r/science: Challenges and Opportunities for Online Science Communication

Ridley Jones
University of Washington
Seattle, United States
rajone@uw.edu

Lucas Colusso
University of Washington
Seattle, Washington
colusso@uw.edu

Katharina Reinecke
University of Washington
Seattle, Washington
reinecke@uw.edu

Gary Hsieh
University of Washington
Seattle, Washington
garyhs@uw.edu

ABSTRACT

Online discussion websites, such as Reddit's r/science forum, have the potential to foster science communication between researchers and the general public. However, little is known about who participates, what is discussed, and whether such websites are successful in achieving meaningful science discussions. To find out, we conducted a mixed-methods study analyzing 11,859 r/science posts and conducting interviews with 18 community members. Our results show that r/science facilitates rich information exchange and that the comments section provides a unique science communication document that guides engagement with scientific research. However, this community-sourced science communication comes largely from a knowledgeable public. We conclude with design suggestions for a number of critical problems that we uncovered: addressing the problem of topic newsworthiness and balancing broader participation and rigor.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*.

KEYWORDS

ACM Reference Format:
Ridley Jones, Lucas Colusso, Katharina Reinecke, and Gary Hsieh. 2019. r/science: Challenges and Opportunities for Online Science Communication. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3290605.3300383

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300383

1 INTRODUCTION

Effective and relevant science communication allows the public to find out about research developments, and make better decisions in their own lives [53]. For scientists, it facilitates the diffusion and application of their research findings, provides useful feedback and perspectives, and increases the recognition and impact of their work [5]. However, the accelerating rate of scientific publishing and the growing specialization in scientific vocabulary raise the barrier for mutually meaningful science communication between scientists and the public [6, 39]. Moreover, even when information is clear and readily available, certain aspects of human cognition can keep people from believing and accepting scientific knowledge [28].

With advances in content dissemination technologies, scientific research can be shared as quickly as it takes to type out a tweet. Significant proportions of scientists now use social media and blogs to discuss science; concurrently, the importance they attribute to gaining media coverage for their work has grown [42]. However, much of this communication stays relatively confined to scientists' personal or professional circles [31]. Building an audience beyond existing circles is often perceived as difficult and labor-intensive [15, 26].

Internet-based technologies have the potential to enable communication between researchers, the general public, government, and all other stakeholders in the processes of science. The importance of this interaction is emphasized by the so-called dialogue and participation models of science communication [32]. Despite this potential, how to foster meaningful science discussions online remains relatively unknown [5]. This raises several questions: Who participates and why, and what types of science dialogues are occurring?

What are some of the challenges and opportunities to foster more effective science communication online?

We sought to answer these questions by studying r/science, the largest science discussion forum on Reddit, with over 19 million subscribers.[1] We performed a quantitative analysis of r/science posts and comments and conducted in-depth interviews with 18 community members to understand the dynamics and characteristics of the community.

Our results reveal a nuanced, collaborative picture of real-world science communication not fully articulated by any of the prevailing models. Active participants in r/science collaborate to do what we call "information stewardship" work: the use of sources, knowledge, skill, and site features to produce a community-sourced and community-centered science communication document. (Compare to [54] for a strikingly similar approach in the video medium.) These collaborative efforts produce a unique science communication document that guides behavior, offers useful perspectives and critiques on the research, and adds additional information about the research that can improve the quality of engagement. However,communication with the general public is still lacking. Instead, r/science's active participants are typically those already interested and invested in science and scientific activities. The stories that are amplified to the broader Reddit readership tend to fall within a few more obviously "newsworthy" topics. To overcome this and to support the information stewarding work taking place, we recommend employing features that allow individual posts to be contextualized in the broader process of scientific research and discovery. Borrowing from the nomenclature of Reddit, we call this a move from aggregation to integration.

2 RELATED WORK

To lay the foundation for our research, we turn to two different but related domains: science communication and citizen journalism.

Science Communication

The study of science communication has developed a variety of theoretical models attempting to answer the question: What are the best ways to bring advances in scientific knowledge to everyone? Burns, O'Connor and Stocklmayer define contemporary science communication as "the use of appropriate skills, media, activities and dialogue" to produce one or more of the following personal responses to science: Awareness, Enjoyment, Interest, Opinion-forming, and Understanding [7]. This variety of goals helps to illustrate how many audiences must be considered.

The *deficit model* was the default approach of science communicators until the early 1990s, when the journal *Public Understanding of Science* was established following the 1986 release of the Royal Society's report of the same title [9]. This model holds that the problem of science communication is simply one of information transmission, or in other words a deficit of knowledge. However, psychology and communications research has shown that simply being given factual scientific information is insufficient to ensure that recipients can apply scientific reasoning to problems [24]. Moreover, for at least some issues, highly educated people are even more polarized in their beliefs [28]. Awareness of this disconnect has led researchers to try to understand what is actually important for the public to know and think—not simply a list of science facts, but an understanding of scientific reasoning and processes [24] and an endorsement of scientific inquiry as a way of knowing [12].

Two currently popular models of science communication that have emerged from this work are the *dialogue* and *public engagement models*. These models both rely on the public themselves being active participants in science communication. According to the *dialogue model*, the process of science communication is one of exchange between scientists and the public [32]. Scientists are accountable to explain their work in a comprehensible way, and the public is informed about scientific research and empowered to talk back [32].

The *public engagement model* (also known as the public participation model) takes this a step further; it requires inclusion of non-scientists in actual scientific work or decision-making [8]. Those producing and consuming scientific information are both learning and contributing; the public engagement model therefore aims to reduce conceptual and structural barriers between scientists and the public [11]. There is a place not only for imparting knowledge, but also for humanizing the people and processes that make up scientific research [11, 30]. Citizen science is the classic example of public engagement (e.g., crowdsourcing the analysis of astronomical data [41]); recent efforts in the HCI community involve the public in the creation of scientific hypotheses [37] and guiding the public through learning domain-specific content to support hypothesis development [38].

Embedded in the definition of the dialogue model are assumptions about what "dialogue" means and who the communicating parties are. Do we consider the dialogue as primarily happening between research producers and a lay general public, specifically about the research? If this is the case, we can also assume that the specific research producers are required to translate their specialized technical vocabulary and processes to non-technical audiences. If we understand this dialogue as happening more broadly—between a research *community* and a variety of publics [7]—then the activities and responsibilities can be quite different. Science communication definitions do not make these specifics clear, and this may have consequences for the interventions they propose.

[1] https://www.reddit.com/r/science/

In our study of r/science, we seek to lay out who these communicating parties are and what kinds of dialogues they are engaging in.

Online Science Communication and Citizen Journalism

The internet has been recruited to assist in many kinds of science communication efforts, including blogs [43], demonstration sites for specific research projects [10], general social media/network sites (e.g., Twitter/Facebook [40]), video-sharing sites (e.g., YouTube [55]), and link aggregation communities (e.g., Reddit). However, the study of science communication as an online community activity, with resulting empirically testable models, has been limited [4]. To bridge this gap, we draw upon work in citizen journalism, as it offers resonant perspectives on how specialists and non-specialists collaborate in disseminating knowledge.

Traditionally, journalists and news outlets have acted as curators and have great control over what becomes "news", giving them the role of gatekeepers [48]. This means that the choices these journalists make influence how people make sense of the world, especially for topics that require significant interpretation for the average reader to understand. This is evident in, for example, how images are portrayed [21, 56] or how scientific uncertainty is depicted [22].

However, this is not the complete story. Individual journalists, while important, are often not the final arbiters of what becomes news, nor do they control sources of information; most get many of their stories from wire services and are subject to "routine" or structural forces in gatekeeping in their news agencies [47]. Further, in the online context, members of the public who rate, share, and comment on news stories functionally contribute to gatekeeping as well, since they preapprove, promote, or contextualize the information to their own networks, hence the term "secondary gatekeeping" [48]. Goode's broad definition of "citizen journalism" encompasses many pratices, such as providing on-the-ground accounts of events, performing independent investigation or otherwise acting as a watchdog, giving personal perspectives, and even sharing and commenting on news [20]. Gatekeeping still exists, but it has become a complex collection of activities that many groups of people participate in, and the role of journalists has become more widely distributed and heterogeneous.

Lingering Participation Issues

However, participation remains a key challenge. Many researchers worry about online comment sections derailing the conversation or promoting unscientific beliefs [57]. Moreover, scientists often perceive those who are active in public communication as less serious scholars [13]. Even when they do consider public outreach to be valuable and important, researchers are not usually incentivized to dedicate the time and effort needed to do it successfully [15, 50]. Non-scientists also face difficulties finding and engaging with scientific information, even when it is available. Difficulty in understanding what scientific uncertainty means [45], as well as a wide range of cognitive mechanisms that help us protect existing beliefs even when they come under direct empirical attack [28], can make it hard to process the deluge of scientific information we now have available.

3 STUDYING R/SCIENCE

r/science is a science news and discussion subforum ("subreddit") on the social news aggregator Reddit. Users can post items (links and text; in other subreddits images and video are also allowed) that represent recent scientific research: either direct links to research papers, or reputable news items that link to the papers. They can also comment on posts, and "upvote" or "downvote" posts and comments. A post or comment's *score* is the difference between the upvotes and downvotes it receives. Each post must be tagged with a *topic flair* (the topic of the post). Users of r/science can also apply to receive *author flair*, a verified scientific credential such as a degree or institutional affiliation. This flair appears next to their username on every post and comment they make in the subreddit (see Figure 1 for an example of a post and comments containing this information). In addition, Reddit users can accumulate *karma* by posting and commenting and receiving more upvotes on these posts and comments. Karma is represented as a number and is visible on a person's user page, but not on posts and comments.

r/science is one of a collection of science-related subreddits, each with somewhat different aims. r/askscience is geared more toward question and answer, and r/everythingscience is more casual and broad in scope. r/science is the oldest and largest, and attempts to balance the largest number of goals, so we chose to concentrate our efforts there.

Similar to other subreddits, r/science has also occasionally hosted "Ask Me Anything" posts (AMAs). An AMA is a comment dialogue between a researcher or research group and r/science members. The post title announces who the researchers are and commenters can ask them questions about their research or interest. The AMA series was shut down in May 2018 due to a drop-off in viewership, attributed to a change in post rankings on Reddit [52]. It is currently unclear whether it will be revived.

To study how science communication is practiced in this community, we paired quantitative analyses of its content with an interview study. We had the following research questions: (1) Who are the participants of r/science and what drives their participation?, (2) What type of science is being

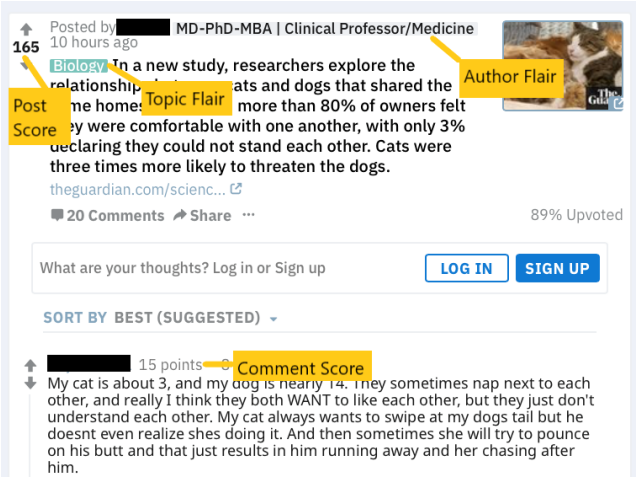

Figure 1: An example post and comment on r/science.

shared on r/science?, and (3) How is dialogue being facilitated?

Analysis of r/science Log-Data and Comments. We first conducted a descriptive analysis of r/science log-data paired with a qualitative analysis of r/science comments. We obtained all posts made to r/science in 2016 (the most recent full year of data when we began our analysis) using the Reddit API for Python [44]. We also scraped as many comments as we could under the constraints of the API. Though we attempted to overcome the constraints of the API as discussed in [17], we acknowledge that some data may be missing. Our full dataset comprises 11,859 posts and 345,789 comments.

To find out who the users of r/science are (our first research question), we started with an analysis of the frequency of author flairs, and patterns of posting by individuals.

To answer our second research question, what type of science is being shared on r/science, we analyzed the posts' *topic flair*. Because the topic flairs vary in their specificity, we grouped all topic flairs into 16 high-level topics (see Table 3).

To answer our third research question, what type of discussions users are having, we analyzed all comments made to a random sample of 100 posts, stratified by comment count. From these threads, we qualitatively coded all first- and second-level comments (replies to the original post, and responses to these replies). The comment codes along with their frequencies of occurance are listed in Table 1. In all, we coded 2,056 comments. For each of these comments, we also calculated the total number of replies it received (the size of its entire reply tree) as a rough measure of how much dialogue each comment stimulated.

Using a modified grounded theory approach [18], we developed a comment codebook, indicating the kind of action taking place in each comment. Four authors individually coded the same 30 comments and discussed any discrepancies until we reached full agreement on 24 final codes. Two authors then coded all 2,056 comments. Since many comments contained multiple components, we allowed two codes per comment, marking one as primary and one as secondary.

Interviews. To gain deeper insight into what motivations and strategies participants bring to r/science, and to triangulate our quantitative analysis, we conducted semi-structured interviews with 18 r/science users representing four groups of people: r/science moderators (M), those who commented or posted in r/science in 2016 (C,P), and those who had previously read r/science posts but had not posted to r/science or considered themselves members of the community (R). Interviews were conducted between January and March 2018. All but one were conducted via phone (one in person), and typically lasted 45-60 minutes.

To recruit the moderators, we used the "message the moderators" function on Reddit. To recruit commenters and posters, we directly contacted an equal number of randomly selected users stratified by activity level. Finally, to recruit general Reddit users, we repeatedly obtained a random comment on a random subreddit (see Supplementary Materials for more detail), and contacted that comment's author.

All r/science community member participants (C,P,M) were asked a core list of questions about their motivations for participating in r/science, what they thought the purpose of r/science was, their usage patterns, how they make use of comments and interface elements, what they think makes r/science effective or ineffective, and what their general views of science communication are. Moderators were additionally asked about their path to becoming moderators and what had informed some of the policy decisions that had happened during their tenure, as well as structural questions about the subreddit. Non-community member redditors (R) were asked about their general Reddit usage, their knowledge and opinions of r/science, and their perspectives on science communication.

Interviews were transcribed verbatim. All four authors generated initial codes for a subset of interviews and discussed the codes. Two authors then coded all interviews and discussed any discrepancies. We then followed an iterative thematic analysis method to cluster codes into themes. Some quotes presented below have been slightly modified for readability.

4 RESULTS

Who are the participants on r/science?

Our collected comments and posts from 2016 were submitted by 2,565 unique posters and 107,033 unique commenters. Much like existing work that has demonstrated

Table 1: Qualitative codes for comments, frequencies, mean scores, and mean number of responses.

Category	Code	Description	Count	Mean Score	Mean Replies
Conversational Moves	Personal	Personal questions or stories or responses to them	252	4.0	1.4
	Argument	Disagreeing or arguing with another commenter	154	12.4	2.1
	Banter	Playful, jokey comment	150	5.4	0.5
	Agreement	Agreeing with parent comment	64	7.6	1.7
	Kudos	Offering thanks or kudos	45	11.8	0.4
	Credential	Claiming a credential or academic association (even loose) with the research	19	4.6	0.3
	Capitulation	Change of opinion/Capitulation to someone's argument	3	5.7	0.3
	Self-clarification	Clarifying or correcting self	1	3	0
Critique	Q-Validity	Questioning the validity of the findings/offering alternative hypotheses	53	25.9	4.2
	Q-Importance	Questioning the novelty, importance, or value of the findings	47	6.1	1.9
	Q-Wording	Questioning the wording of the title or news article	33	19.5	2.0
	Q-Methods	Questioning the methods used	32	94.9	5.5
	General Questioning	Questioning the field, domain, or larger research process	24	3.1	2.8
	Q-Ethics	Ethical issues with the article	16	4.2	1.5
	Critique of Posts	Critique of Poster and their posts (including the wording of the post title)	6	28.9	0.5
	Q-Content	Questioning the presentation or content of the article	5	17.2	0.5
	Q-Source	Questioning the source	5	10	0.3
Positive Engagement	Info	Providing or linking to additional information	350	9.0	1.4
	Reasoning	Extending, applying, or reasoning about the research	222	8.1	2.0
	Educate	An educational response to another comment, including synopsis/simplification	205	9.5	1.6
	Support	Expressing support or excitement about the research	74	4.2	2.0
	Process	Discussing the process of science	63	10.7	7.0
	Summary	Simplifying or summarizing the research (unprompted)	4	176.5	2.0
Information-Seeking	Inquiry	Questions about the research or related work	544	8.5	1.8
Structural	Link	Link directly to the paper or abstract	60	13.4	1.2
	Quote	Quotes from the paper or news article	24	39	2.2
Other	Meta	Discussion of r/science structure or norms (e.g., moderator comments)	38	5.4	0.8
	Other	Other very rare or off-topic comments	14	2.4	3.0

Table 2: Interview participants in our study.

Label	Education	Occupation
M1	Graduate	Professor
M2	Graduate	Science Outreach manager
M3	Graduate	Industry Scientist
P1	Graduate	University Marketing Director
P2	Bachelor	Bookkeeper
P3	Graduate	Industry Scientist
P4	Graduate	Science Writer
P5	Graduate	Retired
P6	Bachelor	Geologist
P7	Bachelor	Laboratory Technician
C1	Graduate	Graduate Student
C2	Graduate	Graduate Student
C3	Graduate	Science outreach professional
C4	Bachelor	Graduate Student
R1	Bachelor	Construction Manager
R3	Bachelor	Documentation Writer
R4	Bachelor	Programmer

inequality in contribution levels to peer production sites (e.g., [2, 16, 25, 34, 36]), we found that the majority of the contributions are made by a small minority of people. Compared to the number of subscribers to r/science in December 2016, 14.5 million, our data suggests that less than 1% of r/science subscribers actively contribute content to r/science. Further, while on average post contributors made 4.6 posts (sd=28.8), 75% of posts (excluding AMAs) were by users who posted five or more times. Most post authors (70%, or 1,791 unique posters) posted only once. Typically, paper authors (i.e., those who did the work described in the linked news articles or academic papers) are not themselves active contributors to the community. In our qualitative comment analysis, disregarding the AMAs, in only five comments (by three different authors, on three of the 100 posts) did people claiming to be involved in the specific research make comments on the post.

An analysis of author flair in our 2016 posts dataset reveals that of the 196 disciplines listed by 305 different flaired authors, the most common were physics (14), neuroscience (13), psychology (10), and medicine and biology (9 each). However, when grouped into 19 main categories of topics, the most common were medicine and anatomy (43), technology and engineering (35), and biology (33).

What drives post and comment contributions? As in other online communities, there are many ways one can participate on r/science. Upvoting, downvoting, and even lurking are all potentially valuable forms of participation [27, 33, 51]. In this section, we focus primarily on the motivations for contributing posts and comments — why people contribute to the content of science discussion on r/science. Through our interviews, we found a variety of motivations for these types of contributions, including wanting to give back, learning about science, and engaging in science discussions, as well as (occasionally) more self-serving motivations.

Reciprocity. This motivation describes a sense of wanting to give back–both to Reddit and to (if applicable) the social

and economic system that allowed them to become scientists in the first place. As a moderator put it following an anecdote about how Reddit helped them find their lost dog:

> But then I felt indebted to this community. How can I give back? I won't know the hottest clubs in [city], or contribute content in that way, but I'm a scientist so maybe I can help with that. M1

A user who has previously posted on r/science further revealed:

> The citizens of my state gave me tuition money to go get degrees and I went to the state school system, and the federal government gave me grant money to do things. So I feel like it's community responsibility to take my knowledge back to the public, especially on controversial issues and ones that affect policy. P3

Learning and keeping up with science news. Another common motivation was to inform themselves, either broadly or in specific areas of interest. One r/science user, for example, talked about how he approaches finding interesting posts:

> Well typically, I mean I check the web page once or twice a day and see if there's any kind of headlines that catches my attention. I mean, that's one of the things I like about Reddit ... because it aggregates from so many things. I can go through and see if there's anything particularly interesting. C2

Other users seek out specific areas of interest:

> I would dive into topics that are in my field, in my sphere of degree work that I've done. I'm kind of in the stay in your own lane thing. I cannot discuss astrophysics. I can read it as anyone else could, but I would never dive into a debate about something like that. P3

Correcting misunderstandings or misinformation. r/science users were also commonly motivated by a sense of responsibility to correct misinformation, such as P3:

> Somebody is wrong on the Internet. That's sort of the problem, right? [...] I feel that I have a responsibility to try to get the correct information out there.
>
> ... When it does come to one of the areas to which I can speak as an expert, I do my best to cite the relevant literature. Try to address misconceptions. Again where I see some sort of brigading[2] and it's in an area of my expertise, I go out and grab a peer type citation for refute or something along those lines. C3
>
> I'm usually attracted to posts that make convincing arguments but are misleading or wrong, from my perspective. P7

Advancing personal interest. Some interviewees also found r/science advantageous for their own work. One user (P4) admitted to using r/science primarily to promote their own work: *It's the one I'm most interested in because it yields the ROI I'm most looking for.* Another (C3), a science communicator, mentioned that science debates on r/science helped them to refine some of their rhetorical strategies. This was a less prominent theme than the others we observed, but it is a reminder that altruism and love of science alone do not sustain this community.

What types of science are being shared on r/science?

r/science posts can link directly to academic papers, or to reputable news sources about them. We wanted to know what sources users posted from, and whether they showed any indication of secondary gatekeeping activities. Our log-data analysis showed that the majority of posts linked directly to popular science publications (4804 / 40.5%) or to general news articles (1673 / 14.1%). Direct links to research papers, including abstracts, comprised 24.3% (2881) of posts. In addition to these, there were a total of 1778 press releases, and 338 Reddit-internal posts, of which 321 were AMAs.

As of 2016, all posts made to r/science, including AMAs, must have topic flair. This affords us an opportunity to analyze the topic balance that arises on a large community-driven science forum. However, the flairs that are added are highly variable in terms of specificity, making direct comparison by flair to be problematic (e.g., Biology vs. Alzheimer's disease). To remedy this, we condensed all topic flairs into 16 high-level topics. The topics that were most commonly posted and discussed were Social Science (1,724 / 14.5%), Biology (1,558/ 13.1%), and Health and Epidemiology (1,505 / 12.7%). Table 3 lists all 16 high-level topics discussed on r/science ordered from most to least commonly posted.

Problems with Newsworthiness. While our quantitative analyses indicate that science content from many fields are being posted to r/science, our results also show that as with online content in general [23], a small subset of the posts tend to attract the majority of users' attention.

Reddit is a *social* news site. This means that the items that are surfaced are those that are upvoted and discussed most

[2] deliberate negative or troll-like traffic coordinated from one subreddit to another, often in response to a specific post.

Table 3: High-level post topics ordered by frequency.

Topic	Post count	% posts
Social Sciences	1,724	14.5%
Biology	1,558	13.1%
Health and Epidemiology	1,505	12.7%
Medicine, Anatomy, and Disease	1,328	11.2%
Ecology, Climate, and Sustainability	1,085	9.2%
Paleontology and Animal Science	996	8.4%
Brain and Cognition	954	8.0%
Weather, Oceans, and Geology	688	5.8%
Physics	688	4.5%
Technology, Eng., and Mat. Science	498	4.2%
Astronomy, Astrophysics, and Space	476	4.0%
Chemistry	287	2.4%
Computer Science	97	0.8%
General Science Discussion	37	0.3%
Mathematics	35	0.3%
Other	25	0.2%
Genetics	18	0.2%
No topic flair provided	12	0.1%

by participants. In 2016, the mean number of comments on a post was 60.1 (sd=260); the median was 3. With the 90th percentile at 83 and the max at 6162, there is significant disparity in the level of engagement different posts receive. The topics that are perceived as most exciting and stimulating naturally float to the top. While it is possible to sort posts in several different ways, less popular posts can quickly be buried unless a user deliberately does this. Several of our interviewees mentioned that this can contribute to the problem that only posts from certain, popular or newsworthy topics, get read and others get overlooked:

> I think that sometimes there are good things that don't get rated up because they're not, unfortunately, click baiting enough. There are some topics that keep coming up [that] I think are hot ... but I think I wish that some of the more obscure but interesting and worthwhile stuff got eyeballs, I guess, but you can't, not based on what people upvote. But I guess that's part of it too. Maybe a curated subset of, "Here are great papers we spotted in this topic today," whether that's health or quantum physics or whatever it is. That might help flag some of the less attractive, initially, things. P3

The problem is further exacerbated by posters who are motivated to attract more people to their posts. To maximize the votes on their posts, they learn over time through observations and trial and error on what topics work.

> Yeah, you have to know the hive mind, right, so the general lay person, especially if it's gonna be on the internet, it just goes to the lowest common denominator, so marijuana, sex, sexual orientations, I'm not kidding ... okay, the latest post that I had ... one of the latest posts was on light crystals, how scientists have actually managed to combine photons together to get solid light. P1

One moderator noted the issue of promoting less-sensationalized or obviously newsworthy research as the largest area for growth in r/science. They have used AMAs to right this balance, but ultimately it is the multitudes of voting readers who decide what is worth discovering:

> That's actually where we would have the biggest room for improvement. Because of Reddit: what is the top post on r/science is what has been voted. We have little control. Nearly every day it's in one or two dozen topic issues: gender, sex, drugs. That's because that's what Reddit upvotes. Two to three pages down, you'll find incredibly fascinating papers that are getting no traction or visibility. M1

A peripheral user [R4] mentioned that it might be hard for posts on r/science that refer to more abstract research topics to cater to the interests of the general public. Since each post stands on its own, it can be difficult to see why a highly specialized research publication is important or interesting in the context of scientific discovery.

> If the topic requires too many prerequisites, that's not something you can market to anyone. If someone's making breakthroughs in something really abstract like formal logic there's no way you could package that info in a way that the general public could appreciate it. You have to be able to explain the way it impacts people, have to have good writers to make it interesting; people like pictures and video. R4

Another interviewee suggested very similar interventions from the perspective of an expert who frequently faces misconceptions and errors in their interactions with less expert readers:

> The "scientific realism" people who say, "Here's some science to prove that black people have smaller brains" or whatever– like the Bell Jar. Right from the 90s. And I have to go track down the same four debunking articles for that time and time again. And so it would be nice to have a repository. You know I originally pitched it as a place where people could go and see for themselves. C3

How are Science Dialogues Facilitated?

Through our analyses of r/science and our interviews, we found that several forms of science discussions are occurring in the comments section. When we analyzed the primary codes for the comments, we noted that most frequent types of comments are related to information exchange (see Table 1 for code descriptions, counts, and mean scores and average size of comment reply tree. See Supplementary Materials for an expanded table with examples of each code). The most frequent types of comments are questions about the research or related work (Inquiry: 544). Next most frequent are linking to additional information (Info: 350), extending, applying, or reasoning about the research (Reasoning: 222), personal questions or stories or responses to such (Personal: 252), and offering an educational response (Educate: 205). Yet the most common types of comments were not always the highest rated. Though summaries of research (Summary) are highly valued (mean score=175), our dataset only contained four of them. Among more commonly occurring codes, it was questioning of research methods (Q-Methods) that performed best on average (m=94.8). These comments tend to be critical of methods, something that requires a certain level of scientific or statistical knowledge. Quotes (m=39) lower the barrier for readers to obtain information from the article, and was the third highest rated type of comment.

Score is only one measure of dialogue success. To see how these comments facilitate discussions, we examined the total number of replies each of our 2,056 comments received (its entire tree, not just those we coded). Here, we found Process, which more broadly discusses the process of science, received the highest number of replies. Q-Methods was second-highest. In contrast, Summary comments tend to be one-off posts, and attract an average of only one reply. Thus, highly valuable comments are often not the most discussed comments.

So far we have examined comments as more and less successful *components* or representations of dialogue. From our interviews, we learned that comments also help to form a secondary communication product themselves, one which *guides* interaction and further dialogue. There are two important themes that we extract from this. One is that the science communication is heavily community-driven, and that the comments are often more important than the articles themselves.

Community Driven. To facilitate high quality science discussions, a great deal of work goes into contextualizing information, correcting fallacious arguments or helping prevent them from happening, upvoting helpful comments, and other informal "traffic direction" activities to shape the conversation. While some of this work is done by moderators, much of the work depends on contributions from the community members. As one interviewee explains:

> Both the user, and the moderators, though theoretically it is the moderators that are there to ensure the conversation is scientifically accurate, and is conducted in a well-mannered/professional fashion. Realistically, however, given the low number of moderators to users, such responsibility being placed solely on the shoulders of the moderators is not feasible. Thus, it must be the responsibility of both the users and the moderators. P6

r/science relies on the knowledge of its community members to ensure a high quality of information. This includes work to correct erroneous claims, but also to help translate content to make the material more accessible to others.

> One of the best behaviors I see [on r/science] is someone who clearly has knowledge or is an expert in the field, and they help someone else out who clearly doesn't. Or maybe that the less knowledgeable person has made a comment that is off base or that they're just ignorant of the subject. Rather than putting them down, the knowledgeable person will offer true information to help them out in the spirit of helpfulness, not to be negative, but to get them more involved and knowledgeable. P5

Other interviewees described the role of the r/science community as bringing in diverse viewpoints and expert knowledge:

> I think what my position does is it gives me a perspective on applied science in a way that some of the more pure academic theoretical kinds of researchers don't see. And so I think that's a different window into science and possibly the way I find science topics and interesting stories. ... So when a vaccine is developed, I'm not as interested in the basic T-cell functions. I'm interested in, okay, now how is it gonna get out to the villages in Africa where it's gonna be needed? What does the whole chain look like? There's a bunch of stuff on that applied end of science that I think I bring to that in a different way. P3

> Because I have a strong background in the psychology, the sociology of the philosophy that many of the educators don't have.

> So I can contextualize in lots of ways that many of my colleagues probably couldn't. C3

Aside from content-related work to ensure high quality dialogues, our interviewees also describe other types of work on the site that require less expertise, and the general community members contribute to improve the quality of r/science. This includes reporting non-relevant content.

> It's like your mom jokes and then people will just key off that and they will go nine comments deep and they go... I'll go in, and I'll screen them, and I'll report them, because you want to get to the scientific questions or people who are genuinely interested, you don't want like ... There's the rest of Reddit for jokes if you really want, not on r/science. P3

Yet it can also include "negabehaviors" [46] such as refraining from commenting unless the comment is relevant and informative or in the person's area of expertise. One interviewee mentioned only contributing content if they were absolutely sure that their contribution would be meaningful:

> I think I'm just one of, you know, hundreds or thousands, or however many people are here. ... Like I said, I don't really contribute a whole lot. I rarely comment. Only if I really think that what I have to say is particularly meaningful. But ... So, no, not really. I'm more of a lurker than I am a participant in that sense. P2

Another participant said:

> I have verified flair. Which means that I've proven to r/science that I hold a PhD. One of the things that they like with people with verified flair to do is to try to limit their conversations mostly to their areas of expertise. C3

Tensions in supporting participation. The standard account of the dialogue model emphasizes bi-directional communication between two very distinct groups of people: active scientists producing research, and a lay public engaging with information about that research. However, by and large, although both our comment codes and our interviews reveal that participants highly value this form of engagement, it makes up a very small proportion of the dialogue that takes place on r/science — typically in the form of AMAs, which have since been discontinued. The labor involved in securing scientists for AMAs, and then managing the conversations, was significant, as one of the moderators who plays a key role in setting them up observed:

> What makes it engaging for the scientists for AMAs is the size of the audience. There's no talk they're ever going to give that's bigger than that. Scientists are inherently curious, engaged people who LOVE to talk about their own research. ... The biggest fear on the scientist side is that things are going to go off the rails and that their message is going to get totally lost. What we offer them is protection from all of that. M1

However, the goal of facilitating broad public dialogue mentioned by the moderators and participants is tested by dealing with the public when they actually do participate in discussions. As described by both moderators and active participants, a post becoming popular enough to reach "r/all" (a user's home page, showing a digest of the most popular posts from their subscriptions) brings its own problems: large proportions of comments are often deleted, and significant moderator labor goes into crowd control, as our interviewees acknowledged:

> Now, one of things that I've noticed in some of the posts that I've been looking at, is when a post does get to r/all and it attracts a lot of attention and that means a lot of comments get deleted there's always someone who's like "Wow, I noticed that three quarters of the comments have been deleted, what's up with that?" P1

A moderator (M1) further discussed the difficulties of reaching specific audiences due to being embedded in the larger Reddit community:

> We'll also get brigading from hate speech subs when we have a topic that touches a nerve with that group. I think about the young, energized, science-interested kid coming to our sub and the first thing they see is something derogatory about their identity group. By being couched in this larger website, where so much bigotry and outright hatred is allowed to grow and flourish, we're losing so much access to the very audience we really want to touch.

In order to maintain a high standard of rigor and discourse, the moderators have come to enforce stricter rules about the sources of posts and the wording of titles:

> Our rules have developed over years. ... We don't want to decide on what science is; anything peer reviewed is fine. But that rule we had to change. People would go

> find some obscure little journals and find some flamebait. People were gaming that. And then there's predatory journals. So we had to adjust that rule to say you had to have an impact factor of 1.5 or greater. M3

Another moderator talked about the struggle of being rigorous while not wanting to discourage participation:

> Since I've joined the changes have been pretty subtle; we've changed the rules about titles so that they have to a have a model organism if relevant. That still gets a little hairy because any psych studies are done in undergrad research populations because of access to participants. So there's some debate, if it was done in male undergrad population–should we be requiring them to say "in undergrad male populations"? There's always a balance between making it strict enough that it's not misleading, but not so strict that users can't engage well. M1

Post title discipline ensures that claims made are presented in a measured way, even to a casual reader. However, as several of our interviewees noted, a total lack of "clickbait" may dissuade less motivated participants to engage at all, and sensationalism could serve as a bridge to reach a broader readership.

The structure of Reddit itself can also inhibit participation. One moderator mentioned that Reddit's content surfacing algorithm was causing declining viewership, making it more difficult to establish a value proposition for busy researchers and university PR people to set up AMAs, which in the months following our interviews were shut down on r/science:

> Now our content is less likely to reach the front page. Now our partners in science discussion can no longer engage with us as often, because they were paying people to talk to us and it may no longer be worth it. M2

r/science is subject to structural forces beyond its control, and this can also have a major impact on the flow of readership and engagement into the community.

Comments First. Another key theme that emerged from our interviews is that readers often review the comments before engaging with the articles. As noted in the previous section, summaries and explanations are often provided in comments, as well as comments that provide or link to other relevant information. For example, one participant cited the ability of other commenters to help them think in new ways:

> Sometimes people raise very good points of view or new ways of thinking about something. C2

P7 additionally mentioned that *"the Reddit format is very conducive for non rigorous scientific debate. Articles get debated and the science checked multiple times by multiple people with different perspectives."*

Reading comments first serves two key purposes. The first purpose, given the tremendous amount of content of Reddit, is that the comments help users decide what articles to spend time on. This cost evaluation made by readers may be supported in three ways by the comments: One, as C5 told us, is that the total number of comments is used to signal general interestingness of the topic:

> The other thing I'm looking for, once it's been posted and I'm looking at comments to see if I want to engage is, how engaged are the other people and how intelligent is this, the discussion.

Two, the summaries in comments can help readers decide if the articles are of interest and worth reading. As M2 noted:

> Comments are important for the takeaway. A lot of people will never click on the article. A lot of people will often (including me), go to a thread, read the comments first, then based on the comments will decide whether the article is worth reading.

Three, comments that are highly critical and that point out numerous flaws in the article might lead readers not to invest the time in reading:

> Sometimes I will read the comments first, especially if the headline is dubious and often the top comment looks like someone who knows what they're talking about and they've cited the scientific literature saying, "Now here's why I don't buy this or here's a flaw in the study." And that can be really useful, like I said, especially stuff that's outside of my area of expertise. C4

The second purpose of reading comments first is that it facilitates reading of the articles. For example, one interviewee who self-described as a non-expert reader told us that comments provide additional context that can situate the work:

> And it's, I don't want to necessarily say bad science, but if it's not on the up-and-up, or if there are studies that contradict it, and a lot of times, you find some really good information in the comment section

> in regards to that, and I kind of like to have that information going in. P2

Similarly, another participant mentioned the importance of having experts who provide summaries and additional information in the comments:

> Generally to me the most valuable are those [comments] that maybe take a more complex article and bring it to a level I can understand better. I think other people in the community appreciate that, someone who is very knowledgeable, maybe an expert in that field, says, "Well, this is what it really means. Here's the TLDR," or, "Here's the summary in layman's speak." P5

Our participants offered many perspectives on how and why they engage on r/science, and what effects the contributions have on them. In the next section, we will connect these insights to higher-level themes in science communication and discuss opportunities for design.

5 DISCUSSION

Much analysis emphasizes the negative effects of social networks, such as bias, echo chambers, or fake news (e.g., [29, 49]), or attempts to port some of the same frameworks from traditional broadcast media. Our analysis of r/science offers a different viewpoint, one with unique challenges as well as new opportunities for contemporary science communication. The content created by this collaborative process of posting, stewardship, moderation, and discussion adds to the original research or news article being posted. Our study also uncovers how "dialogue" can be defined in online science communication, as it identifies a wider range of dialogue participants who are rarely the authors of the research themselves.

Community-Driven Science Dialogues

Much prior work has proposed dialogue as a goal for science communication. However, although many case studies exist, exactly who the parties of this dialogue are and what actions they perform is theoretically underspecified. As more social interaction happens online and new mechanisms for dialogue arise, the need to delineate these roles and activities grows.

Central to community-driven online science dialogues are the comments, which serve multiple purposes. First, they enable valuable dialogue *between* readers. The most frequent types of comments on r/science are on information exchange—questions and answers about research or related work. Second, comments enable broader participation, soliciting peripheral work from the lay public who votes, reports, and even refrains from contributing irrelevant and low quality content. Third, and perhaps most important, these comments become community-sourced and community-centered science communication documents. As our interviews showed, comments not only serve to help explain the science to a broader audience, but they can also heavily influence whether and how subsequent readers engage with and frame the content; the comments section itself becomes the primary artifact that communicates the science, instead of the linked articles. This finding extends recent research that has found that uncivil comments can change people's interpretation of the news story [1] and that most redditors do not actually click through to view the content being rated [19]. Further, our investigation of posts, comments, and interviews demonstrates the importance of *information stewardship*. This heterogeneous collection of citizen journalist-style activities–bringing specialized sources to other readers, guiding the comments section, adding credential-oriented metadata, and imposing careful restrictions post style–allows a relatively large community to produce a living science communication document. "Dialogue," then, functions as both a process and a product.

Our findings add to the ongoing debate on how public comment sections should be used for online science communication [57].There are two important points to note. First, our findings of the value of comments in the r/science community suggests that the commenting section can facilitate the exchange of information. While the cost is high—r/science requires much information stewardship from moderators and the larger community—the comments offer both direct and indirect benefits to the readers. Second, the comments-first approach to how people engage with posts can have an amplifying effect on science communication. A risk is that if a community lacks the expertise, resources, or norms to effectively moderate their posts, the comments can greatly undermine public's engagement with science. Low quality comments not only turn people off from engaging with the science content, but could also negatively bias their interpretation lead them to distrust good science [1, 57]. Yet when effectively moderated, the comments can help translate and complement the valuable insights being discussed in the science articles.

Who is Missing from the Dialogue?

Our study showed that dedicated members of the community contribute a great deal of content and help to maintain community standards. We found that they use r/science to learn, engage in thought-provoking discussion, gain access to working scientists, and help others to do the same. Most active participation in r/science tends to occur among science-interested and science-educated people who are neither the

primary researchers nor a lay "general public" with no specific science training. Burns, O'Connor and Stocklmayer[7] acknowledge with their notion of varied publics that not everyone will have the same level of interest or knowledge about science, but everyone should be able to get something useful out of science communication. They must, however, be present for it.

On one hand, r/science must work on increasing participation from the scientists who are directly involved in the research. Both active r/science community members and the general Reddit participants we spoke to found a great deal of value in the opportunity to have an audience with research producers. Yet this form of dialogue is not a regular occurrence on r/science outside of the now-discontinued AMAs, which required a significant amount of hands-on work with researchers to set up and conduct. We believe there are opportunities to design lighter-weight opportunities to engage the researchers. For example, PIs could be automatically notified and invited to participate when their papers are posted and discussed. To lower researchers' anxiety in participating in such public discussions about their own work, r/science could additionally provide a "quick guide" to Reddit and r/science norms and practices, modeled on prior AMAs. If researchers are encouraged to identify themselves in a recognizable way, a bot could ping the moderators to offer welcome and assistance. This would provide some of the bespoke AMA support in a more automated, as-needed fashion.

Another side of the participation problem is the lay public. For this group of participants, the challenge is in balancing accessibility of the scientific content with rigor. Science communicators must facilitate access to scientific knowledge and process that is impactful and stimulating, but not inaccessible or alienating: an issue the members of r/science grapple with in their daily engagement with the community. For example, our interviewees mentioned that the interactivity or casual nature of some of the sister subreddits (which have a smaller audience) are more appealing and that r/science's rigorous rules lead to frustratingly high barriers to contributing posts or comments. Other online communities, such as StackOverflow, have encountered similar issues where many users refrain from actively contributing if they perceive themselves as non-experts or perceive norms and posting rules as difficult to anticipate and understand [35]. One successful solution to involve such users has been to provide an onboarding process where novice users get assigned a mentor [14]. r/science could offer a similar mentorship program, pairing members of the lay public with users who have previously received author flair. Such mentorships could encourage the lay public to get more involved, grow the community over time, and also keep up the rigor that the r/science community has been trying to maintain. Such relationships could also strengthen the motivation of more senior members who enjoy providing guidance and expertise–further tapping this valuable resource.

From Aggregation to Integration

Though the members of r/science do not rely solely on traditional science gatekeepers, the community still suffers from some of the same pipeline issues that cause most scientific work to go undiscovered by the general public (as in, e.g., [3]). This is caused in part by Reddit's information architecture: less upvoted stories quickly get buried, and the default order of posts is vote-based. Since we found that the majority of posts include links to news sources (both general and popular science), most of the material readers are being exposed to has *already* been judged newsworthy. Title discipline may temper this disparity but, as previously discussed, causes its own problems.

To provide more visibility valuable but overlooked posts, the community could conduct regular reviews of less-amplified research. Moderators could then "pin" a digest of post recommendations ("great posts you may have missed") to the top of the page. This would allow the primary news aggregator functionality of r/science to continue, but to counter some of its ephemerality.

Another, complementary approach might be a move from simple *aggregation* (in the sense of a news aggregator that lists news items) to *integration*. On r/science, each post tends to stand alone. Without the benefit of disciplinary knowledge, it can be challenging for an average reader to see why they should care and what each study contributes to the broader arc of scientific research. As such, topical digests that connect many different posts (especially in unexpected ways) could help general readers understand how small, incremental studies add to the growth of human knowledge. Such digests could also, as our interview participants suggested, offer tools for dedicated information stewards to enhance their capacities: they could turn to these digests of research or argumentation to help steer the conversation in a productive way.

6 CONCLUSION

This paper presented a mixed-methods study of r/science, contributing insights from a science news and discussion forum that has changed the way science communication has traditionally been envisioned. Our findings elucidate who the participants are and what types of work they perform in their community-driven approach to science communication. We also uncovered the central and multifaceted role of comments in this community—not only to support interaction and discussion between community members, but also help guide readers' engagement with scientific research. Our

work reveals the large potential of online science communication for involving broader audiences in scientific research, and contributes improvement opportunities that we hope will foster effective science communication on r/science and beyond.

REFERENCES

[1] Ashley A Anderson, Dominique Brossard, Dietram A Scheufele, and Michael A Xenos. 2012. Online talk: How exposure to disagreement in online comments affects beliefs in the promise of controversial science. *Citizen voices: Performing public participation in science and environment communication* (2012), 119–135.

[2] Judd Antin, Raymond Yee, Coye Cheshire, and Oded Nov. 2011. Gender differences in Wikipedia editing. In *Proceedings of the 7th international symposium on wikis and open collaboration*. ACM, 11–14.

[3] Christopher Bartlett, Jonathan Sterne, and Matthias Egger. 2002. What is newsworthy? Longitudinal study of the reporting of medical research in two British newspapers. *BMJ* 325, 7355 (2002), 81–84. DOI: http://dx.doi.org/10.1136/bmj.325.7355.81

[4] John C. Besley, Anthony D. Dudo, Shupei Yuan, and Niveen Abi Ghannam. 2016. Qualitative Interviews With Science Communication Trainers About Communication Objectives and Goals. *Science Communication* 38, 3 (2016), 356–381. DOI: http://dx.doi.org/10.1177/1075547016645640

[5] Holly M. Bik and Miriam C. Goldstein. 2013. An Introduction to Social Media for Scientists. *PLOS Biology* 11, 4 (04 2013), 1–8. DOI: http://dx.doi.org/10.1371/journal.pbio.1001535

[6] Lutz Bornmann and Rüdiger Mutz. 2014. Growth rates of modern science: A bibliometric analysis. *CoRR* abs/1402.4578 (2014). http://arxiv.org/abs/1402.4578

[7] T. W. Burns, D. J. O'Connor, and S. M. Stocklmayer. 2003. Science Communication: A Contemporary Definition. *Public Understanding of Science* 12, 2 (2003), 183–202. DOI: http://dx.doi.org/10.1177/09636625030122004

[8] Jason Chilvers. 2010. *Sustainable participation? Mapping out and reflecting on the field of public dialogue on science and technology*. Technical Report. University of East Anglia, Norwich, UK.

[9] PMD Collins and WF Bodmer. 1986. *The public understanding of science*. Technical Report. The Royal Institution, London, UK.

[10] Bethan J Davies and Neil F Glasser. 2014. Analysis of www. AntarcticGlaciers. org as a tool for online science communication. *Journal of Glaciology* 60, 220 (2014), 399–406.

[11] Thomas Dietz. 2013. Bringing values and deliberation to science communication. *Proceedings of the National Academy of Sciences* 110, Supplement 3 (2013), 14081–14087. DOI: http://dx.doi.org/10.1073/pnas.1212740110

[12] Aaron Drummond, Matthew A. Palmer, and James D. Sauer. 2016. Enhancing endorsement of scientific inquiry increases support for pro-environment policies. *Royal Society Open Science* 3, 9 (2016). DOI: http://dx.doi.org/10.1098/rsos.160360

[13] Elaine Howard Ecklund, Sarah A James, and Anne E Lincoln. 2012. How academic biologists and physicists view science outreach. *PloS one* 7, 5 (2012), e36240.

[14] Denae Ford, Kristina Lustig, Jeremy Banks, and Chris Parnin. 2018. We Don't Do That Here: How Collaborative Editing with Mentors Improves Engagement in Social Q&A Communities. In *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems*. ACM, 608.

[15] Bev France, Belinda Cridge, and Laura Fogg-Rogers. 2017. Organisational culture and its role in developing a sustainable science communication platform. *International Journal of Science Education, Part B* 7, 2 (2017), 146–160. DOI: http://dx.doi.org/10.1080/21548455.2015.1106025

[16] Adabriand Furtado, Nazareno Andrade, Nigini Oliveira, and Francisco Brasileiro. 2013. Contributor profiles, their dynamics, and their importance in five Q&A sites. In *Proceedings of the 2013 ACM Conference on Computer Supported Cooperative Work*. ACM, 1237–1252.

[17] Devin Gaffney and J Nathan Matias. 2018. Caveat Emptor, Computational Social Science: Large-Scale Missing Data in a Widely-Published Reddit Corpus. *arXiv preprint arXiv:1803.05046* (2018).

[18] Barney G Glaser and Anselm L Strauss. 2017. *Discovery of grounded theory: Strategies for qualitative research*. Routledge.

[19] Maria Glenski, Corey Pennycuff, and Tim Weninger. 2017. Consumers and curators: Browsing and voting patterns on Reddit. *IEEE Transactions on Computational Social Systems* 4, 4 (2017), 196–206.

[20] Luke Goode. 2009. Social news, citizen journalism and democracy. *New Media & Society* 11, 8 (2009), 1287–1305. DOI: http://dx.doi.org/10.1177/1461444809341393

[21] David R. Gruber. 2017. Three Forms of Neurorealism: Explaining the Persistence of the "Uncritically Real" in Popular Neuroscience News. *Written Communication* 34, 2 (2017), 189–223. DOI: http://dx.doi.org/10.1177/0741088317699899

[22] Lars Guenther and Georg Ruhrmann. 2016. Scientific evidence and mass media: Investigating the journalistic intention to represent scientific uncertainty. *Public Understanding of Science* 25, 8 (2016), 927–943. DOI: http://dx.doi.org/10.1177/0963662515625479 PMID: 26802103.

[23] Steven Johnson, Samer Faraj, and Sri Kudaravalli. 2014. Emergence of Power Laws in Online Communities: The Role of Social Mechanisms and Preferential Attachment. *Management Information Systems Quarterly* 38, 3 (09 2014), 795–808.

[24] Dan M. Kahan. 2017. 'Ordinary science intelligence': a science-comprehension measure for study of risk and science communication, with notes on evolution and climate change. *Journal of Risk Research* 20, 8 (2017), 995–1016. DOI: http://dx.doi.org/10.1080/13669877.2016.1148067

[25] Aniket Kittur, Ed Chi, Bryan A Pendleton, Bongwon Suh, and Todd Mytkowicz. 2007. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. *World wide web* 1, 2 (2007), 19.

[26] Janice L. Krieger and Cindy Gallois. 2017. Translating Science: Using the Science of Language to Explicate the Language of Science. *Journal of Language and Social Psychology* 36, 1 (2017), 3–13. DOI: http://dx.doi.org/10.1177/0261927X16663256

[27] Alex Leavitt and Joshua A Clark. 2014. Upvoting hurricane Sandy: event-based news production processes on a social news site. In *Proceedings of the 2014 ACM Conference on Human Factors in Computing Systems*. ACM, 1495–1504.

[28] Stephan Lewandowsky and Klaus Oberauer. 2016. Motivated Rejection of Science. *Current Directions in Psychological Science* 25, 4 (2016), 217–222. DOI: http://dx.doi.org/10.1177/0963721416654436

[29] Q. Vera Liao, Wai-Tat Fu, and Markus Strohmaier. 2016. #Snowden: Understanding Biases Introduced by Behavioral Differences of Opinion Groups on Social Media. In *Proceedings of the 2016 ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 3352–3363. DOI: http://dx.doi.org/10.1145/2858036.2858422

[30] Susana Martinez-Conde and Stephen L. Macknik. 2017. Opinion: Finding the plot in science storytelling in hopes of enhancing science communication. *Proceedings of the National Academy of Sciences* 114, 31 (2017), 8127–8129. DOI: http://dx.doi.org/10.1073/pnas.1711790114

[31] Ashley Rose Mehlenbacher. 2017. Crowdfunding Science: Exigencies and Strategies in an Emerging Genre of Science Communication. *Technical Communication Quarterly* 26, 2 (2017), 127–144. DOI: http://dx.doi.org/10.1080/10572252.2017.1287361

[32] Steve Miller. 2001. Public understanding of science at the crossroads. *Public Understanding of Science* 10, 1 (2001), 115–120. DOI: http://dx.

doi.org/10.3109/a036859

[33] Blair Nonnecke, Dorine Andrews, and Jenny Preece. 2006. Non-public and public online community participation: Needs, attitudes and behavior. *Electronic Commerce Research* 6, 1 (2006), 7–20.

[34] Nigini Oliveira, Nazareno Andrade, and Katharina Reinecke. 2016. Participation differences in Q&A sites across countries: opportunities for cultural adaptation. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. ACM, 6.

[35] Nigini Oliveira, Michael Muller, Nazareno Andrade, and Katharina Reinecke. 2018. The Exchange in StackExchange: Divergences between Stack Overflow and its Culturally Diverse Participants. In *Proceedings of ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*. ACM.

[36] Katherine Panciera, Aaron Halfaker, and Loren Terveen. 2009. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*. ACM, 51–60.

[37] Vineet Pandey, Amnon Amir, Justine Debelius, Embriette R Hyde, Tomasz Kosciolek, Rob Knight, and Scott Klemmer. 2017. Gut instinct: Creating scientific theories with online learners. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 6825–6836.

[38] Vineet Pandey, Justine Debelius, Embriette R Hyde, Tomasz Kosciolek, Rob Knight, and Scott Klemmer. 2018. Docent: transforming personal intuitions to scientific hypotheses through content learning and process training. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*. ACM, 9.

[39] Pontus Plavén-Sigray, Granville James Matheson, Björn Christian Schiffler, and William Hedley Thompson. 2017. The Readability Of Scientific Texts Is Decreasing Over Time. *bioRxiv* (2017). DOI: http://dx.doi.org/10.1101/119370

[40] Jason Priem and Kaitlin Light Costello. 2010. How and why scholars cite on Twitter. *Proceedings of the American Society for Information Science and Technology* 47, 1 (2010), 1–4. DOI: http://dx.doi.org/10.1002/meet.14504701201

[41] M Jordan Raddick, Georgia Bracey, Pamela L Gay, Chris J Lintott, Carie Cardamone, Phil Murray, Kevin Schawinski, Alexander S Szalay, and Jan Vandenberg. 2013. Galaxy Zoo: Motivations of citizen scientists. *arXiv preprint arXiv:1303.6886* (2013).

[42] Lee Rainie, Cary Funk, and Monica Anderson. 2016. How Scientists Engage the Public. (2016). http://www.pewinternet.org/2015/02/15/how-scientists-engage-public/

[43] Mathieu Ranger and Karen Bultitude. 2016. 'The kind of mildly curious sort of science interested person like me': Science bloggers' practices relating to audience recruitment. *Public Understanding of Science* 25, 3 (2016), 361–378. DOI: http://dx.doi.org/10.1177/0963662514555054 PMID: 25361791.

[44] reddit.com. 2018. Reddit API Documentation. (2018). https://www.reddit.com/dev/api/

[45] Laura N. Rickard, Jonathon P. Schuldt, Gina M. Eosco, Clifford W. Scherer, and Ricardo A. Daziano. 2017. The Proof is in the Picture: The Influence of Imagery and Experience in Perceptions of Hurricane Messaging. *Weather, Climate, and Society* 9, 3 (2017), 471–485. DOI: http://dx.doi.org/10.1175/WCAS-D-16-0048.1

[46] Joel Ross and Bill Tomlinson. 2011. Negabehaviors and environmental sustainability. *Journal of Sustainability Education* 2 (2011), 2011.

[47] Pamela J. Shoemaker, Director Martin Eichholz, Eunyi Kim, and Brenda Wrigley. 2001. Individual and Routine Forces in Gatekeeping. *Journalism & Mass Communication Quarterly* 78, 2 (2001), 233–246. DOI: http://dx.doi.org/10.1177/107769900107800202

[48] Jane B Singer. 2014. User-generated visibility: Secondary gatekeeping in a shared media space. *New Media & Society* 16, 1 (2014), 55–73. DOI: http://dx.doi.org/10.1177/1461444813477833

[49] Kate Starbird, Jim Maddock, Mania Orand, Peg Achterman, and Robert M Mason. 2014. Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing. *IConference 2014 Proceedings* (2014).

[50] Marzena Świgoń. 2017. Knowledge sharing practices in informal scholarly communication amongst academics in Poland. *Malaysian Journal of Library & Information Science* 22, 2 (2017), 101–115.

[51] Masamichi Takahashi, Masakazu Fujimoto, and Nobuhiro Yamasaki. 2003. The active lurker: influence of an in-house online community on its outside environment. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*. ACM, 1–10.

[52] Phillip Tracy. 2018. How Reddit killed one of its most popular AMAs. (2018). https://www.dailydot.com/debug/reddit-r-science-ama/, last accessed: 09/11/2018.

[53] Debbie Treise and Michael F Weigold. 2002. Advancing science communication: A survey of science communicators. *Science Communication* 23, 3 (2002), 310–322.

[54] Rajan Vaish, Shirish Goyal, Amin Saberi, and Sharad Goel. 2018. Creating Crowdsourced Research Talks at Scale. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1–11.

[55] Dustin J. Welbourne and Will J. Grant. 2016. Science communication on YouTube: Factors that affect channel and video popularity. *Public Understanding of Science* 25, 6 (2016), 706–718. DOI: http://dx.doi.org/10.1177/0963662515572068 PMID: 25698225.

[56] Brooke Foucault Welles and Isabel Meirelles. 2015. Visualizing Computational Social Science: The Multiple Lives of a Complex Image. *Science Communication* 37, 1 (2015), 34–58. DOI: http://dx.doi.org/10.1177/1075547014556540

[57] Stephan Winter and Nicole C Krämer. 2016. Who's right: The author or the audience? Effects of user comments and ratings on the perception of online science articles. *Communications–The European Journal of Communication Research* 41, 3 (2016), 339–360. DOI: http://dx.doi.org/10.1515/commun-2016-0008