.2014.921572

Crawford, B. (2013). Authentic science. In R. Gunstone (Ed.), *Encyclopedia of science education* (pp. 1–3). Springer. https://doi.org/10.1007/978-94-007-6165-0_144-2

Dunwoody, S. (2016). Communicating uncertainty in our mediated world: A preface. *Communications*, *41*(3), 231–237. https://doi.org/10.1515/commun-2016-0013

Feinstein, N. (2011). Salvaging science literacy. *Science Education*, *95*(1), 168–185. https://doi.org/10.1002/sce.20414

Funk, C., & Goo, S. K. (2015). *A look at what the public knows and does not know about science*. Pew Research Center.

Funk, C., Gottfried, J., & Mitchell, A. (2017). *Science news and information today*. Pew Research Center.

Golumbic, Y. N. (2015). *What makes Citizen Science projects successful, and what can we learn from them for future projects?* Technion Citizen Science Project (TCSP) Report (pp. 1–56).

Golumbic, Y. N., Baram-Tsabari, A., & Koichu, B. (2019). Engagement and communication features of scientifically successful citizen science projects. *Environmental Communication*. https://doi.org/10.1080/17524032.2019.1687101

Golumbic, Y. N., Fishbain, B., & Baram-Tsabari, A. (2019). User centered design of a citizen science air quality monitoring project. *International Journal of Science Education, Part B*, 1–19. https://doi.org/10.1080/21548455.2019.1597314

Guest, G., MacQueen, K. M., & Namey, E. E. (2011). *Applied thematic analysis*. SAGE Publications Inc.

Huxster, J. K., Slater, M. H., Leddington, J., Lopiccolo, V., Bergman, J., Jones, M., McGlynn, C., Diaz, N., Aspinall, N., Bresticker, J., & Hopkins, M. (2018). Understanding "understanding" in public understanding of science. *Public Understanding of Science*, *27*(7), 756–771. https://doi.org/10.1177/0963662517735429

Jordan, R., Crall, A., Gray, S., Phillips, T., & Mellor, D. (2015). Citizen science as a distinct field of inquiry. *BioScience*, *65*(2), 208–211. https://doi.org/10.1093/biosci/biu217

Kahan, D. M. (2017). 'Ordinary science intelligence': A science-comprehension measure for study of risk and science communication, with notes on evolution and climate change. *Journal of Risk Research*, *20*(8), 995–1016. https://doi.org/10.1080/13669877.2016.1148067

Landrum, A. R., Hallman, W. K., & Jamieson, K. H. (2019). Examining the impact of expert voices: Communicating the scientific consensus on genetically-modified organisms. *Environmental Communication*, *13*(1), 51–70. https://doi.org/10.1080/17524032.2018.1502201

Larson, H. J. (2018). Politics and public trust shape vaccine risk perceptions. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-018-0331-6

Lee, S., Kim, S., Hung, Y., Lam, H., Kang, Y., & Yi, J. S. (2016). How do people make sense of unfamiliar visualizations? A grounded model of novice's information visualization sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, *22*(1), 499–508. https://doi.org/10.1109/TVCG.2015.2467195.

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369. https://doi.org/10.1016/J.JARMAC.2017.07.008

Linder, C., Östman, L., & Wickman, O. (Eds.). (2007). *Promoting scientific literacy: Science education research in transaction*. Linnaeus.

Mahr, D., Göbel, C., Irwin, A., & Vohland, K. (2018). Watching or being watched: Enhancing productive discussion between the citizen sciences, the social sciences, and the humanities. In S. Hecker, M. Haklay, A. Bowser, Z.

Makuch, J. Vogel, & A. Bonn (Eds.), *Citizen science: Innovation in open science, society and policy* (pp. 99–109). UCL Press.

Masters, K., Oh, E. Y., Cox, J., Simmons, B., Lintott, C., Graham, G., Greenhill A., & Holmes, K. (2016). Science learning via participation in online citizen science. *Journal of Science Communication*, *15*(3). https://doi.org/10.22323/2.15030207

McCrudden, M. T., & Rapp, D. N. (2017). How visual displays affect cognitive processing. *Educational Psychology Review*, *29*(3), 623–639. https://doi.org/10.1007/s10648-015-9342-2

Murcia, K. (2009). Re-thinking the development of scientific literacy through a rope metaphor. *Research in Science Education*, *39*(2), 215–229. https://doi.org/10.1007/s11165-008-9081-1

National Academies of Sciences Engineering and Medicine. (2017). *Communicating science effectively*. https://doi.org/10.17226/23674

National Academies of Sciences Engineering and Medicine. (2016). *Science literacy: Concepts, contexts, and consequences* (C. E. Snow & K. A. Dibner, Eds.). https://doi.org/10.17226/23595

National Research Council. (2014). *Developing assessments for the next generation science standards*. https://doi.org/10.17226/18409

National Science Board. (2018). *Science and engineering indicators 2018*.

Orr, D., & Baram-Tsabari, A. (2018). Science and politics in the polio vaccination debate on Facebook : A mixed-methods approach to public engagement in a acience-based dialogue. *Journal of Microbiology & Biology Education*, *19*(1), 1–8. https://doi.org/10.1128/jmbe.v19i1.1500

Peters, M. A. (2018). Education in a post-truth world. In M. A. Peters, S. Rider, M. Hyvönen, & T. Besley (Eds.), *Post-truth, fake news* (pp. 145–150). https://doi.org/10.1007/978-981-10-8013-5_12

Ryder, J. (2001). Identifying science understanding for functional scientific literacy. *Studies in Science Education*, *36*(1), 1–44. https://doi.org/10.1080/03057260108560166

Scharrer, L., Rupieper, Y., Stadtler, M., & Bromme, R. (2017). When science becomes too easy: Science popularization inclines laypeople to underrate their dependence on experts. *Public Understanding of Science*, *26*(8), 1003–1018. https://doi.org/10.1177/0963662516680311

Shauli, S., & Baram-Tsabari, A. (2018). The usefulness of science knowledge for parents of hearing-impaired children. *Public Understanding of Science*. https://doi.org/10.1177/0963662518772503

Soleri, D., Long, J. W., Ramirez-Andreotta, M., Eitemiller, R. D., & Pandya, R. (2016). Finding pathways to more equitable and meaningful public-scientist partnerships. *Citizen Science: Theory and Practice*, *1*(1), 9–11. https://doi.org/10.5334/cstp.46

Tufte, E. (2006). *Beautiful evidence*. Graphics press.

Whitmarsh, L. (2011). Scepticism and uncertainty about climate change: Dimensions, determinants and change over time. *Global Environmental Change*, *21*(2), 690–700. https://doi.org/10.1016/J.GLOENVCHA.2011.01.016

Wynne, B. (1991). Knowledges in context. *Science, Technology & Human Values*, *16*(1), 111–121. doi:10.1177/016224399101600108