

# Science, Maddá, and ‘Ilm: The language divide in scientific information available to Internet users

Public Understanding of Science

1–17

© The Author(s) 2021



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/09636625211022975

[journals.sagepub.com/home/pus](https://journals.sagepub.com/home/pus)**Kawther Zoubi\*****Aviv J. Sharon\*** 

Technion—Israel Institute of Technology, Israel

**Eyal Nitzany** 

Independent Scholar, Israel

**Ayelet Baram-Tsabari** 

Technion—Israel Institute of Technology, Israel

## Abstract

The Internet has potential to alleviate inequality in general and specifically with respect to science literacy. Nevertheless, digital divides persist in online access and use, as well as in subsequent social outcomes. Among these, the “language divide” partly determines how successful users are in their Internet use depending on their proficiency in languages, and especially in English. To examine whether the quality of online scientific information varies between languages when conducting searches from the same country, we compared online search results regarding scientific terms in English, Hebrew, and Arabic. Findings indicate that searches in English yielded overall higher quality results, compared with Hebrew and Arabic, but mostly in pedagogical aspects, rather than scientific ones. Clustering the results by language yielded better separation than clustering by scientific field, pointing to a “language divide” in access to online science content. We argue that scientific communities and institutions should mitigate this language divide.<sup>1</sup>

## Keywords

digital divide, digital inequality, language divide, science communication, science literacy

---

\*Equal contribution

<sup>1</sup>*Maddá* and *‘Ilm* mean “science” in Hebrew and Arabic, respectively.

## Corresponding author:

Ayelet Baram-Tsabari, Faculty of Education in Science and Technology, Technion—Israel Institute of Technology, Technion City, Haifa 3200003, Israel.

Email: [ayelet@technion.ac.il](mailto:ayelet@technion.ac.il)

Nur and Talia are high school students who live in the same apartment complex in Haifa, Israel. Nur is a native Arabic speaker who is proficient in Hebrew and English, and Talia is bilingual in Hebrew and English. One day, they read a newspaper article about technological food innovations, such as plant-based meat alternatives. They go online to learn more about the science behind these inventions, but a quick search reveals that popular science content on this topic is mostly available in English, whereas the top search results in Hebrew and Arabic typically offer advertisements, recipes, and technical information for professionals. As they run more searches, they get a stronger impression that the usefulness of the results differs depending on the language they search in.

This fictional vignette reflects a broader issue: Science literacy is considered to benefit the health and well-being of individuals, communities, and society (National Academies of Sciences Engineering and Medicine, 2016). As access to the Internet increases globally, it has the potential to alleviate social inequalities, for example, by increasing access to useful scientific information. However, legacy inequalities remain with us within and between nations of the world, partly due to differences in language proficiency.

## Literature review

### *Science literacy in individuals, communities, and societies*

In recent decades, one of the prime goals of science education has been to prepare non-scientists to make sense of science in their everyday lives, be critical consumers of science information, and make informed decisions about scientific issues (Roberts and Bybee, 2014). The goal of inclusive science education for non-scientists is termed “science literacy.” The US National Academies’ report *Science Literacy: Concepts, Contexts, and Consequences* (National Academies of Sciences Engineering and Medicine, 2016) lists four broad rationales for science literacy: economic (remaining competitive in a global market), cultural (celebrating a human achievement that has changed our understanding of the world), democratic (informing the citizenry to enable democracies to function better), and personal (“[helping] people respond to issues and challenges that emerge in their personal and community context,” p. 24). These issues may include deciding whether to vaccinate a child, to consume genetically modified foods, or to adopt a local policy to mitigate and adapt to climate change. According to this rationale, science and technology (S&T) are intertwined with the lives of people in the 21st century, and therefore, understanding some science and having an ability to engage with it helps people make better decisions and engage in informed actions that lead to richer and healthier lives.

Historically, most conceptualizations of science literacy have focused on the *individual* competencies needed to achieve these outcomes. However, over time, scholars have suggested that science literacy should be viewed as a collective praxis, based on case studies of communities that have developed a shared understanding of science (e.g. Roth and Lee, 2002). Drawing on additional evidence from social activist movements (Brown, 1993; Epstein, 1995), and from the field of citizen science, the National Academies’ consensus report on science literacy suggested that “communities can possess and use science literacy to achieve their goals and may also contribute to new science knowledge in doing so” (National Academies of Sciences Engineering and Medicine, 2016: 83). The consensus report refers to this as *community-level science literacy*.

The National Academies’ consensus report was also influenced by work in the field of health literacy indicating society-level effects beyond the effects of the community. Health literacy studies have conceptualized health literacy as a property not just of the individual, but also of their social and physical context as well, since the demands and complexities of institutions and systems can support or inhibit individual decision-making and actions (Pleasant et al., 2016; Rudd et al.,

2012); for example, “living in a food desert [an area that has limited access to affordable and nutritious food] impairs the ability of an individual to gain access to healthy food, regardless of how much they know about the importance of vegetables” (National Academies of Sciences Engineering and Medicine, 2016: 18). Another example is that of lack of Internet access. Drawing on structural perspectives in sociology, the National Academies’ consensus report suggests that structural factors, such as formal policies and institutions, social and economic stratification, and others, “shape (if not determine) the distribution of science literacy for the communities and individuals therein” (p. 71). The report refers to this as *society-level science literacy*, and notes that very little research has been conducted on science literacy from a structural perspective.

### *The digital divide and the language divide*

The Internet is a major source of information about S&T in developed countries. As of 2018, 57% of US adults cite the Internet as their primary source of S&T information, and 70% say they would go online to find information about a specific S&T issue (National Science Board, 2020). Similarly, in Israel, 77% of adults who mentioned that they were interested in at least one field of S&T cited search engines as a primary source of S&T information (Israel Ministry of Science Technology and Space, 2017).

The Internet offers many affordances for public engagement with science. Unfortunately, not everyone equally benefits from access to information on the Internet, S&T-related, or otherwise, due to disparities collectively named “the digital divide.” This term refers to “any divide or gap between people [ . . . ] in their communication technology awareness, adoption or ownership, use, and skill” (Pearce and Rice, 2014). Research on the digital divide has focused on three topics: physical access (e.g. in terms of hardware and connectivity; the “first-level” digital divide), Internet use (the “second-level” digital divide), and subsequent outcomes (the “third-level”), such as health and educational outcomes (Hargittai and Hsieh, 2013; Robinson et al., 2020).

The literature shows that as Internet access increases worldwide, individuals from higher socio-economic strata tend to benefit from it more than others, since they tend to possess higher levels of skill and social support (OECD, 2015). Similarly, the digital divide is also associated with other social inequalities along lines of “gender, sexuality, race and ethnicity, aging, disability, healthcare, education, rural residency,” and more (Robinson et al., 2020: 1). Thus, in many countries, there are large disparities in early exposure to computers depending on socio-economic status and gender (OECD, 2015). Moreover, a global digital divide is associated with disparities in countries’ wealth, political systems, telecommunication policies, and more (Hargittai and Hsieh, 2013). Segev and Ahituv (2010) have pointed out that the digital divide can manifest itself in terms of the volume of information available and of the competence to derive relevant information and use it skillfully, and found differences in the latter competence between countries based on the most popular search queries used in *Google Search* and *Yahoo!* in 2004–2005. It has been argued that digital inequalities have been exacerbated by the COVID-19 pandemic, as the “digitally disadvantaged” are less able to take advantage of eHealth services and remote learning (Robinson et al., 2020).

One key aspect of the digital divide is the “language divide” in Internet adoption and use (De Jesus and Xiao, 2012), which derives from dominance of a small number of languages on the Internet, with the most dominant one being English. The hegemony of English on the Internet has long been theorized as a barrier to Internet adoption and use in a linguistically diverse world (Chen and Wellman, 2004; Warschauer, 2002). Relatively few studies have investigated this issue, but they indicate that Internet use is correlated with English proficiency in diverse contexts, including Italy and India (reviewed by Pearce and Rice, 2014) and among the Hispanic population in the United States (De Jesus and Xiao, 2012). Similarly, in Israel, language proficiency “explains ethnic

differences in Internet usage as a whole, and, more specifically, in human capital-enhancing Internet use” (Lissitsa and Chachashvili-Bolotin, 2014: 9). Thus, even if users have access to the Internet and have the skills to obtain information, if they cannot obtain information in their language, they are less likely to benefit from it.

### *The intersection of inequality in science literacy and the digital divide*

Based on studies on the digital divide, it seems likely that increasing access to science-related information online material may perpetuate or even exacerbate inequality in science literacy. With respect to the second-level digital divide, the National Academies’ report expresses concern that this may be the case due to “differences in the way that people are supported in their use of Internet technologies” (National Academies of Sciences Engineering and Medicine, 2016: 107). However, this topic has been relatively understudied. Hence, there is both a theoretical and a practical motivation to characterize the S&T-related information available to users in different languages.

## **Research goal and questions**

Our goal is to examine the characteristics and quality of online scientific information and compare them across three languages: English, Hebrew, and Arabic, and across three fields (disciplines): physics, chemistry, and biology. Specifically, we ask: How does the quality of online scientific information concerning core concepts in biology, chemistry, and physics differ when comparing languages and when conducting searches from the same country?

## **Research context**

This study focuses on scientific content in the Hebrew and Arabic languages, compared with content available in English, the dominant language of the Internet (Pearce and Rice, 2014). Modern Hebrew is the official language of Israel and 49% of its population over 20 years old speaks it natively (2011 data). Most of the rest of the population is proficient in Hebrew; hence, it has approximately 4.5 million native speakers and 9 million total speakers (2019 data). Most native Hebrew speakers are Jewish citizens of Israel (Israel Central Bureau of Statistics, 2013, 2019).

By contrast, Arabic has semi-official status in Israel with a large minority of native speakers (18% of the population over 20 years old in 2011 and 21% of the entire population in 2019). Thus, it has approximately 1.9 million native speakers in Israel, as of 2019. While Arabic is a minority language in Israel, it is an official language in 27 other countries and is spoken by roughly 274 million people worldwide, making it one of most widely spoken languages in the world (Eberhard et al., 2020).

Nevertheless, the Arabic language is under-represented on the Internet. Although 5.2% of Internet users are Arabic speakers (Miniwatts Marketing Group, 2020), only approximately 1% of Internet websites are in Arabic, a proportion that is only about twice as large as that of Hebrew websites (0.4%). By comparison, the share of websites in English on the Internet is 60.5% (Q-Success, 2020). The under-representation of Arabic can be partly explained by the fact that Arab countries have been relatively late adopters of the Internet (Warf and Vincent, 2007). Rates of Internet usage still vary considerably between these countries, as over 90% of the population uses the Internet in countries such as Saudi Arabia and the UAE, compared with 34% in Syria and 31% in Sudan. Overall, 47% of individuals in Arab states use the Internet. By comparison, Israel has an 82% Internet usage rate (2017 data; International Telecommunication Union (ITU), 2021).

Several studies point at the existence of a second-level digital divide between Jewish and Arab citizens in Israel. The Program for the International Assessment of Adult Competencies (PIAAC) study found that 34% of Arab citizens aged 16–65 have poor proficiency in accessing, analyzing, and communicating information using common computer applications, compared with 9% of Jewish citizens in the same age range (Israel Central Bureau of Statistics and Israel National Authority for Measurement and Evaluation in Education, 2016). Within Israel, Arab Internet users also report that they use the Internet to search for information less often than Jewish Internet users (46% vs 79%, respectively; Lissitsa, 2015). In addition, within Arab society in Israel, Hebrew and English proficiency correlates with capital-enhancing uses of the Internet, such as searching for information (Lissitsa, 2015). In two surveys conducted between 2011 and 2014, between 61 and 68% of Arab Internet users reported that they preferred reading Arabic-language websites, whereas 25–28% preferred Hebrew-language websites (Ganayem, 2018).

## Methods

### *Sampling search terms*

To measure the quality of scientific information online, a list of scientific terms in three languages was constructed. We focused on core scientific terms, such as “gravity” and “DNA,” to facilitate the comparison between the languages, as we hypothesized that they would be familiar to the average Internet user, that they would have standard translations in all the languages we studied and that there would be a wealth of content about them, compared with terms referring to contemporary science, such as “CRISPR-Cas9.” We also focused on the extent to which the content catered to young learners’ engagement with science because we hypothesized that early exposure to scientific content in one’s own language could contribute to social outcomes later in life.

Hence, the list was constructed in four steps: (1) collection of core scientific terms from school science curricula and from relevant research literature; (2) validation using a panel of secondary school science teachers; (3) translation to English and Arabic; and (4) refinement based on the search results.

First, we collected 365 terms in Hebrew from several sources, including secondary school physics, chemistry, and biology curricula and science content standards from the United States, Israel, and Egypt. We also included terms from scholarly articles about children’s interest in science (Baram-Tsabari and Yarden, 2005) and about public engagement with science online (Segev and Sharon, 2017).

Second, for the validation step, we assembled a panel of nine secondary school science teachers, all native Arabic speakers with professional working proficiency in Hebrew and English. The panel consisted of three smaller panels of three teachers each, for physics, chemistry, and biology. Each panelist held at least a bachelor’s degree in a scientific discipline or in science teaching, and most (seven out of nine) held an advanced degree as well. In addition, each panelist had at least 10 years of teaching experience. The panel members were asked to select the ten most central terms to their scientific domain derived from the list generated in the previous step, with special preference to terms that they considered relevant to everyday life. The panel discussions yielded a list of 30 terms, consisting of ten terms from each scientific domain (Table 1).

Third, we translated the 30 terms to English and Arabic. Since translation often yielded several possibilities, the translations were validated using the multilingual online encyclopedia, Wikipedia. The Hebrew terms were entered into the Hebrew-language edition of Wikipedia, and then equivalent terms in English and Arabic were chosen using the interlanguage links, which point from one article to its equivalent articles in other editions of the encyclopedia. Arabic translations were also

**Table 1.** List of scientific concepts in three languages and three fields.

Field	Item No.	English	Hebrew	Arabic
<b>Physics</b>	1	Electrical insulator	מבודד חשמלי	عازل كهربائي
	2	Time (Physics)*	זמן (פיזיקה)*	زمن (فيزياء)*
	3	Voltage	מתח חשמלי	جهد كهربائي
	4	Electrical network	מעגל חשמלי	دائرة كهربائية
	5	X-ray	קרינת רנטגן	أشعة سينية
	6	Light spectrum	ספקטרום האור	طيف الضوء
	7	Gravity	כבידה	جاذبية
	8	Density	צפיפות	كثافة
	9	Radiation	קרינה	اشعاع
	10	Velocity	מהירות ממוצעת	سرعة متجهة
<b>Chemistry</b>	11	Mass	מסה	كتلة
	12	Volume (chemistry)*	נפח (כימיה)*	حجم (كيمياء)*
	13	State of matter	מצב צבירה	حالة المادة
	14	Gas	גזים	غاز
	15	Liquid	נוזלים	سائل
	16	Chemical elements	יסודות כימיים	العناصر الكيميائية
	17	pH	רמת חומציות	درجة الحموضة
	18	Mixture (chemistry)*	תערובת (כימיה)*	مخلوط (كيمياء)*
	19	Pressure (chemistry)*	לחץ (כימיה)*	ضغط (كيمياء)*
<b>Biology</b>	20	Ozone	אוזון	أوزون
	21	Carbohydrate	פחמימה	سكريات
	22	Fat	שומן	دهن
	23	Protein	חלבון	بروتين
	24	Cell (biology)*	תא (ביולוגיה)*	خلية (علم الأحياء)*
	25	Homeostasis	הומיאוסטזיס	اتزان بدني
	26	DNA	דנ"א	د.ن.أ
	27	Metabolism	מטבוליזם	أيض
	28	Genetic disorder	מחלה תורשתית	مرض وراثي
	29	Enzyme	אנזים	انزيم
	30	Menstrual cycle	המחזור החודשי	الدورة الشهرية

\*The names of the scientific domains were added to the search term in parentheses to obtain more relevant results.

validated with the teacher panels to verify alignment with common usage among Arabic speakers in Israel and within the Arabic version of the Israeli school science curriculum. Hence, for example, the term selected for “pH” was *darajat al-humūḍa* (درجة الحموضة, “acidity level”) rather than the term used in the Arabic Wikipedia article title, *us hīdrūjīnī* (أس هيدروجيني, “power [exponent] of hydrogen”).

Fourth, a final refinement step was conducted to improve the relevance of the search results, in which the names of the scientific domains (“physics,” “chemistry,” and “biology”) were added to the search terms when they had non-scientific meanings in a certain language. For example, the term “volume” in English yielded results referring to both three-dimensional space and to sound pressure; the term *lāhats* in Hebrew (לחץ, “pressure”) yielded results relating to psychological

stress; and the term *makhlūt* in Arabic (مخلوط, “mixture”) yielded results relating to spice mixes and certain food dishes. Thus, if at least four out of the top seven results did not relate to the scientific aspect of the term, the names of the scientific domains were added to the search term in parentheses to obtain more relevant results. For example, the search term “time” was substituted with “time (physics)” across all three languages. This change was done for five search terms (time, volume, mixture, pressure, and cell; Table 1, items 2, 12, 18, 19, and 24).

**Limitations.** The reliance on Wikipedia led to some slightly different translations to English than anticipated, such as item 4 appearing in English as “electrical network” rather than the common term “electrical circuit,” which refers to just one type of network. Similarly, users searching for the Hebrew article for *maḥalá torashit* (מהלה תורשתית, “hereditary disease”; item 28) were redirected to the article titled *pgam genéti* (פגם גנטי, “genetic disorder”). Hence, the English term “genetic disorder” was included in the sample, rather than the direct translation, “hereditary disease.”

### Data collection and analysis

The search terms were entered into *Google Search* from the same computer using an Israel-based Internet connection in December 2018. In total, 630 results were obtained (30 terms  $\times$  3 languages  $\times$  7 results = 630 results). We took measures to avoid surveillance that could personalize the results, including using the browser in a private browsing mode, disabling Google Search customization settings, and deleting browser history before each search.

The scientific relevance of the first seven results was determined and recorded (Table 2, row 1); if all these results pertained to the scientific aspect of the term, they were included in the sample and analyzed (rows 2–12). This occurred in 66 of the 90 searches (73.3%). For the rest of the searches, any irrelevant results were disregarded. Subsequent *relevant* results were included instead, until seven results were reached per term.

The results were coded using a codebook developed based on 40 sources on evaluating electronic information quality in general, and in specific domains, such as health and nutrition. Some common variables include accuracy of the content, frequency of updates and maintenance, and whether the content is freely accessible (e.g. Guardiola-Wanden-Berghe et al., 2011; Savolainen, 2011; Shahbazi et al., 2019). The twelve variables, or metrics, that appeared in the codebook were grouped into three categories: (a) scientific quality (five variables), (b) pedagogical quality (five variables), and (c) quality metrics specific to online content (hereafter “online-specific” quality metrics; two variables).

To assess inter-rater reliability, the first author and a research assistant independently coded a sub-sample of 9.5% of search results ( $n=60$ ). Both are native speakers of Arabic and proficient in Hebrew and English. Cohen’s kappa values were over 0.9 for all but three variables: “Coverage” ( $\kappa > 0.7$ ), “Everyday Life” ( $\kappa > 0.8$ ), and “Last Updated” ( $\kappa > 0.8$ ).

We then turned to examine whether language ( $n_{\text{English}} = n_{\text{Hebrew}} = n_{\text{Arabic}} = 30$ ) and field ( $n_{\text{Physics}} = n_{\text{Chemistry}} = n_{\text{Biology}} = 30$ ) explained differences in information quality, as defined by the three categories of metrics. The first of the analyses we conducted are multivariate analyses of variance (MANOVAs). These analyses allowed us to compare these factors’ effect sizes, meaning the strength of the relationships between these factors and information quality. In addition, we sought to discern whether there were interaction effects between language and field. In other words, we wanted to know whether there are categories of metrics for which these factors have a joint effect that is significantly different than the sum of the parts. Three separate analyses of this type were conducted—one for each of the three categories of quality metrics separately.

**Table 2.** Codebook for assessing the quality of scientific information online.

Variable	Description	Range of possible values	Studies suggesting similar criterion
<b>Scientific quality</b>			
Scientific results <sup>a</sup>	Number of scientific results among the top seven results	4–7	
Accuracy <sup>b</sup>	Extent to which the content is free of scientific errors and imprecision	–1 to 2	Kapoun (1998), Livas et al. (2013), Oliver et al. (1997), Stvilia et al. (2007) and Tate (2019)
Coverage <sup>b</sup>	Comprehensiveness of the explanation	0–4	Ahn et al. (2007), Naumann and Rolker (2000), Oliver et al. (1997), and Savolainen (2011)
No. of sources cited <sup>b</sup>	Number of sources cited, where all values greater than one were recoded as one	0–1	Livas et al. (2013), Naumann and Rolker (2000), and Oliver et al. (1997)
Authority <sup>b</sup>	Whether the author's education and expertise are mentioned and relevant to the domain	0–3	Guardiola-Wanden-Berghe et al. (2011), Livas et al. (2013), and Stvilia et al. (2007)
<b>Pedagogical quality</b>			
Educational results <sup>b</sup>	Search result is an educational website designed for young learners	0 (No), 1 (Yes)	
Everyday life <sup>b</sup>	Search result draws a connection between the scientific term and everyday life	0 (No), 1 (Yes)	Dedeke (2000) and Naumann and Rolker (2000)
Illustration rating <sup>b</sup>	Use of relevant audio and/or visual materials, including animations, simulations, and video	0–3	Oliver et al. (1997) and Shahbazi et al. (2019)
Links to new concepts <sup>b</sup>	Search result contains links to new concepts or defines them	0 (No), 1 (Yes)	Cline and Haynes (2001), Guardiola-Wanden-Berghe et al. (2011), and Oliver et al. (1997)
Further reading <sup>b</sup>	Search result contains references for further reading	0 (No), 1 (Yes)	Cline and Haynes (2001) and Oliver et al. (1997)
<b>Variables specific to online content</b>			
Last updated <sup>b</sup>	Time that has passed since the content was created, in years, where items older than 3 years old were coded as 3 years old	0–3	Ahn et al. (2007), Kapoun (1998) and Savolainen (2011)
Interactivity <sup>b</sup>	Availability of options to contact the author(s)	0–2	Stvilia et al. (2007)

<sup>a</sup>Measured per search term.

<sup>b</sup>Measured per search result, values averaged per search term.

We then conducted 12 one-way ANOVAs, one for each dependent variable separately, to closely inspect the effects of language and field on information quality. This is a similar analysis to the MANOVA, except that each analysis was conducted for a single metric, rather than for an entire

category of metrics. To uncover specific differences between the means, these analyses were followed up by Tukey's post hoc tests (for comparisons that met the assumption of homogeneity of variances) or by Games–Howell post hoc tests (for the rest).

Next, a linear discriminant analysis (LDA) was performed. The goal in mind was to enable easier data visualization and prepare the data for subsequent analyses. This was established by reducing the dimension number from 12 (the number of ratings each search term had) to 2 ( $x, y$ ). LDA was selected as it is known to separate well between requested groups (e.g. search terms labeled as being in English, Hebrew, or Arabic).

Specifically, the LDA algorithm receives the search terms along with the values for each term's 12 metrics (e.g. "enzyme" along with its respective accuracy and coverage ratings, etc.) as well as each the term's group designation (i.e. which group each term belonged to, e.g. "English"). The algorithm then attempts to find a linear transformation of these values that would plot the search terms within each given group most closely together on a two-dimensional ( $x, y$ ) plane while maximizing the separability between the given groups. Grouping the observations in this way is called "clustering" the dataset; if the separability between the groups is high, the dataset is said to be "clustered well."

To accomplish this, the algorithm performed several steps. First, it calculated two axes, or linear discriminants, LD1 and LD2, that were each correlated with sets of the input characteristics: for example, LD1 was strongly correlated with the number of educational results and with authority ratings, whereas LD2 was strongly correlated with interactivity and Further Reading scores. Second, using these two axes, the algorithm then calculated scores (or coordinates) along these axes for each search term, based on its ratings. These scores determined where the point representing the search term was plotted on the two-dimensional ( $x, y$ ) plane. The LDA was run twice: once attempting to cluster search terms by languages and once by fields.

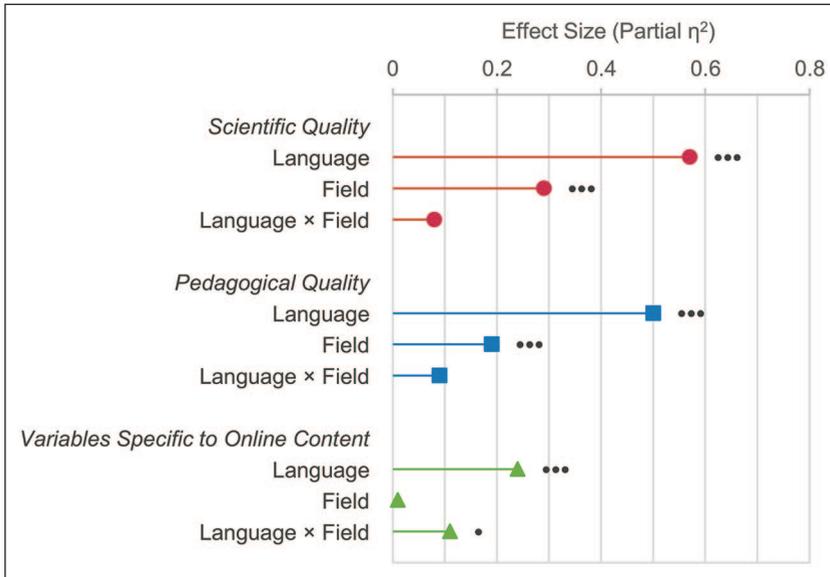
To evaluate whether terms had different characteristics and quality across languages, we then computed triangle areas on the LDA plane for each term, where the vertices of each triangle were the data points representing equivalent terms in each language (e.g. the data points representing "ozone" in English, Hebrew, and Arabic). The areas were computed using the Euclidean distances between the points. Thus, for example, if "ozone" received a small triangle area that would mean that the characteristics of the search results for "ozone" are similar across the three languages, given these specific LD1 and LD2 axes. Conversely, the larger the triangle area, the more different the search results for a given term are depending on the language one searches in, given these specific LD1 and LD2 axes. Here, we drew on similar methods used in the scholarly literature, albeit typically when the chosen technique was not LDA, but another commonly used method, named principal component analysis (e.g. Baram-Tsabari and Yarden, 2011; Huang et al., 2020; Paschou et al., 2007).

## Findings and discussion

### *Are there differences in quality between languages and fields?*

#### *Combined measures of quality*

*Effects of language and field on scientific and pedagogical quality.* By and large, the quality of search results differed both by language and by scientific field, but more so by language. This is evident from two-way MANOVA tests that found that both the language of the search terms ( $p < .001$ ) and their scientific fields ( $p < .001$ ) explained differences in the quality of search results, independently of each other, as regards scientific and pedagogical categories. No significant interaction between these factors was found for either of these combined metrics (Figure 1, Table S1).



**Figure 1.** Effects of language and field on combined measures of information quality—MANOVA results. One, two, and three bullets (\*, \*\*, and \*\*\*) denote statistical significance at the .05, .01, and .001 levels, respectively.

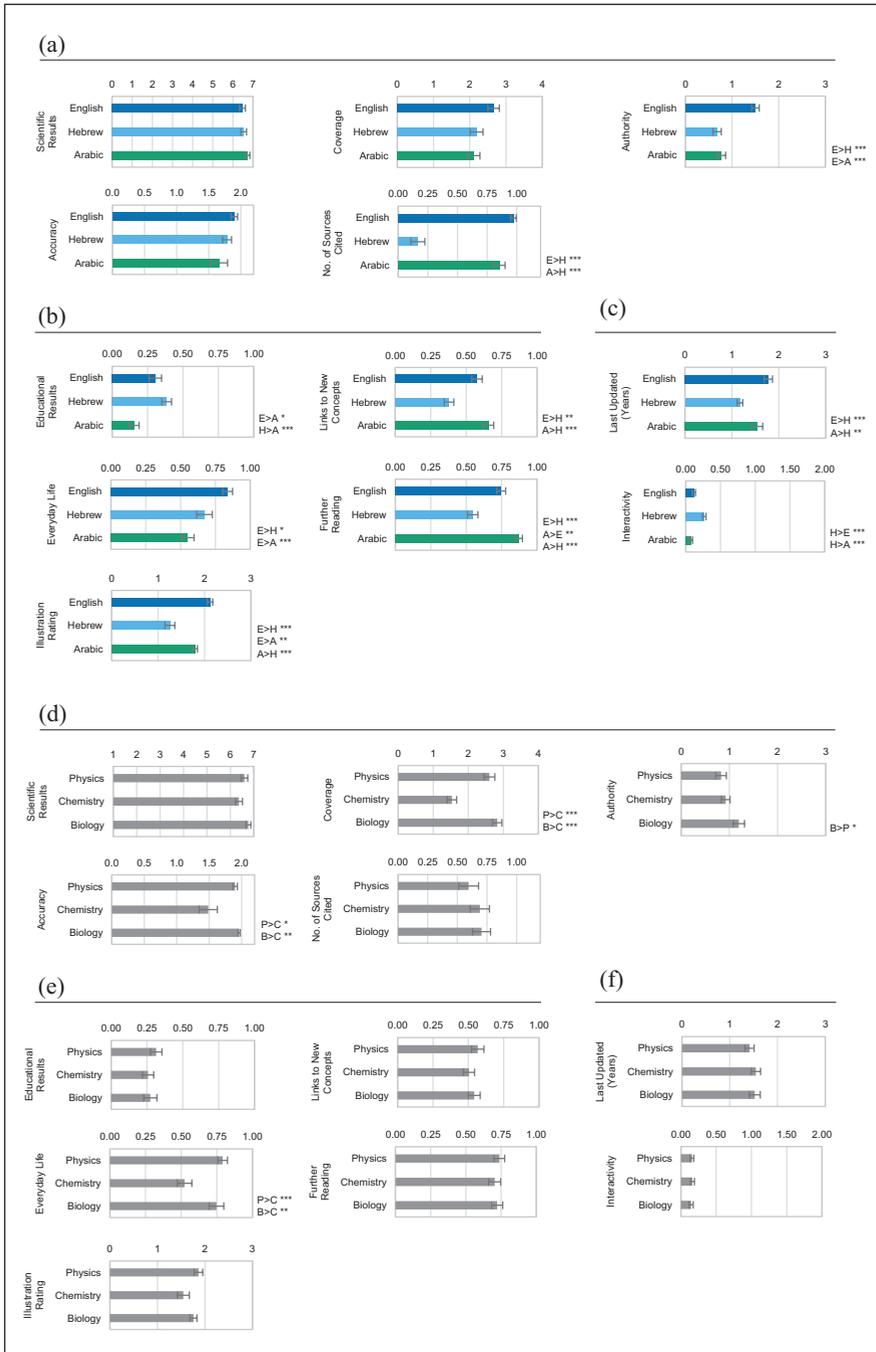
Furthermore, in terms of both scientific and pedagogical quality, the factor with the largest effect size was the language factor, with partial  $\eta^2$  values of .57 and .5, respectively. By comparison, the scientific field factor had smaller, but still significant, effects on these combined measures of information quality, with partial  $\eta^2$  values of .29 and .19 for scientific and pedagogical quality, respectively (Figure 1, Table S1).

*Online-specific quality.* As regards the combined online-specific quality metrics, however, the findings were more complex, since a significant “Language × Field” interaction effect was found ( $p = .016$ ). In other words, language and field have a joint effect on online-specific quality that is significantly different than the sum of the parts (Figure 1, Table S1). Understanding this joint effect entails a close inspection of each of the metrics in this category, which appears in the next section.

#### Individual measures of quality

*Scientific quality by language.* Upon inspection of each individual measure of scientific quality, one-way analysis of variance (ANOVA) tests showed that for most metrics, similar scores were observed across the three languages. The exceptions to this observation were authority ratings and source citations (Figure 2a, Table S2). The average search yielded between six and seven relevant results among the top seven results on average, irrespective of language. In addition, scientific quality, accuracy, and coverage were similar between languages. However, English results had the highest authority ratings ( $p < .001$ ) and Hebrew results cited the fewest sources ( $p < .001$ ).

*Pedagogical quality by language.* By contrast, as regards pedagogical and online-specific quality, the scores differed by language in many ways (Figure 2b, Table S2). English-language results had a consistently high pedagogical quality according to several metrics, and these results scored higher than Hebrew and Arabic results with respect to links to everyday life



**Figure 2.** Information quality by language and by field: (a) scientific quality by language, (b) pedagogical quality by language, (c) variables specific to online content by language, (d) scientific quality by field, (e) pedagogical quality by field, and (f) variables specific to online content by field. E: English; H: Hebrew; A: Arabic; P: Physics; C: Chemistry; B: Biology.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

( $p < .05$ ) and illustration ratings ( $p < .01$ ); however, English results did have some weaknesses compared with Hebrew and Arabic ones. Hebrew results were the most recent ( $p < .01$ ) and interactive ones ( $p < .001$ ), and Arabic results had the most references for further reading. Interestingly, Arabic search results had the fewest educational results ( $p < .05$ ), and Hebrew results had the fewest links to new concepts ( $p < .01$ ).

*Scientific and pedagogical quality by field.* With respect to the differences in quality by scientific field, a much more uniform picture emerged. The average search yielded between six and seven relevant results on average irrespective of field (Figure 2a, Table S2). However, overall, coverage and accuracy were found to be significantly lower for chemistry search results than those of other fields (coverage:  $p < .001$ ; accuracy:  $p < .05$ ). Authority was also found to be higher in biology results than in physics results ( $p < .05$ ). Pedagogical quality was overall similar across fields (Figure 2b) except for references to everyday life, which were less abundant—again—in chemistry results ( $p < .01$ ).

*Online-specific quality by language and field.* The recency of scientific content (Last Updated scores) appeared to vary depending on both language and field. This is evident from a statistically significant interaction effect between language and field for this metric:  $F(4, 81) = 4.49, p = .002$ , partial  $\eta^2 = .18$ . Conversely, in interactivity scores, no such significant interaction was found:  $F(4, 81) = 0.53, p = .72$ , partial  $\eta^2 = .03$ . Differences in interactivity scores were explained by language ( $p < .001$ ) but not by field ( $p = .89$ ; Table S2) and post hoc tests revealed that Hebrew interactivity scores were significantly higher than those of both English and Arabic results ( $p < .001$ ; Figure 2c and f).

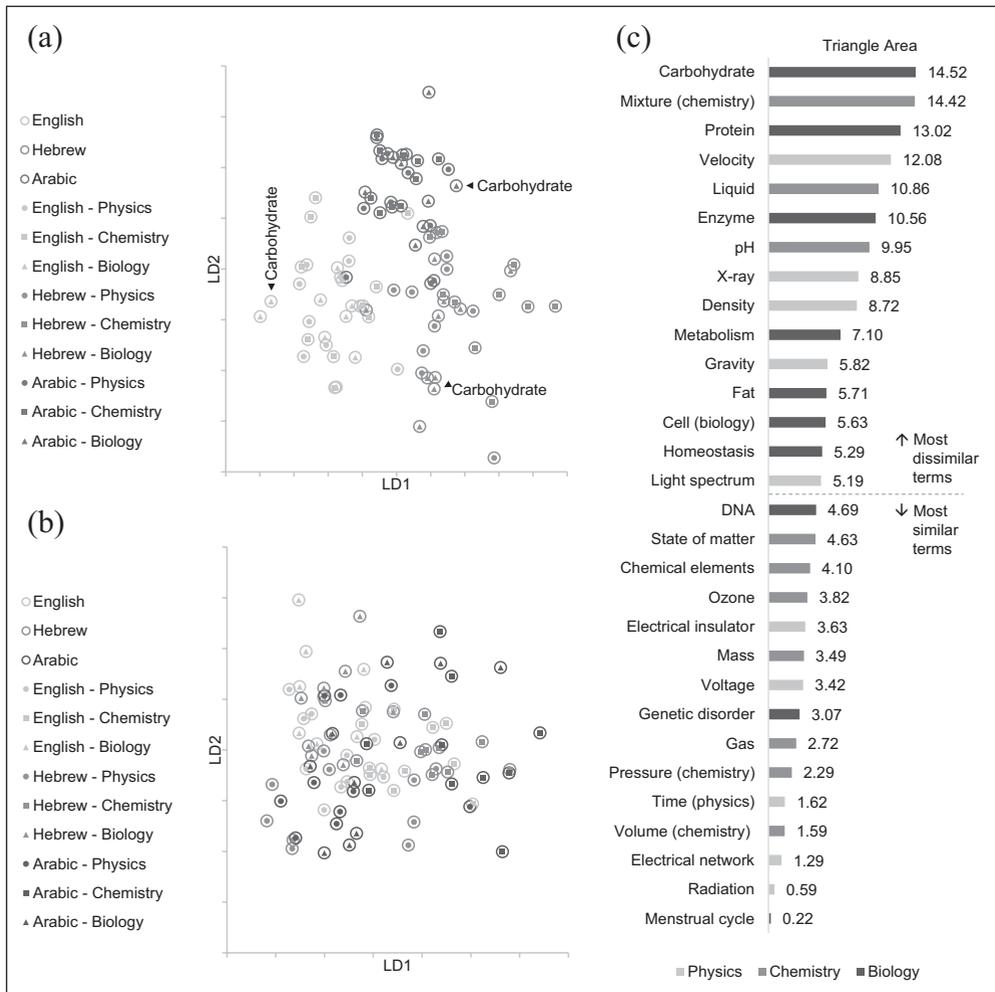
As regards content recency scores, we found significant differences between fields among English and Arabic search results, while Hebrew results were more uniform in this respect. This is evident from the follow-up comparisons we conducted between fields within each language: English:  $F(2, 81) = 3.993, p = .02$ , partial  $\eta^2 = .09$  and Arabic:  $F(2, 81) = 5.56, p = .005$ , partial  $\eta^2 = .12$ . Conversely, for Hebrew results, no significant differences in content recency were found between fields:  $F(2, 81) = 0.12, p = .89$ , partial  $\eta^2 = .003$  (Figure S1). Understanding the effects of language and field on the content recency scores entails a close inspection of the post hoc tests. The results of these are provided in the Supplemental Notes (Figure S1).

### *Were terms clustered together better by language or by field?*

The results of the LDAs show that the chosen search terms were clustered much better by language than by field. When applying the clustering algorithm by language, three almost separate clusters emerged, with only a few points overlapping. This indicates that the search terms in each language share a set of characteristics in common that differs from the other languages (Figure 3a). By contrast, when attempting to cluster the data by field, the clusters overlapped much more. This large overlap indicates that a linear separation of the search terms' characteristics by field is not fully feasible based on the variables measured for this study. Given that the separation by language is successful, we infer that the current selection of characteristics is not sufficient for separation by field (Figure 3b).

### *Which terms have the most similar characteristics across languages and which terms differ the most?*

To measure the distances between equivalent terms in different languages, we calculated the areas of the triangles determined by the three data points representing equivalent terms in Figure 3a (e.g. between the three data points representing “carbohydrate,” when clustering the data by language).



**Figure 3.** LDA of search terms (a) by language (colors) and (b) by field (shapes). (c) Areas of the triangles between equivalent terms in different languages. LD1 and LD2: First and second linear discriminants, respectively. For color version, see Supplemental Notes (Figure S2).

This analysis shows that the list of 15 most similar terms across languages is mostly composed of chemistry and physics terms, with seven and five terms, respectively (Figure 3c). Nevertheless, the “most similar” term across languages overall is “menstrual cycle” (biology).

Conversely, the 15 most dissimilar terms are made up of mostly biological terms, with some of the most dissimilar terms relating to biochemistry and nutrition (“carbohydrate,” “protein,” “enzyme,” and “metabolism”). Other highly dissimilar terms relate to chemistry (“mixture,” “liquid,” and “pH”) and physics (“velocity”).

### Concluding remarks

The main limitation of this study derives from the measurement from a single point of access. This approach may have resulted in location-based results on Google Search and thus biased the data

collection. Future studies could be conducted using pre-programmed searches from multiple servers based in different locations worldwide to control for this effect (as performed by Scherr et al., 2019, for example).

Despite this limitation, this study provides a preliminary characterization of the quality of scientific information available to Internet users in three languages. Our findings raise concern about digital inequalities in opportunities to engage with science between people within the same country along lines of language proficiency, especially with respect to health and nutrition. The findings also suggest that disparities in the quality of online content in the learner's language may contribute to second-level digital divides, especially among young learners. This would add another layer to the "digital inequality stack" of early exposure to computers, on top of known layers, such as socio-economic status and gender (OECD, 2015; Robinson et al., 2020).

These findings are well in line with recent work in informal science education. Studies in science museums have shown how in the United States and the United Kingdom, speakers of minoritized languages are socially excluded when exhibit materials are only available in English (e.g. Dawson, 2014; Yalowitz et al., 2015). The implications can be grave, as:

language has the power not only to prevent people being able to access cultural capital but to contribute to an impression of [informal science education] institutions as spaces of privilege through dominant language use and as a resource for certain groups rather than others (Dawson, 2014: 993).

Based on our findings, it seems likely that many Internet users seeking online scientific content experience similar social exclusion.

Further studies could explore the extent to which disparities in the quality of online content extend to additional languages and to other topics. Contemporary issues, such as COVID-19 and climate change, may be of special interest due to their importance for policymaking and individual decision-making.

Finally, the findings can serve as a call to action for the scientific communities and for institutions that publicly communicate about science and technology. Our findings lend weight to Dawson's (2014) and Márquez and Porras' (2020) calls to make scientific outreach initiatives more inclusive and multilingual. If we wish to reduce inequalities in public engagement with science, we must address their underlying structural causes. One way to do this is to mitigate the language divide in science communication.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Aviv J. Sharon  <https://orcid.org/0000-0003-1326-2168>

Eyal Nitzany  <https://orcid.org/0000-0001-7702-8384>

Ayelet Baram-Tsabari  <https://orcid.org/0000-0002-8123-5519>

## Supplemental material

Supplemental material for this article is available online.

## References

Ahn T, Ryu S and Han I (2007) The impact of Web quality and playfulness on user acceptance of online retailing. *Information & Management* 44(3): 263–275.

- Baram-Tsabari A and Yarden A (2005) Characterizing children's spontaneous interests in science and technology. *International Journal of Science Education* 27(7): 803–826.
- Baram-Tsabari A and Yarden A (2011) Quantifying the gender gap in science interests. *International Journal of Science and Mathematics Education* 9(3): 523–550.
- Brown P (1993) When the public knows better: Popular epidemiology challenges the system. *Environment: Science and Policy for Sustainable Development* 35(8): 16–41.
- Chen W and Wellman B (2004) The global digital divide—Within and between countries. *IT & Society* 1(7): 39–45.
- Cline RJ and Haynes KM (2001) Consumer health information seeking on the Internet: The state of the art. *Health Education Research* 16(6): 671–692.
- Dawson E (2014) “Not designed for us”: How science museums and science centers socially exclude low-income, minority ethnic groups. *Science Education* 98(6): 981–1008.
- De Jesus M and Xiao C (2012) Predicting Internet use as a source of health information: A “language divide” among the Hispanic population in the United States. *Policy and Internet* 4(2): 1–11.
- Dedeke A (2000) A conceptual framework for developing quality measures for information systems. In: *Proceedings of the 5th conference on information quality*, October 20–23, 2000, Cambridge, MA.
- Eberhard DM, Simons GF and Fennig CD (eds) (2020) *Ethnologue: Languages of the World*, 23rd edn. Dallas, TX: SIL International. Available at: <http://www.ethnologue.com/>
- Epstein S (1995) The construction of lay expertise: AIDS activism and the forging of credibility in the reform of clinical trials. *Science, Technology, & Human Values* 20(4): 408–437.
- Ganayem AN (2018) *Hainternet Bahevra Ha'aravit Beyisra'el* [The internet in Arab society in Israel]. Petaḥ Tikva, Israel. Available at: <https://www.isoc.org.il/research/policy-papers/the-internet-in-arab-society-in-israel>
- Guardiola-Wanden-Berghe R, Gil-Pérez JD, Sanz-Valero J and Wanden-Berghe C (2011) Evaluating the quality of websites relating to diet and eating disorders. *Health Information & Libraries Journal* 28(4): 294–301.
- Hargittai E and Hsieh YP (2013) Digital inequality. In: Dutton WH (ed.) *The Oxford Handbook of Internet Studies*. Oxford: Oxford University Press, pp. 129–150.
- Huang S-T, Wang H-R, Yang W-Q, Si Y-C, Wang Y-T, Sun M-L, et al. (2020) Phylogeny of Libellulidae (Odonata: Anisoptera): Comparison of molecular and morphology-based phylogenies based on wing morphology and migration. *PeerJ* 8: e8567.
- International Telecommunication Union (ITU) (2021) Percentage of individuals using the internet. Available at: <https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2021/PercentIndividualsUsingInternet.xlsx>
- Israel Central Bureau of Statistics (2013) *Selected Data from the 2011 Social Survey on Mastery of the Hebrew Language and Usage of Languages*. Jerusalem, Israel. Available at: [https://www.cbs.gov.il/he/mediarelease/DocLib/2013/017/19\\_13\\_017b.pdf](https://www.cbs.gov.il/he/mediarelease/DocLib/2013/017/19_13_017b.pdf)
- Israel Central Bureau of Statistics (2019) *Population of Israel on the Eve of 2020 - 9.1 Million*. Jerusalem, Israel. Available at: [https://www.cbs.gov.il/he/mediarelease/DocLib/2019/413/11\\_19\\_413b.pdf](https://www.cbs.gov.il/he/mediarelease/DocLib/2019/413/11_19_413b.pdf)
- Israel Central Bureau of Statistics and Israel National Authority for Measurement and Evaluation in Education (2016) Preliminary findings from the Israeli survey of adult skills, 2014-2015 (Press Release). Available at: [http://meyda.education.gov.il/files/Rama/PIAAC\\_Takzir.pdf](http://meyda.education.gov.il/files/Rama/PIAAC_Takzir.pdf) (accessed 25 November 2018).
- Israel Ministry of Science Technology and Space (2017) *Tfisot Veemdot Hatsibur Beyisra'el Benos'ey Madda Tekhnologia Vehalal* [Attitudes and Perceptions of the Israeli Public toward Science, Technology and Space]. Jerusalem, Israel.
- Kapoun J (1998) Teaching undergrads WEB evaluation: A guide for library instruction. *ACRL College & Research Libraries News* 59(7): 522–523. Available at: <https://crln.acrl.org/index.php/crlnews/article/view/23707/31079>
- Lissitsa S (2015) Patterns of digital uses among Israeli Arabs—Between citizenship in modern society and traditional cultural roots. *Asian Journal of Communication* 25(5): 447–464.
- Lissitsa S and Chachashvili-Bolotin S (2014) Use of the Internet in capital enhancing ways —Ethnic differences in Israel and the role of language proficiency. *International Journal of Internet Science* 9(1): 9–30.

- Livas C, Delli K and Ren Y (2013) Quality evaluation of the available Internet information regarding pain during orthodontic treatment. *The Angle Orthodontist* 83(3): 500–506.
- Márquez MC and Porras AM (2020) Science Communication in Multiple Languages is Critical to its Effectiveness. *Frontiers in Communication*, Epub ahead of print 22 May. DOI:10.3389/fcomm.2020.00031
- Miniwatts Marketing Group (2020) Top ten internet languages in the world—Internet statistics. Available at: <https://www.internetworldstats.com/stats7.htm> (accessed 9 November 2020).
- National Academies of Sciences Engineering and Medicine (2016) *Science Literacy*. Washington, DC: The National Academies Press.
- National Science Board (2020) Science and technology: Public attitudes, knowledge, and interest. In: *Science and Engineering Indicators 2020*. Alexandria, VA: National Science Foundation. Available at: <https://nces.nsf.gov/pubs/nsb20207/>
- Naumann F and Rolker C (2000) Assessment methods for information quality criteria. In: *Proceedings of 5th international conference on information quality*, October 20–23, 2000, Cambridge, MA.
- OECD (2015) *Students, Computers and Learning: Making the Connection*. Paris, France: OECD Publishing. DOI: 10.1787/9789264239555-en.
- Oliver KM, Wilkinson GL and Bennett LT (1997) Evaluating the quality of internet information sources. In: *ED-MEDIA/ED-TELECOM*, Calgary, AB, Canada, 14–19 June.
- Paschou P, Ziv E, Burchard EG, Choudhry S, Rodriguez-Cintron W, Mahoney MW, et al. (2007) PCA-correlated SNPs for structure identification in worldwide human populations. *PLoS Genetics* 3(9): e160.
- Pearce KE and Rice RE (2014) The language divide—The persistence of English proficiency as a gateway to the internet: The cases of Armenia, Azerbaijan, and Georgia. *International Journal of Communication* 8(1): 2834–2859.
- Pleasant A, Rudd RE, O’Leary C, Paasche-Orlow MK, Allen MP, Alvarado-Little W, et al. (2016) Considerations for a new definition of health literacy. *NAM Perspectives* 6(4). DOI: 10.31478/201604a.
- Q-Success (2020) Usage statistics and market share of content languages for websites, December. Available at: [https://w3techs.com/technologies/overview/content\\_language](https://w3techs.com/technologies/overview/content_language)
- Roberts DA and Bybee RW (2014) Scientific literacy, science literacy, and science education. In: Lederman NG and Abell SK (eds) *Handbook of Research on Science Education*, vol. II. New York, NY: Routledge, pp. 545–558.
- Robinson L, Schulz J, Blank G, Ragnedda M, Ono H, Hogan B, et al. (2020) Digital inequalities 2.0: Legacy inequalities in the information age. *First Monday* 25(7). DOI: 10.5210/fm.v25i7.10842.
- Roth W-M and Lee S (2002) Scientific literacy as collective praxis. *Public Understanding of Science* 11(1): 33–56.
- Rudd RE, McCray AT and Nutbeam D (2012) Health literacy and definition of terms. In: Rowlands G, Begoray D and Gillis D (eds) *Health Literacy in Context: International Perspectives*. New York, NY: Nova Science Publishers, pp. 13–32.
- Savolainen R (2011) Judging the quality and credibility of information in Internet discussion forums. *Journal of the American Society for Information Science and Technology* 62(7): 1243–1256.
- Scherr S, Haim M and Arendt F (2019) Equal access to online information? Google’s suicide-prevention disparities may amplify a global digital divide. *New Media and Society* 21(3): 562–582.
- Segev E and Ahituv N (2010) Popular searches in Google and Yahoo!: A “digital divide” in information uses? *The Information Society* 26(1): 17–37.
- Segev E and Sharon AJ (2017) Temporal patterns of scientific information-seeking on Google and Wikipedia. *Public Understanding of Science* 26: 969–985.
- Shahbazi M, Farajpahlou A, Osareh F and Rahimi A (2019) Development of a scale for data quality assessment in automated library systems. *Library & Information Science Research* 41(1): 78–84.
- Stvilia B, Gasser L, Twidale MB and Smith LC (2007) A framework for information quality assessment. *Journal of the American Society for Information Science and Technology* 58(12): 1720–1733.

- Tate MA (2019) *Web Wisdom: How to Evaluate and Create Information Quality on the Web*. Boca Raton, FL: CRC Press.
- Warf B and Vincent P (2007) Multiple geographies of the Arab Internet. *Area* 39(1): 83–96.
- Warschauer M (2002) Reconceptualizing the digital divide. *First Monday* 7(7). DOI: 10.5210/fm.v7i7.967.
- Yalowitz SS, Garibay C, Renner N and Plaza C (2015) Bilingual Spanish-English intergenerational groups' experiences in bilingual exhibitions. *Museums & Social Issues* 10(1): 35–51.

### Author biographies

**Kawther Zoubi** received her B.Sc.Ed. and M.Sc.Ed. in Science Education from the Technion—Israel Institute of Technology. She has been working as a biology teacher since 2007.

**Aviv J. Sharon** is a research staff member at the Faculty of Education in Technology and Science at the Technion—Israel Institute of Technology. He holds a Ph.D. from the same institution and has professional experience in biological research and in science teaching at the secondary school level.

**Eyal Nitzany** is a computer vision engineer working in the industry. He holds a B.A. in mathematics and computer science, an M.Sc. in biostatistics and a Ph.D. in computational neuroscience from Cornell University. He worked as a research fellow at the University of Chicago and Northwestern University. His research focuses on statistical modeling and interpretations of data from various fields such as neuroscience, cellular biology, machine learning and psychological studies.

**Ayelet Baram-Tsabari** is an Associate Professor at the Faculty of Education in Science and Technology at the Technion—Israel Institute of Technology and heads the “Applied Science Communication” research group. Former elected member of the scientific committee of Public Communication of Science and Technology Network (PCST) and of the Israel Young Academy. Her research focuses on the interactions between science education, science literacy and public engagement with science in real-life scenarios.