Scientific Curiosity and its Role in Forming Beliefs

Thomas Ash, Sarah Cooney, Nikolaos Flemotomos, Amabel Jeon
University of Southern California

Abstract

Despite the wide scientific agreement that climate change is a real and observable scientific phenomenon, the U.S. population shows high rates of polarization on whether or not this is true. Conceptualizing this polarization is crucial in understanding how to make scientific evidence a primary source of policy making. We address this question by looking at two explanations provided by the existing literature using text based methods on news information at the whole country level. We find that scientific literacy does not decrease rates of climate change denial or increase vaccination rates (our measures of belief in scientific phenomenon) consistent with existing research. We also find limited evidence that curiosity provides a population level reason for polarization, in contrast to individual level survey evidence in the literature. Finally, we explore the meaning of curiosity as a psychological concept.

Scientific Curiosity and its Role in Forming Beliefs

Introduction

Over the last few decades, there has been an overwhelming scientific agreement on climate change. Despite the accumulation of scientific evidence that strongly documents human-induced global warming, U.S. public opinion on this issue continues to be polarized. A growing number of people have recognized the severity and the urgency of the climate crisis in recent years, with John Kerry, the former senator and secretary of state, declaring 'World War Zero' on climate change (Friedman, 2019). However, there still exists a spectrum of those who are skeptical or even deny the reality and seriousness of anthropogenic global warming (Leiserowitz, Maibach, Roser-Renouf, & Hmielowski, 2012). The fossil fuel industry has contributed to this polarization through misinformation used to shape public opinion. For instance, 80% of ExxonMobil's public statements from 1989 to 2004 expressed doubt on climate change while 80% of its internal documents acknowledged that climate change was real and human-caused (Supran & Oreskes, 2017).

Misinformation on climate change, such as misleading statistics, disproportionately influences political conservatives while having little to no effect on political liberals (Cook, Lewandowsky, & Ecker, 2017; van der Linden, Leiserowitz, Feinberg, & Maibach, 2015). Therefore, such misinformation plays a role in exacerbating public polarization over climate change (Cook, 2019). According to the most recent CBS News poll, there continues to be a partisan split on views on climate change and its cause. Republicans are more skeptical than liberals about the urgency of the issue as well as the degrees to which climate change is contributed by human activity (De Pinto, Backus, & Salvanto, 2019).

Such polarisation is harmful since it delays action on widely corroborated and accepted scientific evidence. Another heavily polarized topic is vaccination. The rate of anti-vaccine conspiracy theories is on the rise and the consequence of this spread of misinformation is exemplified by the 1,249 reported cases of measles in the U.S., the highest annual number since 1992 (Patel et al., 2019). The question then is what causes this harmful disbelief and could an understanding of its causes lead to potential solutions? One immediate hypothesis that we test in this paper is do differing rates of scientific literacy in the U.S. population explain the acceptance of climate change and vaccination rates by certain individuals?

Contrary to intuition, Kahan et al. (2012) found that there is no positive correlation between scientific literacy/numeracy and concerns about climate change using survey evidence. In fact, they found that polarization in climate change beliefs is greater for those with more general education, more science education, and higher scientific literacy scores (Drummond & Fischhoff, 2017; Kahan et al., 2012). In this same research, Kahan et al. (2012) demonstrated that rather than scientific literacy, scientific curiosity, as an important individual difference in cognitive style, led respondents to uniformly adjustment their assessments of climate change, rather than creating polarization. In this work, we assess these two findings.

Our assessment in this paper also required an examination of the concept of scientific curiosity, and thus we have also made contributions to measuring the psychological concept of curiosity. Since curiosity is of itself imprecisely defined, our work first required us to determine what curiosity is. Our review of the psychological literature found that curiosity is an intrinsic drive that equates with unpleasant feelings of "deprivation" that result from lacking desired knowledge (J. Litman, 2005). Notably, curiosity is not driven by an objective lack of pre-existing knowledge but rather by the perceived knowledge gap based on metacognitive estimates of one's available knowledge (J. A. Litman, 2009). Therefore, scientific curiosity reflects the motivation and information-seeking behaviors for content-specific knowledge about the natural phenomena for personal pleasure (Jirout & Klahr, 2012; Kahan et al., 2012; Spektor-Levy, Baruch, & Mevarech, 2013). A perceived knowledge gap could work as a driving force to seek out new information to challenge one's preconception, decreasing the likelihood of credulously accepting misinformation.

Past research on curiosity and scientific curiosity mainly relied on inventories and scales for measurement (see the supplemental material for a list of examples). We aim to further our understanding of the relationship between scientific curiosity and belief in objective scientific phenomena such as climate change and vaccinations. To do so, this study uses local news media information to measure scientific literacy and curiosity at the population level. We use articles from the hyper-local, online news source Patch (`www.patch.com`), which contains news articles based in various U.S. cities and neighborhoods. As a result, it is necessary to make assumptions about how a population is represented by its media. In general, media outlets follow public opinion, telling viewers and readers what they already believe, to attract and maintain an audience as readers have an economically significant preference for like-minded news (Gentzkow & Shapiro, 2010; George & Waldfogel, 2006). Now, within the contemporary news environment, the emphasis on the engagement with the members of the communities to serve and understand the needs and interests of the audiences has is higher than ever (De Aguiar & Stearns, 2015). Although local newspapers have faced economic hardships in recent years, they have experienced notable resilience and remain the most significant providers of journalism in their communities (Napoli & Mahone, 2019). Local newspapers have embraced the challenge of people turning to social media for new information by working to quickly make sense of a huge range of differing views about an issue and knowing the right questions to ask for their local audience (Greenslade, 2012). Additionally, local people continue to trust their local newspaper more than they trust the national media (Greenslade, 2012).

In order to estimate the presence of the concepts of scientific literacy and curiosity within the news articles, text analysis methods were applied. In particular, we used Distributed Dictionary Representations (DDR) (Garten et al., 2018), as well as Bidirectional Encoder Representation from Transformers (BERT) (Devlin, Chang, Lee, & Toutanova, 2019) as methods for binary text classification. On the one hand, DDR is an efficient and highly-interpretable way to use existing domain knowledge for large text corpora analysis. On the other hand, BERT is a state-of-the-art language modeling technique that we assumed would help us obtain highly reliable estimates of scientific literacy and curiosity. To evaluate our methods, a subset of the data was manually annotated.

Statistical methods were used to compare these representations with data on climate change belief and vaccination rates. We test the same hypotheses as Kahan et al. (2012), using a much more general dataset (covering all U.S. states at an aggregated level), and are able to corroborate the finding that literacy does not increase rates of belief. The implication of this work is that simple, knowledge-based scientific education will not improve belief in objective scientific phenomena. We also test their hypothesis on scientific curiosity finding some, but overall limited and inconsistent, evidence that curiosity can predict belief in scientific phenomena at the whole population level. Therefore our work appears to mute the finding that curiosity can play a role in solving the polarization in belief on objective scientific phenomena.

Local News Data

We collected all available articles from Patch in the Community (around-town) and Health (lifestyle) categories. The Community category was chosen as a general representation of local sentiment. The Health category was chosen for its proximity to science and the likelihood of its containing scientific content. A subset of the data was annotated for scientific curiosity and scientific literacy. Post-processing was performed to attach county designations to patches and to exclude very short articles. The details of each data phase, collection, annotation, and post-processing, are described below.

Collection. Data was scraped from patch.com using a python script, which was a modification of code originally authored by Joe Hoover. For each article, the state, patch (city or neighborhood), category, date, author, title, and text were collected into a mySQL database. Data was collected from 1170 patches from across all 50 states and Washington, D.C. Over 550,000 articles were collected from the Community category and more than 350,000 from the Health category.

Annotation. Articles were evaluated for two distinct concepts: Scientific Literacy and Scientific Curiosity. For both concepts each article was rated on both a five point (1-5) scale as well as given a binary (0-1) classification, with 1 indicating the presence of the concept and 0 indicating its absence. The annotation guide can be found in the supplemental material. The manual annotations were done by the four memebers of the group. The raw datasets corresponding to each of the two article categories were first filtered to exclude the state ‘us’ corresponding to the entire country, and the patches ‘across-xxx’, corresponding to entire states. No other filtering was performed. Each category was split into 4 lists of articles: each list contains articles from all the states, but from different patches (there are overlaps, especially for states with few patches). To create the lists, the initial dataset was sorted with regard to 1) patch, 2) author, 3) year, independently for each state. Based on that, 4 chunks were created for each state, which were later combined to form the final lists that include all the states. Each list was randomly shuffled and then re-sorted with regard to state. Each member of the group was responsible for annotating one article from each category for each state, for a total of 423 annotated articles. However, since some states had very few articles, not every chunk had an article from every state. To make up for this, annotators evaluated extra articles from states with a large number of articles. Table 1 shows examples of text from annotated articles rated as containing scientific literacy, scientific curiosity, both, and neither.

Post-Processing. Documents from the ’across-’ patches, or from the ’US’ patch and those with less than 50 words were discarded, leaving a total of 810,747 documents. For each of the remaining patches, the county designation was attached with an R script that uses the acs (American Community Survey) package. (The script was a modification of code originally written by Joe Hoover). After running the script, the counties for 152 patches were unable to be identified. These mostly consisted of neighborhoods within a larger city, and were hand annotated using Wikipedia seraches.

Patches crossing multiple counties or with no counties—some cities in Virginia have no official county designation—were discarded. Washington DC was given its own county designation for analysis purposes. This left 758,883 documents from 285 counties. (There are a total of 3,242 counties in the US.) Analysis was run on this dataset as well as a subset of the counties having more than 100 articles and more than

Table 1

Examples of annotated articles containing Scientific Curiosity, Scientific Literacy, Both, and Neither. Bold indicates phrases representing scientinfic curiosity. Underliing represents phrases indicating scientific literacy.

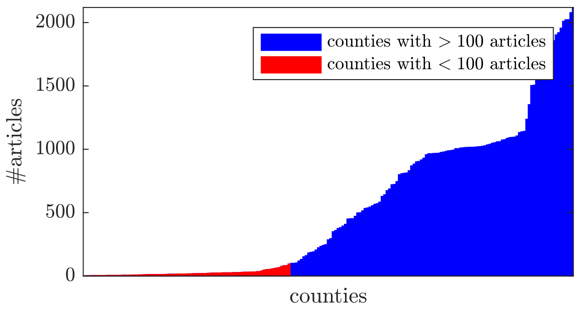
Scientific Curiousity	The Homewood **Science Center is hosting** a PopUp SCIENCE Bubble Bash **where the community can explore** surface tension, chemistry, and geometry from 10 a.m. to 1 p.m.
Scientific Literacy	Diagnosed with HIV in 2009, he suffers from neuropathy, or numbness in his peripheral nerves, and has a hernia in his lower abdomen.
Both	**For reasons no one can quite explain**, eels are ending up in the seals snouts, hanging out like something that should have been mopped away with a sturdy hanky.
Neither	More than 400 signatures were compiled out of concern that the towns proposed changes to the Peddling and Zoning Ordinance would hurt Taqueria Cinco, a food truck in town, the Guilford Courier reports.

50 distinct authors. This set contained 746,371 documents from 172 counties. Figure 1 shows the number of articles and authors per county for those sets.

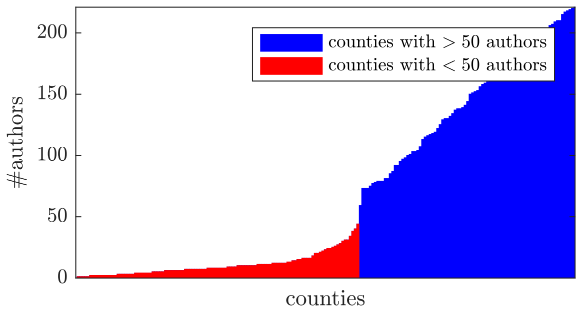
Methods

Distributed Dictionary Representations

Background. Domain dictionaries, coupled with word-counting methods, have been extensively used in several psychological applications (Pennebaker, 2011). This framework, where the Linguistic Inquiry and Word Count (LIWC) is maybe the most notable example (Pennebaker, Francis, & Booth, 2001), involves estimating the occurrences or frequencies of words along particular dimensions, such as “anger”, “health”, “work”, etc, and using this information to analyze large text corpora. Even though such methods have been successfully applied to several studies and seem ideal when domain knowledge from experts is required, they suffer from a number of



(a) Number of articles per county.



(b) Number of authors per county.

Figure 1. Number of articles and authors per county, highlighting the tails of the distributions with counties featuring < 100 articles or < 50 authors. Only the 200 counties with the fewest articles/authors are shown.

limitations, including the biases of the researcher developing the dictionary and the non-static nature of language, which can potentially affect the effectiveness of a dictionary after some period of time (Garten et al., 2018).

More recently, the research interest has been focused on capturing the semantic similarity of linguistic units by constructing neural word (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013; Pennington, Socher, & Manning, 2014) or sentence (Pagliardini, Gupta, & Jaggi, 2018) embeddings in some high-dimensional vector space. Under this viewpoint of semantic vector space representation, words (or sentences) that appear in similar contexts are located closer to each other. Such embeddings can be efficiently trained in an unsupervised way, thus avoiding the burden of creating large dictionaries, and can encode both attributional and relational similarity between linguistic units.

An elegant way of combining the computational and representational advantages of word embeddings with the expert's knowledge and intuition to capture specific high-level constructs through language has been proposed in (Garten et al., 2018) and is known as the Distributed Dictionary Representation (DDR) method. The idea is to represent an entire dictionary category as a point in the semantic vector space and then estimate the similarities of words or documents to that category employing some distance metric instead of word-counting techniques.

In more detail, let's assume we have a dictionary D containing a specific number of words $w_1, w_2, \cdots, w_N$ that are expected to encapsulate some high-level concept. In our case, the two concepts of interest are scientific curiosity and scientific literacy. Let's also assume we have a document T for which we want to extract a metric representing whether that concept is existent or the degree to which it exists. The document is also viewed as a collection of words $\tilde{w}_1, \tilde{w}_2, \cdots, \tilde{w}_{N'}$. After calculating the normalized embeddings $v(w_i)$ and $v(\tilde{w}_j)$of all the words in both D and T, we can estimate an aggregate representation of those:

$$v(D) \triangleq \frac{\sum_{i=1}^{N} v(w_i)}{\left|\left|\sum_{i=1}^{N} v(w_i)\right|\right|}, \quad v(T) \triangleq \frac{\sum_{j=1}^{N'} v(\tilde{w}_j)}{\left|\left|\sum_{j=1}^{N'} v(\tilde{w}_j)\right|\right|}$$

Then, the distance between the document T and the concept in the dictionary D can be simply estimated using the cosine metric:

$$d(T, D) = \cos\angle(T, D) = v(T) \cdot v(D)$$

Since this results in a continuous value, in case we are interested only in the existence of the concept within the document (we can think of it as a binary classification task), we can use a thresholded version of it:

$$d_\theta(T, D) = \mathbb{I}\{d(T, D) > \theta\}$$

where $\mathbb{I}\{\cdot\}$ is the indicator function.

Implementation and Evaluation. The first step before being able to apply DDR is of course to design the dictionaries to be used. Regarding the concept of scientific curiosity, we studied existing literature and agreed on four 4-word dictionaries representing different levels of abstraction (see also the Supplemental Material). As far as the concept of scientific literacy is concerned, we made the assumption that this is reflected by the usage of scientific terms in language. To that end, we used an extended list of terms found online[1] which includes 211 words (only keeping the single-word terms). Additionally, we experimented with a more generic list of 68 words which are "*part of the academic language of science, yet are not specific to science*", "*necessary if*

[1] `https://foxhugh.com/word-lists/list-of-scientific-terms/`

one is to read and understand science literature"[2]. A summary of the dictionaries used is given in Table 2.

Table 2

Dictionaries used with DDR to represent the concepts of scientific literacy and scientific curiosity.

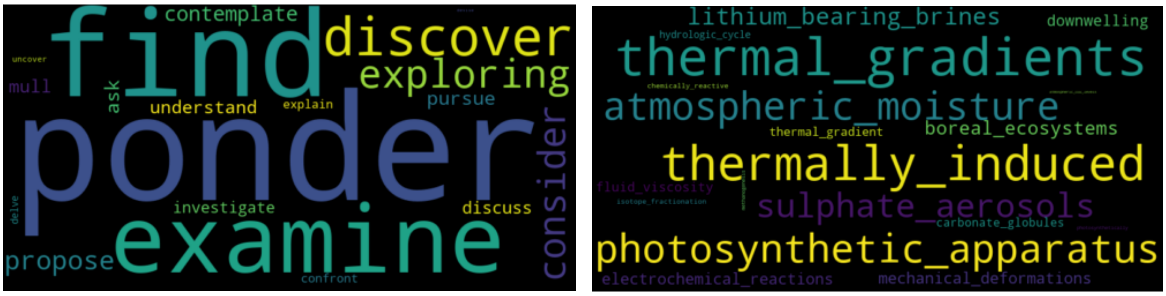
concept	**dict. name**	**#words**	**words (or examples)**
sc. literacy	$D^{lit}_{specific}$	211	circuit, mitosis, spectroscope, taiga
	$D^{lit}_{generic}$	68	assessment, deduction, potential, strategies
sc. curiosity	D^{cur}_{1}	4	museum, outdoor, nature, scientist
	D^{cur}_{2}	4	fact, unpredictable, evidence, experiment
	D^{cur}_{3}	4	explore, cause, visit, question
	D^{cur}_{4}	4	solve, seek, explore, wonder

To get an understanding of how DDR works, Figure 2 shows the nearest neighbors (in terms of cosine similarity) to each of the two concepts of interest using one of the proposed dictionaries. The vector representations used here and in the rest of the work are the word2vec embeddings, pre-trained on $\sim$ 1B words of the Google News dataset[3].

In order to decide which dictionaries to use for the entire dataset, we evaluated all the aforementioned dictionaries on the available annotated articles. To do so, we used the ground truth binary annotations and we also obtained a binary annotation for each article, based on the thresholded distance metric described previously. Since this metric depends on a hyper-parameter θ, the evaluation was based on a 5-fold cross-validation scheme where at every iteration the optimal θ was chosen as the one maximizing the classification accuracy for the corresponding fold. Because of the big gap between the number of words in the dictionaries examined for the scientific literacy and scientific curiosity, for the latter we additionally experimented with the merged dictionary

[2] `https://www.csun.edu/science/ref/language/awl-science.html`

[3] `https://code.google.com/archive/p/word2vec/`



(a) scientific curiosity (D_4^{cur}) (b) scientific literacy ($D_{specific}^{lit}$)

Figure 2. Word clouds with the 20 nearest neighbors of the aggregate representations of the concepts "scientific curiosity" and "scientific literacy" according to some of the dictionaries examined. The font size is proportional to the normalized cosine similarity.

$D_{merged}^{cur} = \bigcup_{i=1}^{4} D_i^{cur}$ (containing 16 words), as well as with seeded dictionaries of variable length. In particular, if we want to create a seeded dictionary of N words, we get the aggregate representation of D_{merged}^{cur}, we find the N nearest neighbors, and those are the N terms of the new dictionary. The results, in terms of cross-validated accuracy, are given in Figure 3 and in the first rows of Tables 3 and 4. The baseline results are the ones derived from a majority-class classifier.

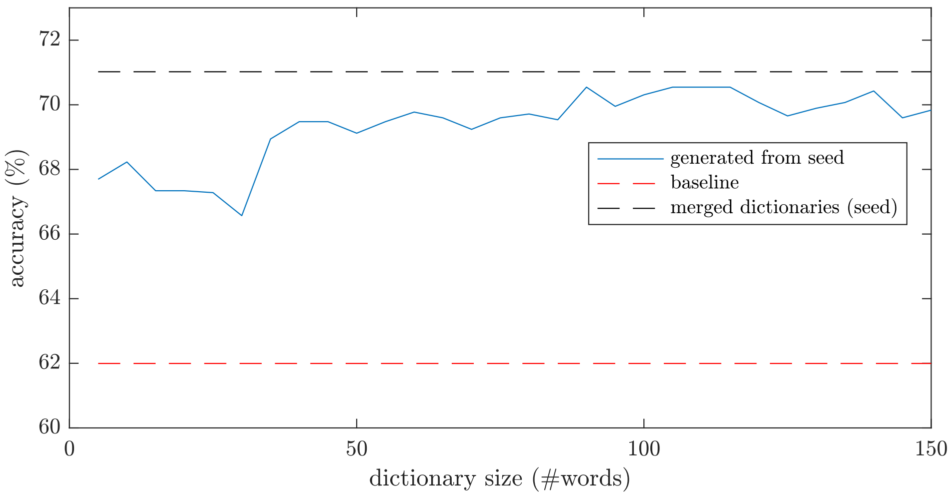


Figure 3. Cross-validated classification accuracy applying thresholded DDR for the concept of "scientific curiosity" when using seeded dictionaries of variable size, where D_{merged}^{cur} is always used as the seed.

Based on those results, we chose to use D_{merged}^{cur} for scientific curiosity and $D_{specific}^{lit}$

for scientific literacy for the rest of the analysis. In order to find the optimal threshold to use, we ran once again the method and evaluated it on the entire annotated set (without cross-validation) again selecting the threshold that maximizes the classification accuracy. The results (for all the dictionaries) are given in the second rows of Tables 3 and 4.

Table 3

Classification accuracy applying thresholded DDR for the concept of "scientific curiosity" using different dictionaries. First row shows cross-validated results, second row shows the overall results based on the optimal threshold after a line search.

dictionary	D_1^{cur}	D_2^{cur}	D_3^{cur}	D_4^{cur}	D_{merged}^{cur}	baseline
5-fold	64.73%	68.68%	65.62%	66.03%	**71.02%**	62.00%
all	64.61%	68.65%	65.32%	66.03%	**71.02%**	

Table 4

Classification accuracy applying thresholded DDR for the concept of "scientific literacy" using different dictionaries. First row shows cross-validated results, second row shows the overall results based on the optimal threshold after a line search.

dictionary	$D_{generic}^{lit}$	$D_{specific}^{lit}$	baseline
5-fold	68.88%	**80.52%**	66.03%
all	68.88%	**80.29%**	

Once we decided on the dictionaries and the corresponding thresholds to use, we were able to run DDR on the entire dataset of available articles. It is worth noting that before running the DDR method, each article was normalized to remove punctuation, non-printable characters, and stopwords. After this procedure, articles with less than 20 words were removed from the analysis (usually empty documents or only hyperlinks). That way, each normalized article received a binary label (0 or 1) representing whether or not it reflects scientific curiosity and a binary label (0 or 1) representing whether or not it reflects scientific literacy. By averaging the results over all the articles at the county level, we could extract a continuous value in $[0, 1]$ reflecting how "scientifically curious" or "scientifically literate" this county is. The results are visually depicted in Figure 4.

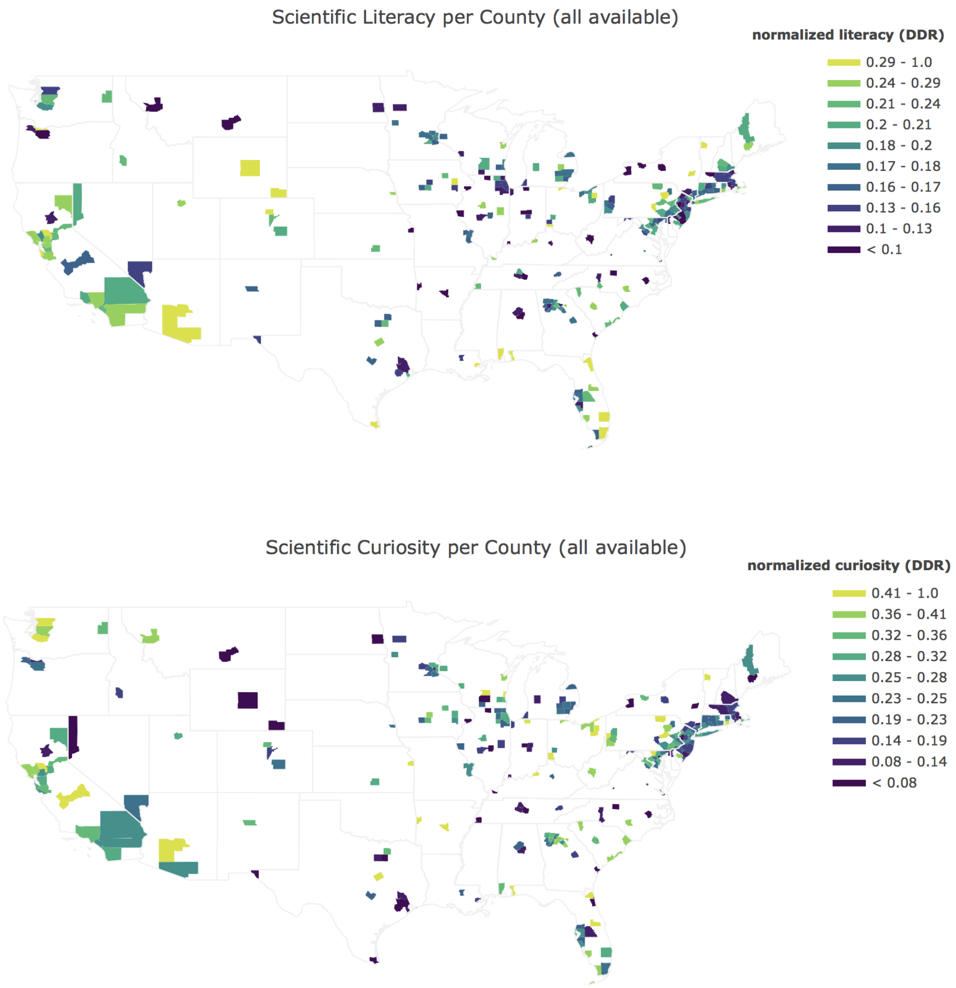


Figure 4. Scientific literacy and scientific curiosity per county after the DDR analysis for all the available 285 counties. The results were further normalized to cover the entire $[0, 1]$ range and clustered into 10 bins for visualization purposes.

BERT

Background. BERT, which stands for Bidirectional Encoder Representations from Transformers, is a language representation model recently introduced (Devlin et al., 2019), which has already achieved state-of-the-art performance in a wide variety of natural language processing tasks, including text classification, question-answering, sentence completion, and natural language inference.

As opposed to the semantic vector representations which serve as a way to extract linguistic features to be used in a subsequent module (e.g. a distance-based classifier in the case of DDR), BERT has been proposed to be used as a pre-trained end-to-end model the parameters of which can be fine-tuned for the desired task. That means that the architecture details remain the same for all the different tasks and the researcher is not concerned with task-specific methodologies and parameters. During the pre-training phase, BERT is trained in an unsupervised way over a variety of tasks. During fine-tuning, on the other hand, the model is initialized with the pre-trained parameters and is further updated using task-specific labeled data.

BERT is built on top of the Transformer architecture (Vaswani et al., 2017). Transformers were initially introduced as sequence-to-sequence models which extend the widely used encoder-decoder structure. The idea is that given an input sequence, the encoder will transduce it into a latent representation which will be further processed by the decoder to produce an output sequence. A Transformer uses 6 identical encoding layers as the "encoder" part and 6 identical decoding layers as the "decoder" part. For BERT, since the goal is to generate a language representation model, only the encoder mechanism is necessary. Each encoding layer consists of a multi-head self-attention mechanism which helps the network attend to specific neighboring words of the input word being encoded and a feed-forward 2-layer subnetwork.

Viewing the entire Transformer encoder as a black box, BERT uses stacked such blocks in a bidirectional manner. This deep bidirectionality is a big advantage compared to previously proposed architectures which use either left-to-right or right-to-left language models or a concatenation of independently trained unidirectional models.

As already mentioned, BERT has been proposed to be used for a variety of tasks involving minimal changes in the architecture. This includes single-document classification tasks, as well as sequence transduction tasks which require pairs of documents. Since in this work we are interested in document classification, we will be focusing on that realm. Given a sequence of words, this is first transformed into a sequence of tokens by a pre-trained tokenizer. The tokens `[CLS]` and `[SEP]` are added by default at the beginning and at the end, respectively. The output of the BERT model is a sequence of hidden vectors, each one corresponding to an input token. In order to do the final classification, we added a pooling layer to get a representation of the entire input sequence which was fed to a dense layer before a final inference layer giving the probability that our input is a member of the concept class under examination (e.g. scientific curiosity or scientific literacy). The architecture is illustrated in Figure 5.

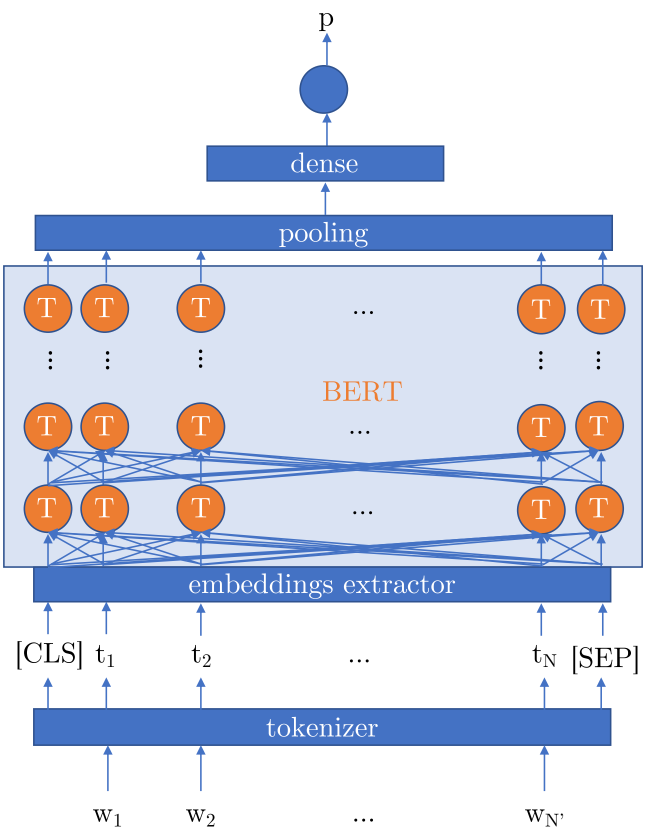


Figure 5. BERT-based architecture for binary text classification. The input sequence of words $\{w_i\}_{i=1}^{N'}$ is parsed into a sequence of tokens which is then passed to BERT. The BERT output is pooled to get a sequence representation and we finally get an output probability p after a dense layer. Each module T here corresponds to a Transformer encoder block.

Implementation and Evaluation. For our experiments we used the cased $\mathrm{BERT}_{\mathrm{BASE}}$ model[4] which uses 12 stacked Transformer layers, 12 heads for self-attention, and the output hidden vector is 768-dimensional. On top of that we added a 256-dimensional dense layer as shown in Figure 5. For pooling, the most common practice is to just use the first hidden vector (the one corresponding to the `[CLS]` token) as the input sequence representation, as recommended in the original paper (Devlin et al., 2019). We experimented both with that option, as well as with averaging over all the output vectors. The latter option yielded slightly better results, so this is what we used for our analysis. The output results were binarized based on a $p = 0.5$ probability threshold.

The original paper suggests fine-tuning all the network parameters. However, we opted for tuning only the top 3 Transformer layers of BERT (and of course the dense layer), mainly due to computational limitations. This resulted in a network with $\sim$ 110M parameters, out of which $\sim$ 22M are trainable. Assuming that we can trust the extreme decisions made by DDR, we fine-tuned the model using the top and bottom 10% of the articles in our dataset, after sorting them with respect to the DDR scores (before binarization). As in the DDR case, each article was normalized to remove punctuation, non-printable characters, and stopwords, and we also removed articles with less than 20 words after normalization. Additionally, duplicate samples (i.e. reposted articles) were removed during fine-tuning to avoid over-fitting. This led to $\sim$ 71K articles used for this procedure.

Fine-tuning was done using the SGD optimizer with a learning rate equal to 0.01 and mini-batch size equal to 32. 20% of the articles used for fine-tuning were held out as a validation set. The results after 1 training epoch are given in Table 5. It should be highlighted that those results do not reflect how good the BERT model is with respect to some ground truth but only how similar the BERT output is to the DDR output. As observed, the BERT-based model yields almost identical results with the DDR-based approach for the extreme articles (the most (dis)similar to our concepts according to

[4] `https://tfhub.dev/google/bert_cased_L-12_H-768_A-12/1`

DDR). However, this is not the case for the rest of them ("test" set), especially as far as the scientific curiosity is concerned.

One practical consideration with our model is that BERT can only handle sequences of a maximal length of 512 tokens. Although we could have used a hierarchical model that can consume the entire article, or some BERT variant specifically proposed for document classification, e.g. (Adhikari, Ram, Tang, & Lin, 2019), we chose just to crop the available articles and only use the first 510 tokens (and also include the `[CLS]` and `[SEP]` tokens). This decision was informed by the observation that usually the initial part of an article contained enough information for a human annotator to decide whether the concept of interest exists in the article or not.

Table 5

Classification accuracy comparing the BERT binary outputs with the DDR binary outputs for scientific literacy and scientific curiosity. 10% *of the articles with the highest and* 10% *of the articles with the lowest DDR score with respect to each concept were used for fine-tuning. Out of those,* 20% *was used as the "validation" set and* 80% *as the "training" set. All the rest are regarded as the "test" set.*

	training	validation	test
sc. literacy	99.59%	99.95%	65.87%
sc. curiosity	99.55%	99.88%	47.30%

Similarly to the DDR method, our BERT-based approach assigned a binary label (0 or 1) to each article representing whether or not it reflects scientific curiosity and a binary label (0 or 1) representing whether or not it reflects scientific literacy. Those results were subsequently averaged at the county level, with the final output illustrated in Figure 6.

Statistical analysis

We use a simple statistical cross-section approach using Ordinary Least Squares (OLS) regression. We present a range of different statistical models with different

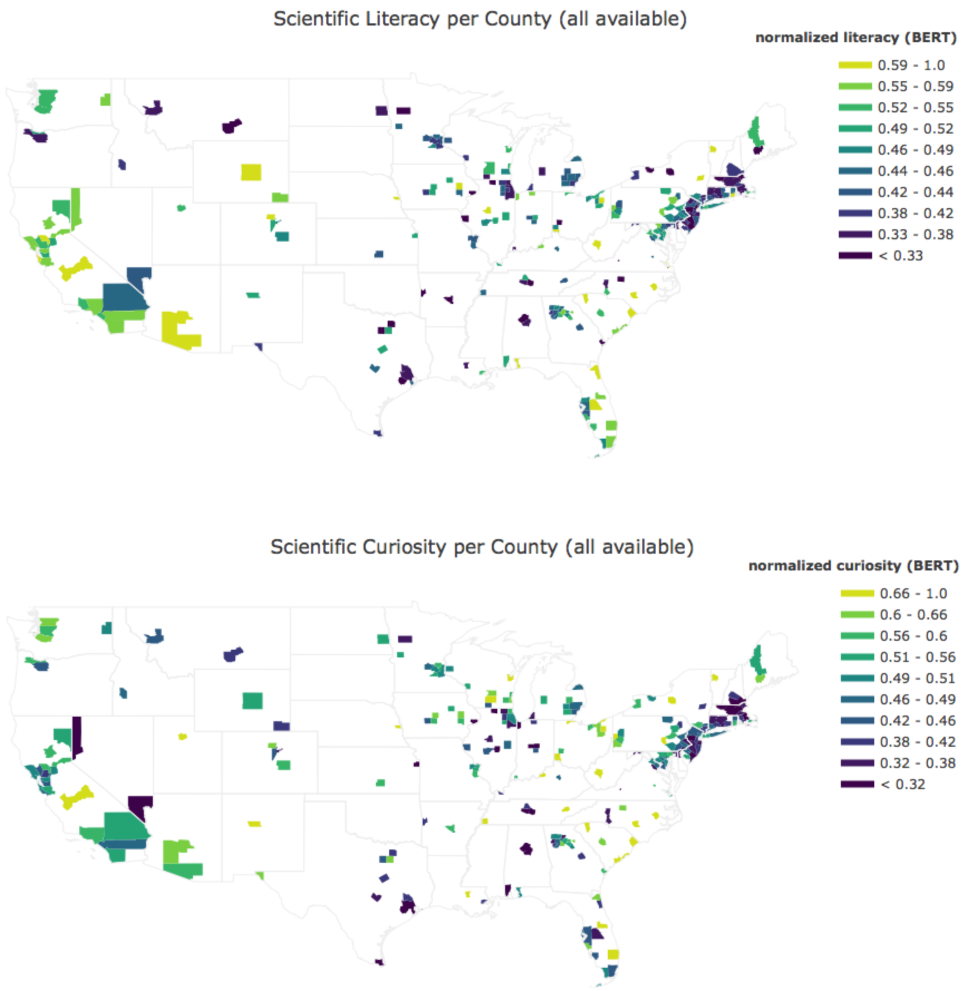


Figure 6. Scientific literacy and scientific curiosity per county after the BERT analysis for all the available 285 counties. The results were further normalized to cover the entire $[0, 1]$ range and clustered into 10 bins for visualization purposes.

dependent and independent variables to take a broad assessment of the available evidence on our research question and ensure that our results are not dependent on model specification.

The statistical model we use is as shown in the equation below:

$$\text{Beliefs}_i = \alpha + \beta_1 \cdot \text{ScientificLiteracy}_i + \beta_2 \cdot \text{ScientificCuriosity}_i + \beta_3 \cdot X_i + \epsilon_i$$

In the above equation, i specifies the geographical unit in question (either State or County). X_i specifies a matrix of regional explanatory variables including average education levels, income levels, among other things (see Figure 8 for a complete list).

Beliefs$_i$ are from four sources: the Yale Climate Survey in both 2016 and 2018 at the County level; vaccination rates at the state level; and Google Trends measure of searches for "Climate Change Real" in 2018 across states. Finally, scientific literacy and curiosity are measured from either DDR or BERT, as described in the previous sections.

It is important to examine correlations among dependent and independent variables before using OLS regressions. Doing so avoids problems with measurement error and multi-collinearity. Below we present such correlations. We find that our measures of belief in objective scientific phenomena are quite uncorrelated – suggesting some potential measurement issues with the concept we are attempting to measure. This concern is present in all the existing research on this subject to date e.g. (Kahan et al., 2012). A scatter examination of state climate change beliefs against vaccination rates suggests there are few states with high climate change belief and low vaccination rates, however there are states with high vaccination rates and low rates of climate change belief. This suggests vaccination rates may be a noisier (and therefore worse) measure of scientific beliefs, since climate change belief moves correctly with low vaccination rates.

	Yale Survey *CC happening - 2018*	Yale Survey *CC happening - 2016*	Vaccinations *Rates (%)*	Google trends *CC real*
Yale Survey *CC happening - 2018*	1			
Yale Survey *CC happening - 2016*	0.98	1		
Vaccinations *Rates (%)*	0.1	0.1	1	
Google trends *CC real*	0.21	0.16	-0.23	1

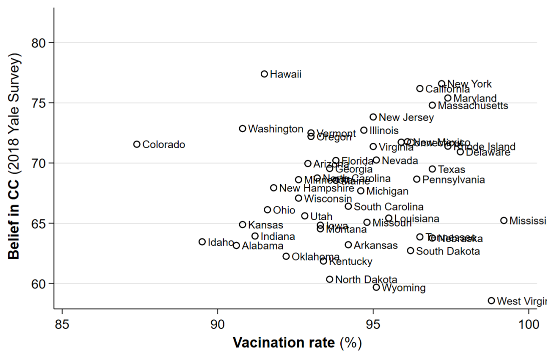


Figure 7. Dependent variables do not consistently measure belief in Scientific phenomena.

Finally, examining correlations among explanatory variables shows that most are uncorrelated. Of particular importance literacy and curiosity scores have reasonable separate variation with a correlation coefficient of 0.47. Multi-collinearity concerns therefore appear minimal.

	ScLiteracy	Curiosity	Unempl	Med_HH_Inc	Poverty_%	BA_Degree+	Population	Area
ScLiteracy	1							
Curiosity	0.47	1						
Unempl	0.19	0.12	1					
Med_HH_Inc	-0.08	-0.25	-0.33	1				
Poverty_%	0.12	0.23	0.41	-0.8	1			
BA_Degree+	0.03	-0.06	-0.1	0.33	-0.28	1		
Population	0.16	-0.01	0.12	0.14	-0.08	0.21	1	
Area	0.42	0.26	0.26	-0.1	0.09	0.01	0.36	1

Figure 8. Generally, explanatory variables are uncorrelated with each other.

A general concern with work of this nature is causality. If we wish to draw any conclusions about if curiosity causes belief in objective scientific phenomena then these beliefs must not affect curiosity. Theoretically, we may expect this to be broadly true – since curiosity is a personality characteristic it is unlikely to be affected by beliefs in specific phenomena. However, more testing is needed to truly establish a causal link and over-interpretation of our results as causal may result in a biased estimate of the true causal effect. This is a useful subject for further research.

Results and Discussion

Statistical results

Below, we present the results of our statistical analysis. Results are shown first at the state level (i.e. one observation per state), yielding a small sample size. State level scientific literacy and curiosity scores are the combined averages of the county level DDR data for each state. Second and third we present results at the county level using all available counties using DDR and BERT respectively. We also do analysis using just the subset of counties that have a sufficiently large number of articles and authors as a check. This is provided in an appendix but has similar results.

Each table first presents four columns with a simple model including scientific literacy and curiosity only for each of our four dependent variables. The second set of four columns includes models with the best explanatory variables measured by statistical significance at the 10% level. The final set of four columns just includes all explanatory variables as a further check.

Dependent variables:	NO OTHER CONTROLS				LIMITED CONTROLS				ALL CONTROLS			
	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends
	CC happening - 2018	CC happening - 2016	Rates (%)	CC real	CC happening - 2018	CC happening - 2016	Rates (%)	CC real	CC happening - 2018	CC happening - 2016	Rates (%)	CC real
Explanatory variables												
literacy_mean	0.19	0.19	-0.21	-0.01	0.09	0.09	-0.2	0	0.08	0.07	-0.2	0.08
	0.2	0.2	0.16	0.96	0.36	0.39	0.19	0.98	0.48	0.56	0.21	0.59
curiosity_mean	0.1	0.1	-0.08	0.3	0.18	0.18	-0.08	0.3	0.18	0.22	-0.1	0.25
	0.51	0.51	0.62	0.05	0.07	0.09	0.61	0.05	0.12	0.07	0.61	0.14
bachelor_plus_pct					0.67	0.66	-0.02	0.03	0.68	0.66	-0.1	-0.02
					0	0	0.89	0.82	0	0	0.56	0.88
pop_2018					0.32	0.28	0.21	0.19	0.31	0.27	0.35	0.25
					0	0.01	0.16	0.2	0	0.02	0.07	0.1
unemp									-0.01	-0.1	-0.08	0
									0.92	0.46	0.69	0.98
med_hh_inc									0.05	0.02	-0.01	0.05
									0.74	0.92	0.95	0.82
pov_pct									0	-0.02	0.06	0.14
									1	0.9	0.79	0.54
area									0.04	0.04	-0.27	-0.38
									0.73	0.7	0.16	0.02
N	47	47	46	47	47	47	46	47	47	47	46	47
R-sq	0.055	0.055	0.058	0.088	0.636	0.596	0.103	0.126	0.64	0.609	0.165	0.268
AIC	278.83	271.76	213.73	30.59	238.04	235.88	215.51	32.6	245.53	242.32	220.18	32.27
BIC	284.39	277.31	219.21	36.15	247.29	245.13	224.65	41.85	262.18	258.97	236.64	48.92

Standardized beta coefficients; p-values in second row

Figure 9. Statistical results at the state level using the DDR scores.

Dependent variables:	NO OTHER CONTROLS				LIMITED CONTROLS				ALL CONTROLS			
	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends
	CC happening - 2018	CC happening - 2016	Rates (%)	CC real	CC happening - 2018	CC happening - 2016	Rates (%)	CC real	CC happening - 2018	CC happening - 2016	Rates (%)	CC real
Explanatory variables												
literacy_all	0.08	0.09	-0.07	-0.09	0.01	0.03	-0.04	-0.06	0	0.03	-0.04	-0.07
	0.21	0.14	0.26	0.15	0.9	0.56	0.5	0.3	0.94	0.6	0.48	0.29
curiosity_all	-0.04	-0.03	-0.1	0.04	0	0	-0.08	0.03	-0.01	-0.01	-0.09	0.02
	0.54	0.57	0.08	0.55	0.97	0.95	0.17	0.63	0.83	0.84	0.14	0.76
unemp					0.18	0.2	0.21	0.01	0.15	0.17	0.2	-0.01
					0	0	0	0.84	0.01	0.01	0	0.86
med_hh_inc					0.15	0.18	0.3	0	0.27	0.27	0.31	0.12
					0.01	0	0	0.95	0	0	0	0.27
bachelor_plus_pct					0.36	0.32	0.06	-0.01	0.34	0.3	0.05	-0.02
					0	0	0.31	0.93	0	0	0.43	0.79
area					0.06	0.02	0.11	-0.05	0.05	0.01	0.1	-0.06
					0.26	0.66	0.05	0.39	0.41	0.93	0.1	0.34
pov_pct									0.17	0.14	0.03	0.15
									0.08	0.14	0.73	0.14
pop_2018									0.06	0.08	0.05	0.03
									0.27	0.17	0.41	0.61
N	284	284	283	284	278	278	278	278	278	278	278	278
R-sq	0.006	0.008	0.017	0.008	0.183	0.168	0.124	0.008	0.195	0.18	0.126	0.017
AIC	1814.87	1789.86	1232.22	-91.48	1723.72	1704.5	1164.05	-96.07	1723.3	1704.31	1167.22	-94.59
BIC	1825.82	1800.81	1243.15	-80.54	1749.12	1729.9	1189.44	-70.67	1755.95	1736.96	1199.87	-61.94

Standardized beta coefficients; p-values in second row

Figure 10. Statistical results at the county level using the DDR scores.

Scientific literacy does not predict beliefs

Consistent with (Kahan et al., 2012), across all models and specifications scientific literacy is not statistically significant at the 10% level. This provides strong evidence

Dependent variables:	NO OTHER CONTROLS				LIMITED CONTROLS				ALL CONTROLS			
	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends
	CC happening - 2018	*CC happening - 2016*	*Rates (%)*	*CC real*	*CC happening - 2018*	*CC happening - 2016*	*Rates (%)*	*CC real*	*CC happening - 2018*	*CC happening - 2016*	*Rates (%)*	*CC real*
Explanatory variables												
literacy_bert_all	0.05	0.05	-0.06	-0.08	-0.01	-0.01	-0.04	-0.07	-0.03	-0.02	-0.06	-0.07
	0.42	0.45	0.34	0.24	0.81	0.84	0.51	0.28	0.68	0.75	0.37	0.3
curiosity_bert_all	0.01	0.02	-0.06	0.01	0.04	0.06	-0.07	0.03	0.04	0.05	-0.07	0.02
	0.86	0.74	0.37	0.93	0.46	0.36	0.28	0.71	0.49	0.4	0.28	0.76
unemp					0.17	0.18	0.22	-0.01	0.16	0.17	0.21	-0.01
					0.01	0	0	0.91	0.01	0	0	0.89
med_hh_inc					0.27	0.28	0.31	0.13	0.27	0.27	0.31	0.12
					0	0	0	0.21	0	0	0	0.27
pov_pct					0.16	0.14	0.02	0.16	0.16	0.14	0.03	0.16
					0.08	0.14	0.81	0.13	0.08	0.14	0.79	0.13
bachelor_plus_pct					0.36	0.33	0.07	-0.02	0.34	0.31	0.05	-0.02
					0	0	0.23	0.8	0	0	0.4	0.75
pop_2018									0.06	0.08	0.05	0.04
									0.3	0.19	0.4	0.58
area									0.05	0.01	0.11	-0.06
									0.37	0.85	0.07	0.33
N	284	284	283	284	278	278	278	278	278	278	278	278
R-sq	0.003	0.004	0.01	0.006	0.189	0.175	0.111	0.013	0.197	0.181	0.128	0.017
AIC	1815.7	1791.13	1234.23	-90.9	1721.41	1701.93	1167.97	-97.44	1722.84	1703.88	1166.81	-94.52
BIC	1826.64	1802.08	1245.17	-79.96	1746.81	1727.33	1193.36	-72.04	1755.49	1736.53	1199.46	-61.87

Standardized beta coefficients; p-values in second row

Figure 11. Statistical results at the county level using the BERT scores.

from a range of text based methods of geographic scientific literacy that scientific literacy is unrelated to belief in scientific phenomena.

This conclusion further contributes to the literature's current published research suggesting that general knowledge-based scientific education is unlikely to generate a populace with greater belief in objective scientific phenomena or concepts. This evidence provides a useful further exploration of this topic as (a) it is evaluated at the whole country level rather than at the smaller scale survey level as in previous research; and (b) it is based on data from an indirect, live dataset expressing people's characteristics rather than the primed environment of a survey study.

Curiosity has limited ability to predict beliefs

We do not find consistent evidence that curiosity explains beliefs in scientific phenomena. This finding is in contrast to previous work in the research that suggests this as a potential factor causing beliefs (Kahan, Landrum, Carpenter, Helft, & Hall Jamieson, 2017).

In the county-level data, we find that curiosity only shows promise predicting beliefs for the vaccination rate based belief data (statistically significant at the 10% level; note some caution needed, since here the dependent variable is only at the state level) – however here the coefficient on curiosity is negative, suggesting unintuitively that curiosity leads to less belief in scientific phenomena. With the BERT data, curiosity fails to predict beliefs in any model and with any measure of beliefs.

Finally, at the state level, there is some evidence that curiosity predicts beliefs. When controlling for other state explanatory variables curiosity does predict beliefs at the 10% or more levels of significance. The change in significance when adding explanatory factors suggests some omitted variable bias when these are not included – this bias would have to negate the effect of curiosity on beliefs. Although as in the existing literature, curiosity does achieve more here than the scientific literacy measure, this is far from a consistent finding. The fact that this result is not at all corroborated at the county level suggests further concerns about how well curiosity is associated with beliefs.

Several reasons could explain why the finding is present at the survey level in existing research but not in our analysis. One explanation is that although curiosity is part of the underlying psychological mechanisms that lead to beliefs, these effects are washed out at a more aggregated level, or are negated by some other real world process that is not being replicated in the laboratory. Nevertheless, these inconsistencies warrant investigation in future research.

Conclusion

In this study, we investigated the relationship between belief in objective scientific phenomena and population levels of scientific literacy / curiosity measured via local news articles. While previous research on scientific curiosity mainly focused on creating self-reported inventories and scales, we attempted to create a more objective, text-based measure of this psychological concept. Using this measure, we tested the hypotheses that 1) scientific literacy does not predict belief in scientific facts (Drummond & Fischhoff, 2017; Kahan et al., 2012), and 2) scientific curiosity is better able to predict belief in these facts (Kahan et al., 2012).

Our results are consistent with the hypothesis that there is no relationship between scientific literacy and belief in climate change or efficacy of vaccinations. However, we find that text-based measures of scientific curiosity have limited predictive power in determining belief in scientific phenomena, providing only partial support for Kahan et al.'s (2012) findings. One potential explanation for this discordance which should be the focus of future work is that although curiosity is part of individual level mechanisms in belief adjustment, population level factors may offset this.

A further explanation that warrants future research is that only using partisan split to explain the polarization of scientific beliefs limits our understanding of the issue. Psychologists posited that closer psychological distance can encourage concerns for climate change (e.g. day feels warmer than usual; Li, Johnson, and Zaval (2011)) but can backfire if used to provoke fear, resulting in avoidant emotional reactions (for review see McDonald, Chai, and Newell (2015)). Another popular explanation is the Cultural Cognition Thesis (CCT) which argues that individuals tend to form perceptions of how society should be structured that coincide with values of groups with which they identify – the more individuals adopt hierarchical and individualistic cultural values, less they believe in climate change (Hornsey, Harris, Bain, & Fielding, 2016; Kahan, Braman, Cohen, Gastil, & Slovic, 2010). Therefore clearly more psychological research is warranted here. Other limitations of the work include the use of a single news source (Patch), lack of locations available in this news source (only 285/3000+ counties), and possible annotation bias due to a lack of outside annotators.

References

Adhikari, A., Ram, A., Tang, R., & Lin, J. (2019). Docbert: Bert for document classification. *arXiv preprint arXiv:1904.08398*.

Cook, J. (2019). Understanding and countering misinformation about climate change. In *Handbook of research on deception, fake news, and misinformation online* (pp. 281–306). IGI Global.

Cook, J., Lewandowsky, S., & Ecker, U. K. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PloS one*, *12*(5), e0175799.

De Aguiar, M., & Stearns, J. (2015). Declaration of dependence: Communities and news organizations working together will transform local journalism. here's how. *The Local News Lab Blog*.

De Pinto, J., Backus, F., & Salvanto, A. (2019, Sep). *Most Americans say climate change should be addressed now — CBS News poll.* `https://www.cbsnews.com/news/cbs-news-poll-most-americans-say-climate-change-should-be-addressed-now-2019-09-15/`. (Online; accessed 1 December 2019)

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the north american chapter of the association for computational linguistics: Human language technologies, volume 1 (long and short papers)* (pp. 4171–4186).

Drummond, C., & Fischhoff, B. (2017). Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. *Proceedings of the National Academy of Sciences*, *114*(36), 9587–9592.

Friedman, L. (2019, Nov). John Kerry Launches Star-Studded Climate Coalition. *New York Times*. Retrieved from `www.nytimes.com/2019/11/30/climate/john-kerry-climate-change.htm`

Garten, J., Hoover, J., Johnson, K. M., Boghrati, R., Iskiwitch, C., & Dehghani, M.

(2018). Dictionaries and distributions: Combining expert knowledge and large scale textual data content analysis. *Behavior research methods*, *50*(1), 344–361.

Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? evidence from us daily newspapers. *Econometrica*, *78*(1), 35–71.

George, L. M., & Waldfogel, J. (2006). The new york times and the market for local newspapers. *American Economic Review*, *96*(1), 435–447.

Greenslade, R. (2012, Jun). *Local news crisis: why newspapers remain so important to the public.* `https://www.theguardian.com/media/greenslade/2012/jun/25/marketingandpr-local-newspapers`. The Guardian. (Online; accessed 1 December 2019)

Hornsey, M. J., Harris, E. A., Bain, P. G., & Fielding, K. S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change*, *6*(6), 622.

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental review*, *32*(2), 125–160.

Kahan, D. M., Braman, D., Cohen, G. L., Gastil, J., & Slovic, P. (2010). Who fears the hpv vaccine, who doesn't, and why? an experimental study of the mechanisms of cultural cognition. *Law and human behavior*, *34*(6), 501–516.

Kahan, D. M., Landrum, A., Carpenter, K., Helft, L., & Hall Jamieson, K. (2017). Science curiosity and political information processing. *Political Psychology*, *38*, 179–199.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature climate change*, *2*(10), 732.

Kashdan, T. B., & Roberts, J. E. (2004). Trait and state curiosity in the genesis of intimacy: Differentiation from related constructs. *Journal of Social and Clinical Psychology*, *23*(6), 792–816.

Kashdan, T. B., & Roberts, J. E. (2007). Social anxiety, depressive symptoms, and post-event rumination: Affective consequences and social contextual influences.

Journal of Anxiety Disorders, *21*(3), 284–301.

Kashdan, T. B., Stiksma, M. C., Disabato, D. J., McKnight, P. E., Bekier, J., Kaji, J., & Lazarus, R. (2018). The five-dimensional curiosity scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *Journal of Research in Personality*, *73*, 130–149.

Landrum, A., Hilgard, J., Akin, H., Li, N., & Kahan, D. (2016). Measuring interest in science: The science curiosity scale. In *Cogsci.*

Leiserowitz, A., Maibach, E., Roser-Renouf, C., & Hmielowski, J. (2012). Global warming's six americas, march 2012 & nov. 2011. *Project on Climate Change Communication.*

Li, Y., Johnson, E. J., & Zaval, L. (2011). Local warming: Daily temperature change influences belief in global warming. *Psychological science*, *22*(4), 454–459.

Litman, J. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & emotion*, *19*(6), 793–814.

Litman, J. A. (2009). Curiosity and metacognition. *Metacognition: New research developments*, 105–116.

Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of personality assessment*, *80*(1), 75–86.

McDonald, R. I., Chai, H. Y., & Newell, B. R. (2015). Personal experience and the 'psychological distance'of climate change: An integrative review. *Journal of Environmental Psychology*, *44*, 109–118.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111–3119).

Napoli, P., & Mahone, J. (2019, Sep). *Local newspapers are suffering, but they're still (by far) the most significant journalism producers in their communities.* `https://www.niemanlab.org/2019/09/local-newspapers-are-suffering-but-theyre-still-by-far-the-most`

-significant-journalism-producers-in-their-communities/. Neiman Lab. (Online; accessed 1 December 2019)

Naylor, F. D. (1981). A state-trait curiosity inventory. *Australian Psychologist*, *16*(2), 172–183.

Pagliardini, M., Gupta, P., & Jaggi, M. (2018). Unsupervised learning of sentence embeddings using compositional n-gram features. In *Proceedings of the 2018 conference of the north american chapter of the association for computational linguistics: Human language technologies, volume 1 (long papers)* (pp. 528–540).

Patel, M., Lee, A. D., Clemmons, N. S., Redd, S. B., Poser, S., Blog, D., . . . others (2019). National update on measles cases and outbreaks—united states, january 1–october 1, 2019. *Morbidity and Mortality Weekly Report*, *68*(40), 893.

Pennebaker, J. W. (2011). The secret life of pronouns. *New Scientist*, *211*(2828), 42–45.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, *71*(2001), 2001.

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (emnlp)* (pp. 1532–1543).

Spektor-Levy, O., Baruch, Y. K., & Mevarech, Z. (2013). Science and scientific curiosity in pre-school—the teacher's point of view. *International Journal of Science Education*, *35*(13), 2226–2253.

Supran, G., & Oreskes, N. (2017). Assessing exxonmobil's climate change communications (1977–2014). *Environmental Research Letters*, *12*(8), 084019.

van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D., & Maibach, E. W. (2015). The scientific consensus on climate change as a gateway belief: Experimental evidence. *PloS one*, *10*(2), e0118489.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998–6008).

Weible, J. L., & Zimmerman, H. T. (2016). Science curiosity in learning environments: developing an attitudinal scale for research in schools, homes, museums, and the community. *International Journal of Science Education*, *38*(8), 1235–1255.

Supplemental Material

County subsample – counties with many authors/articles

	NO OTHER CONTROLS				LIMITED CONTROLS				ALL CONTROLS			
Dependent variables:	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends	Yale Survey	Yale Survey	Vaccinations	Google trends
	CC happening - 2018	*CC happening - 2016*	*Rates (%)*	*CC real*	*CC happening - 2018*	*CC happening - 2016*	*Rates (%)*	*CC real*	*CC happening - 2018*	*CC happening - 2016*	*Rates (%)*	*CC real*
Explanatory variables												
literacy_all	0.12	0.13	0.32	-0.15	0.04	0.05	0.27	-0.15	0.02	0.02	0.3	-0.21
	0.16	0.14	0	0.08	0.6	0.55	0	0.09	0.86	0.82	0	0.02
curiosity_all	-0.11	-0.13	-0.34	-0.15	-0.02	-0.03	-0.26	-0.17	-0.03	-0.03	-0.24	-0.19
	0.21	0.14	0	0.07	0.78	0.72	0	0.05	0.75	0.7	0	0.03
unemp					0.21	0.22	0.16	0.03	0.17	0.18	0.14	-0.01
					0.01	0	0.04	0.72	0.03	0.03	0.08	0.9
med_hh_inc					0.17	0.23	0.32	-0.04	0.29	0.35	0.44	0.02
					0.03	0	0	0.68	0.02	0	0	0.91
bachelor_plus_pct					0.35	0.29	0.05	-0.04	0.34	0.28	0.03	-0.03
					0	0	0.53	0.63	0	0	0.65	0.68
pov_pct									0.16	0.16	0.19	0.06
									0.18	0.18	0.12	0.65
pop_2018									0.05	0.06	0.1	-0.06
									0.56	0.48	0.17	0.5
area									0.05	0.05	-0.14	0.21
									0.55	0.57	0.08	0.02
N	171	171	171	171	169	169	169	169	169	169	169	169
R-sq	0.014	0.018	0.118	0.068	0.204	0.193	0.216	0.073	0.218	0.208	0.248	0.105
AIC	1108.23	1090.14	660.32	-113.13	1061.62	1046.67	639.41	-104.93	1064.65	1049.49	638.47	-104.76
BIC	1117.65	1099.56	669.75	-103.71	1080.4	1065.45	658.19	-86.15	1092.82	1077.66	666.64	-76.59

Standardized beta coefficients; p-values in second row

Figure 12. Statistical results at the county level using the DDR scores - counties with larger datasets only (> 100 articles, > 50 authors).

Curiosity and Scientific Curiosity Inventories and Scales

- Curiosity and Exploration Inventory (CEI; (Kashdan & Roberts, 2004))
- Curiosity and Exploration Inventory II (CEI-II; (Kashdan & Roberts, 2007))
- Epistemic Curiosity Inventory (EC; (J. A. Litman & Spielberger, 2003))
- Five-Dimensional Curiosity Scale (5DC; (Kashdan et al., 2018))
- Melbourne Curiosity Inventory (MCI; (Naylor, 1981))
- Scientific Curiosity in Learning Environment Scale (SCILE; (Weible & Zimmerman, 2016))
- Scientific Curiosity Scale (SCS; (Kahan et al., 2017))

- The Science Curiosity Scale ((Landrum, Hilgard, Akin, Li, & Kahan, 2016))

Scientific Curiosity Word Dictionaries

D1. museum; outdoor; nature; scientist

D2. fact; unpredictable; evidence; experiment

D3. explore; cause; visit; question

D4. solve; seek; explore; wonder

Annotation Guide

Measure(1): Scientific literacy: indicate whether the author of the article is expressing a degree of scientific awareness or knowledge. This may include but is not limited to:

- Use of scientific terminology.
- Use of scientific framework in analysis of an issue or topic.
- References to scientific literature.

Measure(2): Scientific curiosity: indicate whether author is demonstrating curiosity, separate from Scientific literacy. For example:

- Expressing learning things as exciting
- Trying to learn about the unknown/uncertain
- Expressing interest in exploring new places, ideas, and things.
- Some level of speculation (e.g. "I wonder if...", "Why is that...", "What causes...")
- Expressing that current knowledge is insufficient

Team Member Contributions

Amabel. Worked primarily on literature review on curiosity and the importance of media. Created the scientific curiosity word dictionaries and the annotation guide based on literature and a collection of curiosity/scientific curiosity inventories/scales. Annotated articles. Wrote report section on introduction.

Nikos. Worked primarily on text analysis. Ran and evaluated the DDR method on the available data. Built the BERT-based model and used part of the DDR outputs as labeled data to fine-tune it. Produced the final county-level representations for the 2 concepts of interest based on those methods. Created the initial 4 chunks of articles for the group members to annotate. Annotated articles. Wrote report section on text methods (DDR and BERT).

Sarah. Worked primarily on the data collection phase. Modified and ran the code to collect the Patch around-town and lifestyle articles. Modified and ran the code to append county designations and hand annotated missing counties. Annotated articles. Wrote report section on data.

Thomas. Worked primarily on belief data collection and statistical analysis. Collected data and researched different potential data sources at the state and county level. Conducted statistical analysis, structured the findings and drew conclusions. Wrote report section on statistical analysis and results.