



Flooded by jargon: how the interpretation of water-related terms differs between hydrology experts and the general audience

Gemma J. Venhuizen¹, Rolf Hut², Casper Albers³, Cathelijne R. Stoof⁴, and Ionica Smeets¹

¹Science Communication and Society, Leiden University, Leiden, the Netherlands

²Delft University of Technology, Delft, the Netherlands

³Heymans Institute for Psychological Research, University of Groningen, Groningen, the Netherlands

⁴Soil Geography and Landscape Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, the Netherlands

Correspondence: Rolf Hut (r.w.hut@tudelft.nl)

Received: 28 May 2018 – Discussion started: 25 June 2018

Revised: 17 December 2018 – Accepted: 19 December 2018 – Published: 22 January 2019

Abstract. Communication about water-induced hazards (such as floods, droughts or levee breaches) is important, in order to keep their impact as low as possible. However, sometimes the boundary between specialized and non-specialized language can be vague. Therefore, a close scrutiny of the use of hydrological vocabulary by both experts and laypeople is necessary.

In this study, we compare the expert and layperson definitions of 22 common terms and pictures related to water and water hazards, to see where misunderstandings might arise both in text and pictures. Our primary objective is to analyze the degree of agreement between experts and laypeople in their definition of the used terms. In this way, we hope to contribute to improving the communication between these groups in the future. Our study was based on a survey completed by 34 experts and 119 laypeople.

Especially concerning the definition of words related to water there are some profound differences between experts and laypeople: words like “river” and “river basin” turn out to have a thoroughly different interpretation between the two groups. Concerning the pictures, there is much more agreement between the groups.

the way these hazards are communicated to the general public. The availability of newspapers, magazines, television, radio and the internet has enabled more hydrogeocommunication, thus possibly contributing to a better informed society.

In particular, communication about water-induced hazards is becoming more and more important. A key aspect of increasing climate change is the expectation that natural hazards related to water, like floods and levee breaches, will occur more frequently in the future (Pachauri and Meyer, 2014).

Geoscientific studies (e.g., hydrological studies) are sometimes being ignored in policy and public action, partly because of the fact that scientists often use complicated language that is difficult to understand (Liverman, 2008). Other studies show that policy makers are more willing to take action if they understand why a situation could be hazardous (Forster and Freeborough, 2006). To be effective, early warning systems for natural hazards like floods need to focus on the people exposed to risk (Basher, 2006).

One way to improve communication with nonexperts is to avoid professional jargon (Rakedzon et al., 2017). However, sometimes the boundary between specialized and non-specialized language can be vague. Some terms are used both by experts and by laypeople, but in a slightly different way. A term like “flood” might not be considered jargon since it is quite commonly used, but could still have a different meaning in the scientific language than in day-to-day language.

In the health sciences, clear communication by doctors has been linked to better comprehension and recall by patients (Boyle, 1970; Hadlow and Pitts, 1991; Castro et al., 2007;

1 Introduction

Water-related natural hazards have impacted society throughout the ages. Floods, droughts and changing river patterns all had their influence on where and how people lived. One thing that has changed throughout the last centuries, however, is

Blackman and Sahebjalal, 2014). Similar benefits from effective communication can be expected in other scientific areas as well. An important factor is the degree to which people have the capacity to understand basic information – in the health sciences, this is referred to as health literacy (Castro et al., 2007) and in the geosciences as geoliteracy (Stewart and Nield, 2013). We prefer to avoid the term “literacy” in this article, since it is a limited way of addressing shared comprehension of science concepts (Kahan et al., 2012). We prefer to focus more on the divergent definitions of jargon.

In our research, we choose to study both the understanding of textual terms and the understanding of pictures. Some interesting work has been done about alternate conceptions in oceanography, focusing on students and using both textual and pictorial multiple-choice questions (Arthurs, 2016). Arthurs’ study also focuses on the topic of intermodality, i.e., switching between modes of communication (textual vs. pictorial).

However, no studies have been done about the extent to which geoscientists use jargon in interaction with the general audience (Hut et al., 2016). Therefore, a close scrutiny of hydrological vocabulary and the interpretation of common water-related terms by both experts and laypeople is necessary. In this article, we define “water related” as “associated with water and sometimes also with water hazards”.

Health scientific studies show that a significant difference in the interpretation of specific definitions (both in text and images) can be found between doctors and patients (Boyle, 1970). A similar difference between experts and laymen can be expected in the communication in other scientific areas, e.g., hydrology. Experts can be unaware of using jargon, or they may overestimate the understanding of such terminology by people outside their area of expertise (Castro et al., 2007).

Knowledge about which terms can cause misunderstanding could help hydrogeoscientists in understanding how to get their message across to a broad audience, which will benefit the public.

The word “jargon” derives from Old French (back then, it was also spelled as “jargoun”, “gargon”, “ghargun” and “gergon”) and referred to “the inarticulate utterance of birds, or a vocal sound resembling it: twittering, chattering”, as noted by Hirst (2003). In the same article, the author comes up with several general definitions of jargon, with the two main ones being (1) “the specialized language of any trade, organization, profession, or science”; and (2) “the pretentious, excluding, evasive, or otherwise unethical and offensive use of specialized vocabulary”. The first one can be considered neutral definition, the second one has a negative connotation (Hirst, 2003).

Within the geosciences, no specific definition of jargon is available. As noted by Somerville and Hassol (2011), scientists often tend to speak in “code” when communicating about geosciences to the general public. The authors refer in their article to climate change communication and encour-

age scientists to use simpler substitutes and plain language, without too much detail – as an example they suggest “human caused” instead of “anthropogenic”. However, they do not suggest a specific definition of jargon.

Nerlich et al. (2010) write that climate change communication (as part of geocommunication) shares features with various other communication enterprises, amongst which is health communication. Since there is no specific definition of jargon in geosciences and since the definitions by Hirst are very broad and not science-specific, we chose to adopt the definition from medical sciences (Castro et al., 2007) in which jargon is defined as both (1) technical terms with only one meaning listed in a technical dictionary and (2) terms with a different meaning in layperson contexts. In other words, jargon has a broader definition than some scientists think. It can be expected that hydrogeological terms sometimes have a less strict meaning for laypeople than for experts, meaning that hydrologists should be aware of this second type of jargon (Hut et al., 2016).

In this article, we compare the expert and layperson definitions of some common water-related terms, in order to assess whether or not these terms can be considered jargon and to see where misunderstandings might arise. With this goal in mind, we developed a questionnaire to assess the understanding of common water-related words by both hydrology experts and laypeople. Our primary objective is to analyze the degree of agreement between these two groups in their definition of the used terms. In this way, we hope to contribute to improving the communication between these groups in the future.

To our knowledge, no study has measured the agreement between experts and laypeople in understanding of common water-related terms. A matched vocabulary could increase successful (hydro)geoscientific communication.

2 Methodology

We started by analyzing the water-related terms frequently mentioned in the 12 Water Notes (European Commission, 2008). These notes contain the most important information from the European Water Framework Directive (European Commission, 2008), a European Union directive which commits European Union member states to achieve good qualitative and quantitative status of all water bodies. This was done by counting how often each term related to water appeared in the text. We chose these notes because they are a good representation of hydrogeocommunication from experts to laypeople: they are meant to inform laypeople about the Water Framework Directive. From this list, 20 of the most frequently used terms were chosen (10 of these were also present in the definition list of the Water Framework Directive itself), such as river, river basin, lake and flood. The questionnaire (including the chosen terms) can be found in the Supplement. Although the word “water” was the hydro-

logical term most frequently used in the notes, we decided to exclude this from the survey, because it is too generic a term.

A focus group was carried out at the American Geophysical Union fall meeting in San Francisco in December 2016, to check the list of terms and to come up with appropriate definitions. Eight participating hydrology experts were asked to describe the above-mentioned terms on paper, and to discuss the outcomes afterwards. The focus group consisted of experts, which mimics the process of science communication: the experts choose and use the definitions, which are then communicated to laypeople. This discussion was audio recorded, with consent of the participants. This focus group was important because we wanted to generate reasonable answers for our survey.

Ten of the terms that turned out to be too-Framework-Directive-specific (for example “transit waters”, which was not recognized as common hydrological language by the focus group participants) were left out of the survey. The 10 other terms, which generated some discussion (like whether the word “dam” only relates to man-made constructions), were deemed to be fit for the survey, because they were recognized by the experts as common water-related words. Two additional, less frequented terms (discharge and water table) were also chosen, based on the focus group. The focus was only on textual terms; the 10 pictorial questions (see below) were chosen by ourselves, based on water-related pictures we came across in various media outlets. The pictures were chosen by two of the authors: one of them a hydrologist, one of them a layperson in terms of hydrology.

2.1 Survey

Our survey contained 22 multiple-choice questions about commonly used terms by water experts. Twelve of these were textual questions: participants were asked to choose (out of four options; USGS Glossary, 2019; WMO Glossary, 2019) which answer described a specific term best, in their opinion. Ten of these were pictorial questions: participants were asked to choose (out of four options) which full-color photo depicted a specific term best, in their opinion. In addition, we asked some demographic data (gender, age, level of education, postcode area + country). The complete survey can be found in the Supplement. Pictures were found using the Wikimedia Commons feature. An example of both types of questions can be found in Fig. 1.

2.2 Participants

We developed a flyer with a link to the survey, which we handed out to experts at the international hydrology conference IAHS (International Association of Hydrological Sciences) in South Africa in July 2017. Furthermore, the link to the survey was sent via email to hydrology experts around the globe: members of the hydrology division of the European Geosciences Union and professional hydrologists (studying

(a) What is, in your opinion, the best definition of a dam?

- A. Barrier constructed across a valley to store water or raise the water level
- B. Barrier that prevents a river to flow into a lake
- C. Man-made, giant concrete structure to regulate water flow
- D. Man-made object to keep rivers or seas from overflowing land

(b) Which of the following photos is, in your opinion, the best depiction of a geyser?



Figure 1. Example of a textual multiple-choice question (a) and a pictorial question (b) from the survey.

for PhD or higher) at various universities. The total number of respondents from the experts was $n = 34$.

The laypeople were approached in a different way. In the first week of September 2017, one researcher went to Manchester to carry out the survey on various locations on the streets, to make sure that native English-speaking laypeople would participate. Manchester was chosen because it is a large city in the UK, meaning that it would be convenient to find participants from a general population who were also native English speakers. In total, the number of laypeople that were incorporated in the study was $n = 119$. In the initial Google form results, the number of laypeople was $n = 131$, but 22 participants were excluded because they did not fill out the electronic consent or because they accidentally sent the same electronic form twice or thrice (in that case, only one of their forms was incorporated in the study).

The participants could fill out the survey on an iPad. If there were more participants at the same time, one would fill the survey out on the iPad and the other ones filled out an A4-sized printed full-color handout. In this way, multiple participants could fill out the survey at the same time.

All participants, both experts and laypeople, were asked to fill out an electronic consent form stating that they were above 18 years of age and were not forced into participating. The questionnaire was of the forced-choice type: participants were instructed to guess if they did not know the answer.

2.3 Analysis

In order to detect definition differences between experts and laypeople, we wanted to analyze to what extent their answers differed from each other for each question. As pointed out before, it was not about giving the right or wrong answer, but about analyzing the match between the resemblance between the answering patterns of the laypeople and the experts.

For each term, the hypotheses were as follows:

H_0 : laypeople answer the question the same as experts.

H_1 : laypeople answer the question differently than experts.

A statistical analysis was carried out in *R* (R Core Team, 2017) by using Bayesian contingency tables. A contingency table displays the frequency distribution of different variables, in this case a two by four table showing how often which definition of a specific term was chosen by experts and laypeople.

For each term, the hypothesis is tested using a so-called Bayes factor (BF; computed using Morey and Rouder, 2015). A value of the $BF < 1$ is evidence towards H_0 : it is more likely that laypeople answer questions the same as experts than differently. A value of the $BF > 1$ is evidence towards H_1 : differences are more likely than similarities. The BF can be interpreted as the so-called likelihood ratio: a BF score of 2 means that H_1 is twice as probable as H_0 , given the data. $BF = 0.5$ means that H_0 is twice as probable as H_1 . For example, aquifer has a $BF = 7801$. This means it is almost 8000 times as probable with these data that there is indeed a difference between laypeople and experts in defining this term. As the values can become very large, one often interprets their logarithm instead.

The Bayes factors can be interpreted as follows (Jeffreys, 1961):

- $BF > 10$: strong evidence for H_1 against H_0 .
- $3 < BF < 10$: substantial evidence for H_1 against H_0 .
- $1/3 < BF < 3$: no strong evidence for either H_0 or H_1 .
- $1/10 < BF < 1/3$: substantial evidence for H_0 against H_1 .
- $BF < 1/10$: strong evidence for H_0 against H_1 .

An additional benefit of the use of Bayes factors is that, unlike their frequentist counterpart, no corrections for multiple testing are necessary (Bender and Lange, 1999).

In addition to a Bayes factor for the significance of the difference, we also calculated the “misfit”: the strength of the difference. The misfit was calculated by a DIF score (differential item functioning), in which $DIF = 0$ means perfect

match, and $DIF = 1$ means maximum difference. This DIF score was operationalized as

$$DIF = \sqrt{\frac{1}{2} \sum_{i=1}^4 (p_{E,i} - p_{L,i})^2}, \quad (1)$$

where $p_{E,i}$ is the proportion of experts choosing option i , and $p_{L,i}$ is the proportion of laypeople making that choice. Thus, DIF is based on a sum-of-squares comparison between the answer patterns of laypeople and experts.

Subsequently, we plotted the posterior distribution of DIF, for each term. This posterior distribution indicates the likelihood for a range of DIF scores, based on the observed data.

For example, if the answering pattern would be A: 50 %, B: 50 %, C: 0 % and D: 0 % for both the experts and the laypeople, there would be a perfect match ($DIF = 0$). The misfit was plotted in graphs, ranging from the largest to the smallest misfit. The higher the misfit, and the higher the BF, the more meaningful a difference between laypeople and experts. Low values of misfit indicate agreement between laypeople and experts. The R code and data used for the analyses are available from Albers et al. (2018).

3 Results

For the overall view of all the 22 terms (both texts and illustrations), there is extreme evidence for differences between laypeople and experts. This can be quantified by multiplying the BFs with each other, leading to a 10 log value of 33.50 (H_1 is approximately 3×10^{33} more probable than H_0).

However, this difference is only visible when looking at the textual questions, with a combined 10 log value of 46.14. For the pictorial questions, there is very strong evidence for the absence of differences, with a negative 10 log value of -12.63 .

Interestingly enough, there was a lot of internal disagreement for both experts and laypeople on the term “stream” (47 % agreement of experts on the most chosen answer, “C. Small river with water moving fast enough to be visible with the naked eye”; 37 % agreement of laypeople on the most chosen answer, “D. General term for any body of flowing water”) and on the picture of a sewer (56 % agreement of experts on answer D – see Supplement for the picture; 55 % agreement of laypeople on answer D).

Concerning the text questions, there was full agreement between the experts on “discharge” (100 % agreement, $N = 33$ answered B, $N = 1$ answered blank) and almost full agreement on “downstream” (97 % agreement, $N = 33$ answered D). This can be seen in Fig. 2 and Appendix B.

Concerning the pictures, there was full agreement between the experts on “geyser” (100 % agreement, $N = 34$ answered B) and on “river” (100 % agreement, $N = 34$ answered B). High levels of agreement were found on the pictures for “flood” (97 % agreement, $N = 33$ answered C),

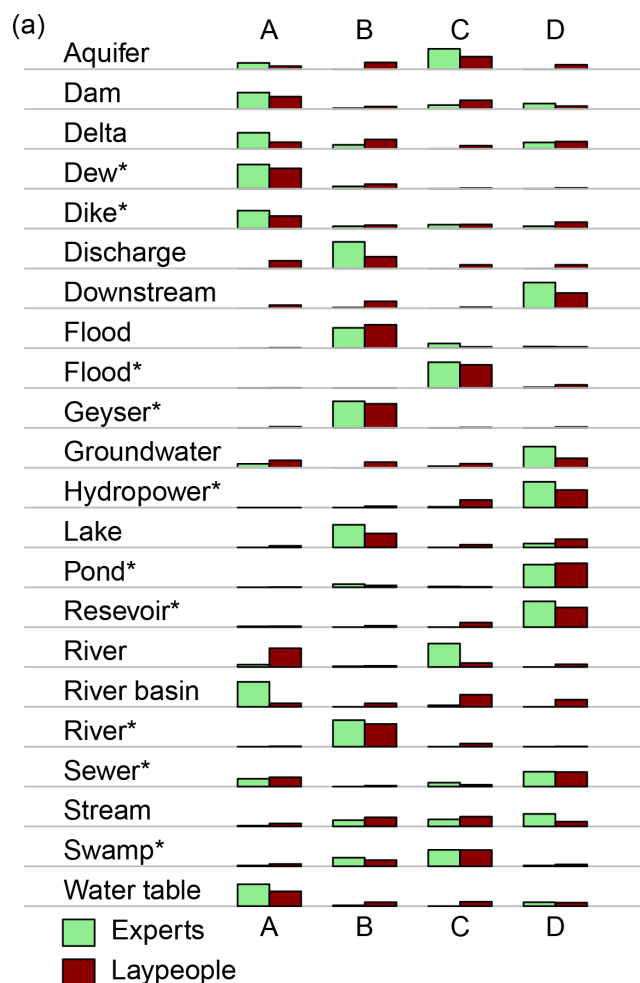


Figure 2.

“hydropower” (97 % agreement, $N = 33$ answered D) and “reservoir” (97 % agreement, $N = 33$ answered D). This can be seen in Fig. 2. The complete table with an overview of the multiple-choice answers (and the number of laypeople and experts that chose that specific answer) can be found in Appendix B.

3.1 Misfit between laypeople and experts

The most prominent misfit between laypeople and experts was found in the textual questions, for the definitions of river basin (10 log BF of 14.9), river (10 log BF of 11.9), discharge (10 log BF of 6.2), aquifer (10 log BF of 3.9) and groundwater (10 log of BF 3.4) (for more BF values, see table in Appendix A).

For these words, we have clear evidence that there is disagreement between experts and laypeople on the interpretation. This can be seen in Fig. 3. None of the pictorial questions made it to the top 10 of the biggest misfits. The pictorial questions that lead to the most prominent misfits were hydropower, reservoir, dike, sewer and swamp.

The broader and flatter the distribution, the stronger the Bayes factor. If both experts and laypeople have a high internal agreement (above 90 %) the misfit is smaller than if there is a lot of internal disagreement.

This can be seen in the graph: the posterior distribution of the misfit parameter is visible. It is important to note that under H_0 the misfit is not exactly equal to 0, because there is a certain degree of randomness. In other words, the misfit describes to what extent the answering patterns of the laypeople and the experts are similar to each other.

4 Discussion and conclusion

In total, we collected 119 questionnaires from native English-speaking laypeople and 34 questionnaires from (not necessarily native) English-speaking experts. Fifteen of the experts were native English (American) speakers (two others came from South Africa, where English is also a major language, two others did not fill this question out and the rest of the experts came from the Netherlands, Belgium, Germany, Turkey, Switzerland, Luxembourg, Brazil, France and Italy. All experts were of PhD level or above and were thus considered to have sufficient knowledge of the English scientific language. Nevertheless, two participants wrote in the comments that they found some of the terms difficult to understand due to the fact that they were nonnative English speakers.

This could be a limitation to our study, because possibly the nonnative English-speaking experts would have answered differently if they had been native English-speaking experts. However, since the majority of the experts ($n = 32$) did not have trouble understanding the questions (or at least did not write a comment about this), we do not consider this a major limitation and we did not exclude these experts because they did meet our criteria (PhD level or above).

Our definition from jargon, which is as mentioned before adopted from Castro et al. (2007), is not influenced by a distinction between native and nonnative English speakers. However, it can be expected that hydrogeological terms sometimes have a less strict meaning for nonnative English speakers in general, and especially for nonnative English-speaking laypeople, due to the difference in understanding between laypeople and experts (Hut et al., 2016). This is why we excluded nonnative English-speaking laypeople.

A disadvantage of the survey was that some of the text questions were still quite ambiguous. The interpretation of some terms changes depending on the context, the specific background and the exact definitions. Due to the limitations of a multiple-choice format, in some cases none of the definitions might seem to have a perfect fit, whereas with the pictures it is the other way around and sometimes more than one picture could fit a generic term. Giving only four pre-defined options could seem a bit misleading and restricted.

(b)

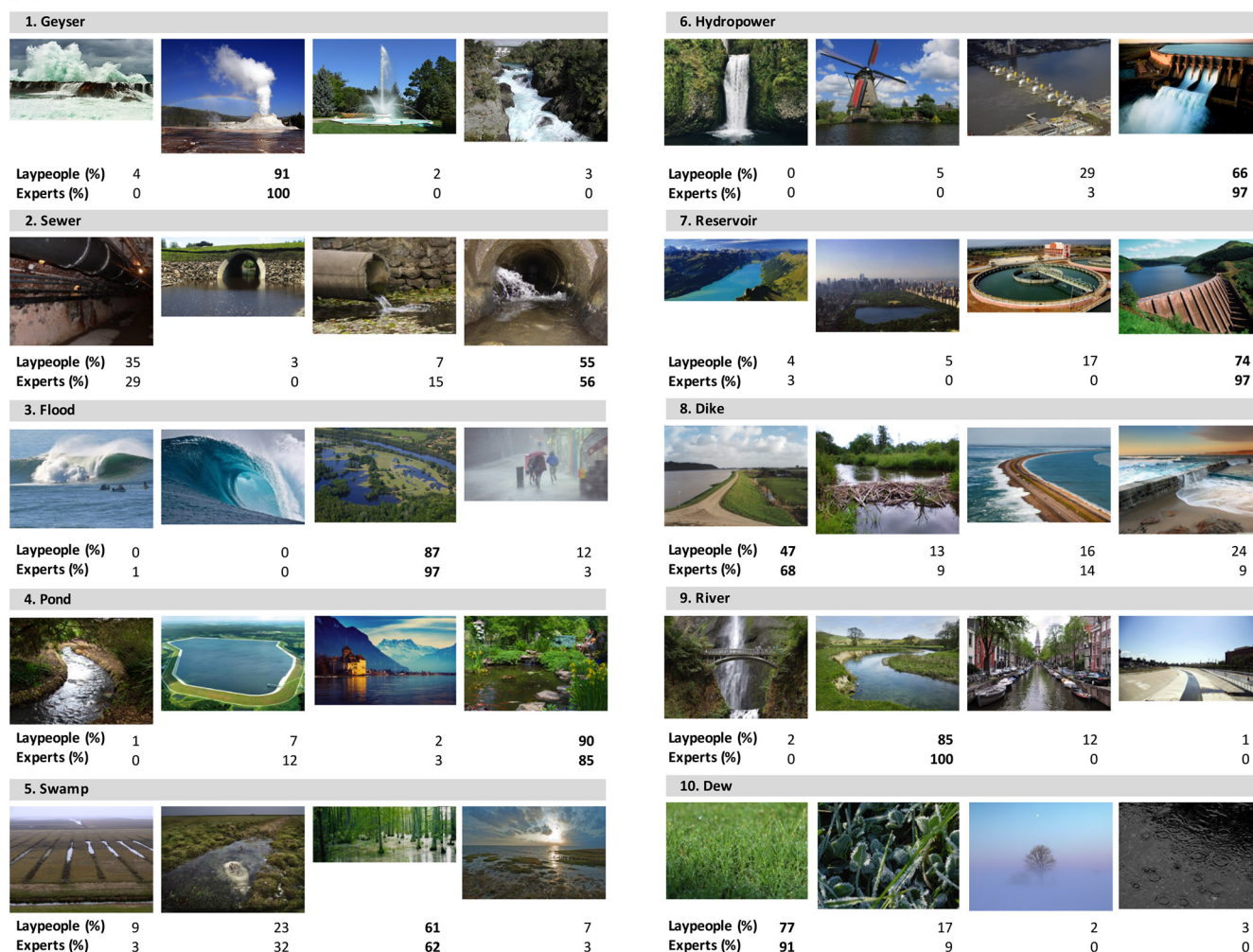


Figure 2. (a) Bar charts showing the answer distribution of both textual and pictorial questions (pictorial questions are marked with an asterisk, *). (b) Answer distribution of pictorial questions (the number of layperson respondents was 115 to 117: $N = 115$ for hydropower, reservoir; $N = 116$ for geyser, pond, swamp, dike, dew; $N = 117$ for sewer, flood, river.)

Moreover, nonnative speaking experts could be confused by some of the English definitions.

In this study, we have chosen to use terms as defined by experts, because it mimics the real-life situation in which scientists use specific terms for communication to a broader audience. As suggested by one of the reviewers, in future research it would be interesting to adopt a broader perspective by also incorporating terms as defined by laypeople. This could be done by organizing a focus group consisting of laypeople and discuss with them the meaning of specific terms.

Concerning the surveys of the laypeople, a disadvantage of the handouts was the fact that the pictures could not be enlarged. In addition, the prints were two-sided, and in some cases participants overlooked some of the questions. Even though the survey was of the forced type, not all people answered all the questions. As one of the reviewers suggested, in a next survey we could ask people to describe their expe-

riences with flooding – people who are familiar with water-related hazards may answer differently from people who do not have this experience.

The answering pattern within a group (laypeople or experts) could be inherent to the specific answers. In some cases, the answers were quite similar to each other; in other cases, the difference was quite big. However, this could not explain the misfit between laypeople and experts, since they both filled out the same survey.

We expected there would be no difference between people who filled out the survey on paper and people who filled out the survey on iPad. However, we did not test for this, so we cannot take into account any possible influences of the material used. This might be a topic for future research.

Of course, this research is only a first step in investigating the possibilities of a common vocabulary. By introducing our method to the scientific community (and making it accessi-

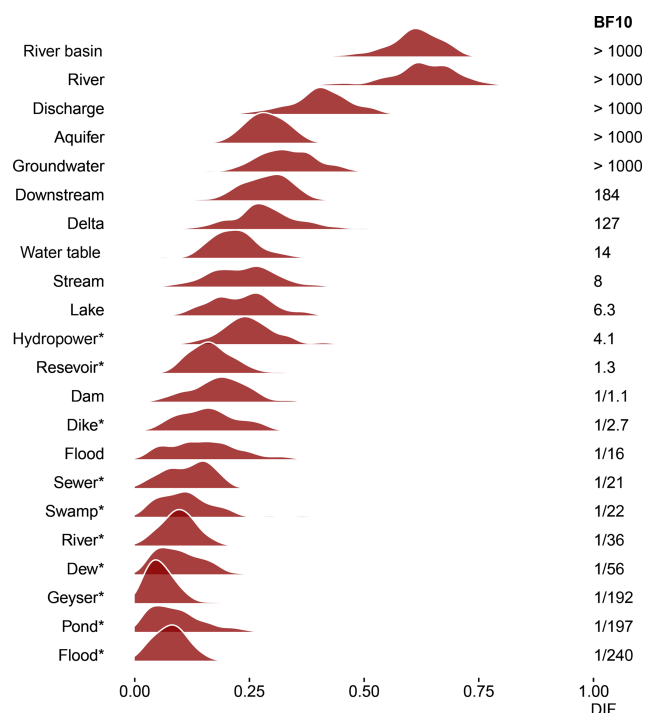


Figure 3. Graph showing the posterior distribution of the misfit between laypeople and experts by using a Bayes factor (BF) for every term used in the survey. Pictorial questions are marked with an asterisk. A value of the BF $< 1/10$ is strong evidence towards H_0 : it is more likely that laypeople answer questions the same as experts than differently. A value of the BF > 10 is strong evidence towards H_1 : differences are more likely than similarities. In addition to a Bayes factor for the significance of the difference, we also calculated the misfit: the strength of the difference. The misfit was calculated by a DIF score (differential item functioning), in which DIF = 0 means perfect match, and DIF = 1 means maximum difference.

ble via open access) we hope to encourage other scientists to carry out this survey with other terminology as well.

Since relatively little is known about the interpretation of jargon by laypeople and experts (especially in the natural sciences), additional research in this field is recommended.

Concluding, this study shows that there exists a strong difference between laypeople and experts in the definition of common water-related terms. This difference is more strongly present when the terms are presented in a textual way. When they are presented in a visual way, we have shown that the answer patterns by laypeople and experts are the same.

Therefore, the most important finding of this study is that pictures may be clearer than words when it comes to science communication around hydrogeology. We strongly recommend using relevant pictures whenever possible when communicating about an academic (hydrogeological) topic to laypeople.

Our findings differ from medical jargon studies which take into account both textual terms and images. For example, Boyle (1970) finds that there is a significant difference between doctors and patients when it comes to the interpretation of both terms and images. However, these images differed in various ways from the pictures in our study: they were hand drawn and only meant to indicate the exact position of a specific bodily organ.

What makes a “good” picture for science communication purposes would be an interesting topic for further research. Also, more research could be done on the textual terms: how could the existing interpretation gap between experts and laypeople be diminished? What impact would the combination of pictures and textual terms have – would the text enhance the pictures and vice versa? All in all, a broader research which incorporates more terminology and pictures (from various scientific disciplines) would be a very valuable starting point. Also, in line with Hut et al. (2016), it would be interesting to analyze the understanding of motion pictures (e.g., documentaries) in geoscience communication, while TV is a powerful medium.

Data availability. The anonymized result of the survey and an anonymized aggregated data file on which the analyses is based are available through the Open Science Framework. Both the data files and the computer code used to generate the figures and results in this work are persistently available through Albers et al. (2018, <https://doi.org/10.17605/OSF.IO/WK9S6>).

Appendix A

Table A1. Bayes factors (BFs) and their base-10 logarithms.

Term	BF	10 log BF
Aquifer	7.801e+03	3.892
River basin	7.428e+14	14.871
Dam	8.783e-01	−0.056
Delta	1.273e+02	2.105
Dew	1.798e-02	−1.745
Dike	3.685e-01	−0.434
Discharge	1.531e+06	6.185
Downstream	1.841e+02	2.265
Flood (text)	4.165e-03	−2.380
Flood (picture)	6.403e-02	−1.194
Geyser	5.209e-03	−2.283
Groundwater	2.418e+03	3.383
Hydropower	4.070e+00	0.610
Lake	6.324e+00	0.801
Pond	5.069e-03	−2.295
Reservoir	1.274e+00	0.105
River (text)	2.784e-02	−1.555
River (picture)	7.094e+11	11.851
Sewer	4.790e-02	−1.3197
Stream	8.046e+00	0.906
Swamp	4.601e-02	−1.337
Water table	1.360e+01	1.134

Appendix B

Table B1. Answer distribution for textual questions.

Term with possible definitions	Answer distribution (%)	
	Laypeople ^a	Experts ^b
1. River		
A. Path of fresh water flowing into the ocean	71	9
B. Water flowing only on the surface of the land and never underground	4	3
C. Large stream which serves as the natural drainage for a basin	15	88
D. Flow of surface water within a straight channel	10	0
2. River basin		
A. Area having a common outlet for its surface runoff	13	94
B. Dry river channel which may be flooded during high water events	13	0
C. Catchment which a river flows into	47	6
D. Body of water (lake, sea, ocean) a river flows into	27	0
3. Groundwater		
A. All water stored in the ground	28	15
B. All water which is in direct contact with the ground	21	0
C. Water flowing under ground	15	6
D. Subsurface water occupying the saturated zone	36	79

Table B1. Continued.

Term with possible definitions	Answer distribution (%)	
	Laypeople ^a	Experts ^b
4. Aquifer		
A. Subsurface water body	11	24
B. Groundwater that reaches the surface through a permeable rock layer	25	0
C. Geological formation capable of storing, transmitting and yielding water	47	76
D. Man-made structure first built by the Romans to transport water	17	0
5. Lake		
A. Man-made body of standing surface water of significant extent	6	0
B. Inland body of standing surface water of significant extent	53	85
C. Small body of water encompassed by high mountains	10	0
D. Area of variable size filled with water	31	15
6. Dam		
A. Barrier constructed across a valley to store water or raise the water level	47	62
B. Barrier that prevents a river to flow into a lake	9	3
C. Man-made, giant concrete structure to regulate water flow	33	15
D. Man-made object to keep rivers or seas from overflowing land	11	20
7. Delta		
A. Feature resulting from an alluvial deposit at a river mouth	25	61
B. River mouth that spreads out a little bit, like the shape of a Greek letter Delta	35	15
C. Triangular-shaped island in a river	12	0
D. Landform that forms from deposition of sediment carried by a river	28	24
8. Downstream		
A. Heavy intensity rainwater falling down	12	0
B. Direction from which a fluid is moving	26	3
C. Stream that branches off from the main stream	4	0
D. Direction in which a fluid is moving	58	97
9. Flood		
A. Large wave of moving water	2	0
B. Overflow of water onto lands that are not normally covered by water	88	76
C. Rise in the water level to a peak from which it recedes at a slower rate	5	18
D. Unusually large runoff event that leads to economic damage	5	6
10. Stream		
A. River that drains into another river and not into a lake, sea or ocean	11	3
B. Watercourse that flows into a larger watercourse or into a lake	34	24
C. Small river with water moving fast enough to be visible with the naked eye	37	26
D. General term for any body of flowing water	18	47
11. Discharge		
A. Volume of water that passes through the whole river in 1 day	29	0
B. Volume of water flowing through a river cross section per unit time	45	100
C. Water with enough sediment in it to limit visibility to less than 1 feet	13	0
D. Flowing water in a reservoir used to generate electricity	13	0
12. Water table		
A. Top surface of the zone of saturation	56	82
B. Saturated part of an aquifer	15	3
C. Tide table kept at water authority	16	0
D. Height to which water raises in a well	13	15

^a The number of layperson respondents varied from 115 to 119: $N = 115$ for aquifer, water table; $N = 116$ for lake, delta; $N = 117$ for stream; $N = 118$ for river basin, groundwater, dam, downstream, flood, discharge; $N = 119$ for river. ^b The number of experts respondents was $N = 33$ for delta and discharge and $N = 34$ for all other terms.

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/hess-23-393-2019-supplement>.

Author contributions. We use the CRediT taxonomy defining scientific roles to specify the author contributions. The definition of these roles can be found at <https://casrai.org/credit/> (last access: 8 January 2019). GJV performed the following roles: conceptualization, methodology, data curation, investigation (survey), and writing – original draft. RH performed the following roles: conceptualization, methodology, investigation (focus group), and writing – review & editing. CA performed the following roles: formal analysis, methodology, software, visualization, and writing – review & editing. CRS performed the following roles: conceptualization, visualization, and writing – review & editing. IS performed the following roles: conceptualization, methodology, funding acquisition, and writing – review & editing.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Our special thanks goes to Sam Illingworth, senior lecturer in Science Communication at Manchester Metropolitan University, for contributing to this project in various ways: from thinking along about the questionnaire to helping out with the logistics while carrying out the survey in Manchester.

Also, we would like to express our thanks to all the participants in the survey and to the members of the AGU focus group.

Finally, we would like to express our thanks to the reviewers, Hazel Gibson (University of Plymouth) and Charlotte Kämpf (Karlsruhe Institute of Technology), who were so kind to improve our paper with their helpful comments, and to Erwin Zehe (Karlsruhe Institute of Technology) as the editor of this paper.

Cathelijne R. Stoof has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 706428

Gemma J. Venhuizen has received funding from the Gratama Stichting and from the Leids Universiteits Fonds.

Edited by: Erwin Zehe

Reviewed by: Hazel Gibson and Charlotte Kaempf

References

- Albers, C., Venhuizen, G. J., Hut, R., Smeets, I., and Stoof, C. R.: Flooded by jargon, <https://doi.org/10.17605/OSF.IO/WK9S6>, 2018.
- Arthurs, L.: Assessing Student Learning of Oceanography Concepts, *Oceanography*, 29, 18–21, <https://doi.org/10.5670/oceanog.2016.68>, 2016.
- Basher, R.: Global early warning systems for natural hazards: systematic and people-centred, *Philos. T. R. Soc. A*, 364, 2167–2182, <https://doi.org/10.1098/rsta.2006.1819>, 2006.
- Bender, R. and Lange, S.: Adjusting for multiple testing-when and how?, *J. Clin. Epidemiol.*, 54, 343–349, [https://doi.org/10.1016/S0895-4356\(00\)00314-0](https://doi.org/10.1016/S0895-4356(00)00314-0), 1999.
- Blackman, J. and Sahebjalal, M.: Patient understanding of frequently used cardiology terminology, *British Journal of Cardiology*, 21, 102–106, <https://doi.org/10.5837/bjc.2014.007>, 2014.
- Boyle, C. M.: Difference Between Patients' and Doctors' Interpretation of Some Common Medical Terms, *Br. Med. J.*, 2, 286–289, 1970.
- Castro, C. M., Wilson, C., Wang, F., and Schillinger, D.: Babel babble: physicians' use of unclarified medical jargon with patients, *Am. J. Health Behav.*, 31, S85–S95, <https://doi.org/10.5555/ajhb.2007.31.supp.S85>, 2007.
- European Commission: Water Notes, available at: http://ec.europa.eu/environment/water/participation/notes_en.htm (last access: 8 January 2019), 2008.
- European Water Framework Directive: available at: http://ec.europa.eu/environment/water/water-framework/index_en.html, last access: 8 January 2019.
- Forster, A. and Freeborough, K.: A guide to the communication of geohazards information to the public, British Geological Survey, Nottingham, Urban Geoscience and Geohazards Programme, Internal Report IR/06/009, 2006.
- Hadlow, J. and Pitts, M.: The understanding of common health terms by doctors, nurses and patients, *Soc. Sci. Med.*, 1991, 193–196, 1991.
- Hirst, R.: Scientific jargon, good and bad, *J. Tech. Writ. Commun.*, 33, 201–229, 2003.
- Hut, R., Land-Zandstra, A. M., Smeets, I., and Stoof, C. R.: Geoscience on television: a review of science communication literature in the context of geosciences, *Hydrol. Earth Syst. Sci.*, 20, 2507–2518, <https://doi.org/10.5194/hess-20-2507-2016>, 2016.
- Jeffreys, H.: Theory of probability. Oxford, UK, Oxford University Press, 1961.
- Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Larrimore Ouellette, L., Braman, D., and Mandel, G.: The polarizing impact of science literacy and numeracy on perceived climate change risks, *Nature Climate Change*, 2, 732–735, 2012.
- Liverman, D. G. E.: Environmental geoscience – communication challenges, Geological Society, London, Special Publications, 305, 197–209, <https://doi.org/10.1144/SP305.17>, 2008.
- Morey, R. D. and Rouder, J. N.: BayesFactor: Computation of Bayes Factors for Common Designs, available at: <https://CRAN.R-project.org/package=BayesFactor> (last access: 8 January 2019), 2015.
- Nerlich, B., Koteyko, N., and Brown, B.: Theory and language of climate change communication, *Wiley Interdisciplinary Reviews, Climate Change*, 1, 97–110, 2010.
- Pachauri, R. K. and Meyer, L. A. (Eds.): IPCC Climate Change 2014 Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, IPCC, Geneva, Switzerland, 2014.
- Rakedzon, T., Segev, E., Chapnik, N., Yosef, R., and Baram-Tsabari, A.: Automatic jargon identifier for scientists engaging with the public and science communication educators, *PLoS ONE*, 12, e0181742, <https://doi.org/10.1371/journal.pone.0181742>, 2017.
- R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, available at: <https://www.R-project.org/> (last access: 8 January 2019), 2017.

- Somerville, R. C. and Hassol, S. J.: Communicating the science of climate change, *Physics Today*, 64, 48, <https://doi.org/10.1063/PT.3.1296>, 2011.
- Stewart, I. S. and Nield, T.: Earth stories: context and narrative in the communication of popular geoscience, *P. Geologist. Assoc.*, 124, 699–712, <https://doi.org/10.1016/j.pgeola.2012.08.008>, 2013.
- USGS Glossary: available at: <https://water.usgs.gov/edu/dictionary.html>, last access: 8 January 2019.
- WMO Glossary: available at: http://www.wmo.int/pages/prog/hwarp/publications/international_glossary/385_IGH_2012.pdf, last access: 8 January 2019.