How Features of a Civic Design Competition Influence the Collective Understanding of a Problem

BRIAN MCINNIS, Cornell University, USA
XIAOTONG (TONE) XU, University of California, San Diego, USA
STEVEN P. DOW, University of California, San Diego, USA

From Fortune 500 companies to local communities, organizations often strive to build a shared understanding about complex problems. Design competitions provide a compelling approach to create incentives and infrastructure for gathering insights about a problem-space. In this paper, we present an analysis of a two-month civic design competition focused on transportation challenges in a major US city. We examine how the event structure, discussion platform, and participant interactions affected how a community collectively discussed design constraints and proposals. Ninety-two participants took part in the competition's online discussion, hosted on Slack. Applying a mixed-methods analysis, we found that participants shared less as they settled into teams and, due to the discussion system, had difficulty seeing how topics connected across channels; we also learned that certain messages led participants to add depth to existing topics. Based on the findings we provide recommendations for civic competitions aimed at building knowledge around a problem.

CCS Concepts: • **Human-centered computing → Field studies**; **Empirical studies in HCI**;

Additional Key Words and Phrases: Collective intelligence, civic design, competition, deliberation

ACM Reference Format:
Brian McInnis, Xiaotong (Tone) Xu, and Steven P. Dow. 2018. How Features of a Civic Design Competition Influence the Collective Understanding of a Problem. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 120 (November 2018), 25 pages. https://doi.org/10.1145/3274389

1 INTRODUCTION

An important topic in organizational science is how to coordinate and communicate shared understanding about different facets of a problem [51, 71]. One mechanism that organizations use to rapidly generate solutions to specific problems is team-based design competition. While the winning solution typically responds to a narrow set of requirements to address a specific problem, the conversations that led up to a final design can capture how teams develop a shared understanding about a problem-space. Systems that coordinate team-based competitions can reveal how teams examine and talk about a problem-space.

Malone and Crowston [1990, pg. 24] wrote about how such coordination systems provide incentive mechanisms that encourage people to share their understanding of the problem so "[...] that many people could *swarm* around, filling in things they know, until a complete picture of what to do emerges." For example, the *ForumReader* system [14] developed for IBM's Innovation Jams [3] provides participants with a way to navigate the hundreds of thousands of comments in massive online discussions via visual search and automatic topic extraction tools. A useful example from

Authors' addresses: Brian McInnis, Cornell University, USA, bjm277@cornell.edu; Xiaotong (Tone) Xu, University of California, San Diego, USA, xt@ucsd.edu; Steven P. Dow, University of California, San Diego, USA, spdow@ucsd.edu.

2573-0142/2018/11-ART120 $15.00
https://doi.org/10.1145/3274389

online policy deliberation is *Deliberatorium* [33], where participants contribute specific types of information about a policy (e.g., ideas, supporting evidence, counter-arguments) so that the system can assemble participant contributions into a hierarchical argument map (see examples *gIBIS* [13] and *Cohere* [61]).

Many social and technical design factors alter how people talk about and explore a problem space through an online discussion forum. Competitions encourage in-group interactions, but discourage sharing across groups [27]. Discussion can influence how people talk with each other by adding message reply and thread features [61, 72] and hash-tags [23]. Additionally, some types of messages posted to a discussion system can invite more [1, 29, 65] and deeper discussion than others [49, 66]. Our research applies a mixed-methods approach to explore how aspects of the competition structure, discussion system, and participant interactions factor into how a local community talks about a problem-space.

In this paper, we examine the online discussions of a two month (62-day) civic design competition that we organized within a large metropolitan United States city. The competition was designed to generate prototype solutions to major transportation challenges, such as biking safety and how to prepare the city for autonomous vehicles. In total, twenty-three teams formed and 92 people took part in the online discussions, hosted on a Slack work-space.[1] Here we focus on 476 messages posted to four public channels intended for general discussion about the civic issues. The case study documents how the competition structure, discussion system, and participant interactions influenced the breadth and depth of discussion about the transportation problem-space.

To understand how deeply participants built on existing topics, we measured thread response rate and topic coherence [52, 66]. We found that 28% of the main channel messages led to a thread response, though the level of activity and the types of prior contributions affected the likelihood of a response. The implication is a missed opportunity for depth when the topics are not well timed or crafted. Not surprisingly new topics emerged through participant interaction; however, we discuss how the pattern of new topics linked earlier topics together in new ways.

We discuss the outcome of the competition as more than prototype solutions, but as an opportunity to gain a broader and deeper view into a problem-space. Design competitions with discussion forums provide the motivation and infrastructure for cities to talk about wicked problems; more effective socio-technical system design can help organizations collectively learn about a problem-space through conversation.

2 RELATED WORK

Systems have been designed to generate ideas at scale [3, 40], but dealing with real communities and real problems requires focus on a range of cultural, social, and technical challenges as well as the motivators that move people to participate [57]. In crisis response situations, the drive to recover and reconnect through community building can be a powerful impetus [44, 69, 74]. To address civic challenges in less dire circumstances, many cities have turned to design competition to motivate civic action [25, 27, 73].

In this paper, we consider the role that online systems play in supporting the collective understanding of a civic issue that develops during a civic design competition. Online systems can reveal the social processes that individuals, groups and teams adopt to address a problem [67]. Online medical forums provide a useful example, as the thread of responses to a simple medical question can surface deep dialogue about the strategies people use to ease their recovery, but unearthing this knowledge from a discussion is not easy [49]. In large online collaborations, like Wikipedia, systems also ease the process of collaborating around knowledge work with features to assign

[1] Slack is a popular online team communication platform that operates on multiple devices

tasks and flag problems [31]. In a similar way, the online social activity related to a civic design competition might capture how a series of design teams develop a collective understanding of a civic problem-space.

However, the effectiveness of an online discussion system to capture the collective understanding of a problem-space can be significantly affected by the organizer's preconceived discussion topics and by interaction between participants. Existing HCI research demonstrates many ways to represent open-ended discussion to support organizational learning—as key topics [14, 66], its social network [5, 28] or arguments [13, 33, 61]. Below, we review the research related to hosting design competitions, constructing a knowledge structure within a forum, and managing an online discussion. Each subsection concludes with a list of the design guidelines that were used to organize our civic design competition and online discussions.

2.1 Create a timeline for collaboration and competition

An important consideration for community problem-solving systems design is how to motivate participation. While crowds of people may be involved, the work is not paid crowd work [32, 53], but rather community work [17]. Regarding the design of online deliberation systems, Coleman and Gøtze [12, pg. 32] observed that "People do not simply go online to deliberate about policy; they go online because they are social animals who like relating to other people." For this reason, system designers not only need to think about how users navigate an information system [15, 68], but also how to design for the regular rhythm of social activities that sustain community problem-solving over time [12, 17].

Competition is an example of a social activity that can sustain community engagement over time [20]. While individual competitions discourage information disclosure, team-based competitions can foster a sense of collaboration among team members which can help to sustain individual participation [54]. However, team-based competition can be at odds with collaboration. To address the tension between team-based competition and collaboration, Hutter et al. [27] recommend generating a lot of different ideas, encouraging constructive criticism, and emphasizing that there are multiple ways to address a community challenge.

While developing a timeline for a civic design competition, an organizer should also consider how each event and deadline contributes to facilitating social activity around the civic issue. By "facilitating" we draw from the participatory design definition, "the art of moving people through processes to agreed-upon objectives in a manner that encourages participation, ownership and creativity from all" [62]. In recent years, participatory design research has shifted from traditional workplace studies of how workers engage with technology [62, 70] to how design might support community and commitment around a societal concern [19, 41]. Light and Akama [42] argue that with this shift in emphasis, more attention should be paid to the role that the facilitator plays in generating activity around a social concern, as their manner and approach affect the collaborative activity at different scales of the interaction—from the set up of the social environment, to the activity, to "the rhetoric of power, control, voice, history and culture [used throughout]" [43]. For this reason, constructing a timeline involves thinking about how social activity might develop as the participants and facilitators become familiar with each other over time.

Design Guidelines: (1) foster experiences that build connection among participants, (2) plan the timeline so that collaboration can occur before team competition sets in, (3) structure the competition so that there are several challenges with no clear answers.

2.2 Provide structure to the problem, but listen for new topics

Complex topics can lead to an incoherent discussion as new comments shift the focus of the discussion from one subtopic to the next [64, 66]. For this reason, online discussion organizers

often plan a discussion so major topics are separated into specific channels [4] or time periods [17, 60]. However, dividing a discussion topic into separate channels can create a technical barrier, preventing conversation at the intersection of major topics [4, 15, 68]. This problem is similar in face-to-face settings, where the participants in a discussion room may not know about the relevant conversation happening down the hall [76]. Existing HCI research has tested various visual systems [14, 66] and interfaces to assist users as they make sense of a discussion [33, 61].

Another approach is to stage a discussion to address specific topics for set periods. Many policy deliberation systems are modeled after *Robert's Rules of Order* [58, 68] and provide an online editable agenda so that participants can decide which topics to address, in what order, and when to move onto the next—see *e-Liberate* [60] and *OpenDCN* [17, 18] as examples. At some point in the discussion of a topic, it is often useful to stop and reflect on the existing positions or accumulated insight. The *RegulationRoom* platform [56] incorporates such a reflection stage, by inviting people who posted a comment on a public rule a few weeks earlier, back to respond to a draft summary of all comments. However, not all participants are adequately prepared to engage directly in discussion about a social problem [30, 59]; several systems integrate independent learning stages so that participants have time to study a problem before they take part in its debate [9, 17, 33].

As a discussion takes shape, the volume and breadth of topics can be challenging to navigate. HCI researchers have experimented with ways to overcome this challenge by aggregating participant contributions in novel ways, such as in pro-con lists [35, 36] or in a tree-like structure [13, 33, 61]. However, the systems tend to require participants to contribute specific types of information to a problem-solving process (e.g., issues, questions, evidence, positions of support or opposition) [39], which may detract from the social aspect of community problem-solving [12, 17].

Design Guidelines: (1) separate complex topics into separate channels or periods, (2) stage the discussion so that there is a regular rhythm of activity around the topics, (3) pay attention to how the topics shift over time (e.g., by polling participants about the most important topics, looking for opportunities to connect different discussions).

2.3 Craft comments that invite a response from newcomers

When new participants do make a contribution to the discussion, receiving a response increases their commitment to the discussion community [34]. However, the content of some messages can also affect whether other members are likely to reply [7, 29]. For example, posting on-topic messages, introducing oneself to others' in the discussion, asking questions, and being succinct tend to elicit a reply [1]. This presents an important concern for system designers, as people arrive at an online discussion with different levels of ability to contribute [30, 59].

Moderators play a pivotal role in welcoming people into a discussion and helping them to make their best contribution [21, 50]. The concept of newcomer socialization in online discussion communities describes a series of tactics that moderators might use to foster commitment among newcomers, by encouraging them to take part in the work of the community (e.g., raising questions, constructive critique) [10, 34, 55].

While receiving a response can boost a participants commitment to an online community, the coherence of a response to the topics affects how the discussion examines a problem-space [8, 64]. The discussion cannot deeply consider a problem-space when responses distract from, rather than add to, existing topics. Therefore, eliciting threads that build on the existing discussion is essential toward facilitating a constructive dialogue about a problem-space. In online health forums, collective sensemaking is obtained after back and forth negotiation of opinions and resolution of conflicts in coherent and deep discussion of a topic [49]. When people add comments that do not build on the existing topics, the result can be an incoherent discussion [52, 66].

Design Guidelines: (1) be at the door, introducing and welcoming new participants to the discussion, (2) recruit moderators and existing participants to help newcomers improve on their posts (e.g., restating a question, asking for more detail), (3) encourage responses that build on existing design topics.

3 DESIGNING A CIVIC DESIGN COMPETITION

Together with support from the city of San Diego, California and other regional partners, we created the Design for San Diego (D4SD) civic design competition focused around transportation challenges, such as making transit more accessible and preparing for autonomous vehicles.[2] As a large sprawling city with imperfect public transit systems, mobility provided an ideal topic for city-wide exploration. Pursuing this challenge also has the potential to give rise to innovative solutions that could be adopted as part of the city infrastructure or launched as a start-up. No matter how people choose to commute, they have relevant experience with the transportation issues and personal insights into potential solutions, regardless of their design background.

3.1 Event Organization

The D4SD competition events included promotion through social media, a kickoff event with the city's mayor, a hackathon, studio sessions, a judging period, and the awards ceremony at a local design summit.

3.1.1 Competition Timeline. We planned the 2-month competition timeline so that participants would have enough time to spend several weeks discussing the issues, forming teams and developing proposals in advance of a major design summit where the winners would be showcased. Following the design guidelines in section 2.1, we planned several in-person and online activities early in the timeline to connect participants, while pushing out the registration and submission deadlines so that collaboration might occur before team competition set in.

Day	Date	Event
0	Aug. 25th	Online discussion forum went live
27	Sept. 21st	First in-person gathering / kickoff event (mayor speaking)
28-29	Sept. 22nd - 23rd	Two-day hackathon
37	Oct. 1st	On-campus design course begins
47	Oct. 11th	Team registration deadline
58	Oct. 22nd	Project submission due
62	Oct. 26th	Finalist showcase and award ceremony at Design Summit

Table 1. A glance of key dates and events of our 62-day civic design competition

3.1.2 Recruitment. On day 27 (Sept 21st), we hosted a kickoff event in a downtown co-working space and invited the city mayor and municipal officials to give warm-up speeches. The goal of involving city officials at the kickoff was to increase recruitment and to motivate participants. Immediately following the kickoff, we hosted a two-day hackathon where several dozen people studied the civic topics and partnered with others eager to brainstorm and build prototypes. Thirteen concepts emerged, but only a few teams continued past the hackathon.

In addition, we offered a studio class at a local public research university to mentor participants who were less familiar with human-centered design. The class was offered twice per week to

[2]For more information about D4SD, please visit https://www.d4sd.org/

university students for course credit, but we encouraged participants from outside of the university to join and a few attended regularly. On day 37 (Oct 1st), the students joined the Slack workspace as a way of introducing them to the civic challenges and they were asked to contribute to the online discussions. The students added relevant newspaper articles, replied to some of the existing posts and shared stories about their personal struggles with transportation. Aside from this initial assignment, the students were instructed to select the communication tools that worked best for their team (e.g., Facebook messaging, text message, etc.).

Prior to the competition we worked with the university Institutional Review Board (IRB) to request a waiver for informed consent ahead of the research. Following a full IRB review, the waiver was granted on the grounds that: (a) requesting informed consent may be a barrier to public participation in the competition, and (b) it would not be possible for research purposes to separate the discussion contributions and influence of the participants who did not consent to participate in our research, but wanted to join in the competition.

3.2 Competition Structure

3.2.1 Topics. We chose to structure the competition on public concern about the mobility issues facing our city, specifically the following four broad topics: Accessibility, Autonomous Vehicles, Commuter UX, and Walking & Biking. Following the design guidelines in section 2.2, the complex topic of mobility was separated into sub-topics that involve a mix of immediate and long-term concern with no clear answers.

3.2.2 Teams. Inspired by the concept of co-operating and competing forces to team-based design competition, discussed by Hutter et al. [27], we allowed teams of 1-5 members and encouraged team formation at all the in-person events (e.g., Hackathon, design studio). To help participants find each other online, we created a Slack channel called *#team-formation*.

3.2.3 Incentives. Participants received several benefits from their participation. In addition to the opportunity to win a monetary reward for the best submission ($5,000), all teams who submitted a final proposal were provided with tickets to a community mixer at the design summit (a $50 value), and 8 finalist teams received a free ticket to the conference (a $150 value). The top teams were provided with an opportunity to have their idea supported and funded by a local innovation accelerator (an estimated $2,000 value). To select the top teams, we chose four judging criteria: novelty, feasibility, impact, and human-centered process.

In addition to the extrinsic rewards, there were several intrinsic motivations to participate. The competition provided a chance to work through real problems with a community of other interested designers, to learn human-centered design practices from professional designers and educators, and to network with other designers, engineers, entrepreneurs and civic leaders.

3.3 Discussion System

3.3.1 Socio-Technical Environment. We selected Slack as the official online discussion platform for several reasons. Slack is a virtual team environment that, like an online chat-room, enables synchronous chatting in both public and private channels. For the purpose of the competition, we created several public channels to host general discussions about the four major mobility challenges (e.g., accessibility, walking and biking) and three channels for process type questions and for forming teams. When teams formally registered for the competition, they were assigned a private team channel in the Slack environment. These decisions derive from the design guidelines to create separate public spaces for the general discussion of each topic (section 2.2), while providing space for teams to develop their winning ideas in private (section 2.1).

3.3.2 Reply-Threads. In addition to synchronous chatting, Slack also includes several features that support asynchronous discussion. When participants return to a channel, they are presented with a *readline* indicating the beginning of the new posts since their last visit. To engage with older posts without disrupting the current flow of the discussion, participants can initiate a reply-thread from any main channel post. Reply-threads are used to facilitate a side discussion associated with topics in a specific main channel post. When a reply-thread discussion becomes relevant to the current discussion, participants can post their replies back to the channel. As described in section 2.3, the design of reply mechanisms are critical to encouraging participant interaction.

3.3.3 Human and Bot Supported Discussion Facilitation. As described in section 2.3, by welcoming newcomers and being generally responsive to the discussion, moderators play a pivotal role in fostering community online [34]. However, managing large online discussions can take an entire team of people [21]. For that reason, we developed a Slack ChatOps bot (called *Helper Bot*) to engage with participants at a scale and speed that our human moderators could not match.

As an example, Helper Bot was used to initiate *Ideation Sprints*, which were 24hr online brainstorming efforts to address a single question related to a specific civic design topic, e.g., "*How might we detect that a traffic intersection is becoming dangerous?*" Eight Ideation Sprints were launched during the competition, two for each of the general channels. With Helper Bot we were able to send personalized direct messages to each participant. Similar direct messaging techniques have been used to gather synchronous crowds for real-time crowdsourcing [2]. We used this feature to generate activity around an ideation sprint. In practice, we worked with Helper Bot to elicit a response from participants, so that human moderators could follow up later.

4 METHOD

In the Related Work section, we presented a set of design guidelines that our research team considered as we developed the civic design competition with city leaders months ahead of the official launch (as outlined in section 3). To consider how participants explored the civic design problem-space, we examined the following research questions.

4.1 Research Questions

4.1.1 RQ1. How did competition structure affect the general discussion of the problem space? As presented in section 2.1, competition can help to sustain a regular rhythm of activity among an online community, though too much can be detrimental to collaboration. For this reason, it was important to consider how the level of discussion about the problem-space may have changed from one stage to the next, and throughout the design competition timeline.

4.1.2 RQ2. How does the discussion system factor into exploring the breadth of topics? The discussion environment affects how people find, share and draw connections to the information related to a problem-space (see section 2.2). Though Slack's design integrates familiar features from chat-room and online discussion systems, the integration is somewhat novel. For this reason it was important to consider how elements of Slack's design (e.g., public-private channels, reply-threads) relate to the breadth and depth of topics raised in the general discussion.

4.1.3 RQ3: How does participant interaction affect the coherence of a topic discussion? People also make choices about what they respond to and how they respond. However, as presented in section 2.3, factors related to the content of a message can yield different responses. Following this logic, it was important for us to consider how the content of a message can both elicit a response, and lead to responses that are topically coherent, i.e., adding depth to the existing discussion.

In some ways, these are familiar questions and reflect existing CSCW research about competition and conflict [20], discussion systems [15, 34] and how people talk with each other online [1, 6, 16]. The online aspect of this civic design competition provides a new context to apply mixed-methods research to understand how these questions interact.

4.2 Participants and Participation

The online discussions went live on the evening of Aug 25th (Day 0), although most participants joined around day 30 of the 62 day competition. A typical participant joined and then posted their first message within a few days (average 3-4 days) and within a few more days had found a team (7-10 days from start to joining a team) (See Table 2). This made for a brief window of opportunity for collaborative discussion before setting into their team work.

	Public	Student	Total
Participants	209	43	252
Lurkers	155	5	160
Posted	54	38	92
One time posters	9	1	10
Joined a team	19	39	58
On a winning team	15	8	23
Duration (days)	*Mean (SD)*	*Mean (SD)*	
Start day # (1-62)	29.19 (10.9)	34.37 (5.3)	
Start to first post	3.46 (6.0)	4.03 (4.6)	
Start to join team	7.77 (7.8)	9.81 (7.7)	

Table 2. Participants and participation. Participants were recruited from a local University and from the general public through a publicity campaign that included an article in the city newspaper, online advertisements and direct messaging. Most participants joined the competition near day 30, made a first post within 3-4 days and joined a team within 7-10 days.

Participants used the discussion forum for several purposes. In the process channels (e.g., Events, FAQ, Team Formation), participants posted pictures from the kickoff event with the Mayor and the hackathon, they asked about parking, and they used the *#team-formation* channel to look for a team (@P8): "Hello I'm a product designer (UX and UI design) with focus in accessibility. If any one need help bring concept to hi-fidelity prototype. Please feel free to reach out. I'm always up for the fun [...]" or to recruit team members, "Hi (@P8), our group is currently looking for a video editor and it'd be sick if you could join. We currently have 3 designers and a programmer. We're aiming to win [...]" The in-person events were an opportunity for people to find a team, but also to meet with a team they formed online.

The pattern of messaging in the private Team channels was different than that in the General or Process channels, which were public (see Table 3). A few student teams made the most use of the Team channels, posting at a significantly higher rate—average 65.6 messages per participant—and a greater volume than in other channels. Participants used this space to gather information (e.g., news articles, links to data), share early stage sketches and wire-frames, as well as to coordinate schedules and post funny memes. Team channels were more social spaces than the public General and Process channels, with more *emoji*.[3] Links to external resources were more common in the General channels.[4]

[3]Nearly 2 times as many messages contained Emoji in the Team channels than in other channels: χ^2 (2, 1502)=9.82 **
[4]Participants shared 43% more links in the General channels than in the Team channels: χ^2 (1, 1338)=61.85 ***

	Public		Student		Moderator		
Channels	*Total*	*M (SD)*	*Total*	*M (SD)*	Human	Bot	Total
General	216	5.54 (5.84)	155	4.19 (2.59)	76	29	476
Process	63	2.52 (1.87)	34	2.12 (1.50)	65	2	164
Team	22	3.67 (3.14)	786	65.6 (80.23)	54	0	862
Total	301	5.54 (5.84)	975	4.19 (2.59)	195	31	1502

Table 3. Volume of messages posted by public participants, students, and moderators in the online discussion channels, by type of channel (i.e., General, Process, Team).

Public and student participants talked with each other about the civic design challenges in the General channels, but the level of activity varied by channel. Specifically, fewer of the messages posted to the Accessibility channel led to a reply-thread than messages posted to the Commuter UX and Walking & Biking channels (see Table 4).[5] This observation meant that we needed to statistically control for the differences between channels when modeling participant interaction related to reply-threads.

	Accessibility	Autonomous Veh.	Commuter UX	Walking & Biking	Total
Main channel post	64	54	80	66	264
. . . Initiates thread	13	12	23	26	74 (28%)
Reply-thread post	34	46	61	71	212
Total	98	100	141	137	476

Table 4. Post and reply-thread engagement by Channel.

Just 74 messages (28%) in the General channels led to a reply-thread. Though reply-threads generated a lot of activity (44% of all General messages). Reply-threads varied in length[6] and authorship.[7] Although in only 20% (15) did the author of the initiating message take part in the reply-thread. To better understand how participants talked with each other about the design challenges in the General channels, we hand coded the comments for markers of talk about the design problem and talk with each other (see Table 5).

4.3 Coding for problem-talk, participant interaction and design topics

To understand how participants used the discussion forum to talk about the design challenges, we adapted a method for measuring a deliberations content to examine the "problem-talk" about a civic design issue [64]. When participants in a policy deliberation raise an opinion or offer evidence in support or opposition to a position, they are contributing to the group's analysis of the policy issues—referred to as problem-talk. Problem talk can be explicit, with statements such as "I disagree,

[5]Significantly fewer threads were initiated in the Accessibility channel: $\chi^2(3, 481)=7.84$ *
[6]On average reply-threads included 2.85 messages (2.9 SD) with a maximum of 18 messages
[7]On average reply-threads involved 1.92 (1.16 SD) unique participants with a maximum of 7

but ..." or expressed with more subtlety, such as by raising a rhetorical question. However, when a design team deliberates, they use proposals to elicit questions and constraints as they talk about a problem and its possible solutions [24, 63].

Table 5 defines each of the codes developed to capture problem-talk about a design topic. To code for discussion about the design problem (i.e., constraints, proposals, questions) and how participants talked with each other about designs (i.e., agreement, disagreement, social-talk), two human coders trained on a representative sample of 100 messages until they reached an acceptable level of inter-rater reliability (Cohen's Kappa ≥ 0.80) and then tested on a holdout set of 50 messages. After reaching an acceptable level of reliability, the remainder of the messages were split evenly among the two human coders. To identify the design topic of a message, first each of the human coders independently performed an open coding of the subject of design (e.g., autonomous vehicle, public transit, physical environment), then the codes were compared to create a list of nine topic codes—four for the structured topics, and five that emerged through participant interaction (See Table 7). In total, 476 messages were made to the General channels (e.g., #acc-accessibility); however, only 75% (359) of the messages referenced a design topic.

The design topic codes were also applied to create a measure of topic coherence. Topic coherence refers to "[...] the consistency of the topics within a thread of comments." Discussions with low topic coherence struggle to deeply consider any one topic before shifting to the next [52]. Rather than code for topic coherence [66] we derived a measure of topic coherence by evaluating how each design topic advances (or drifts off) through a sequence of posts. The specific measures are introduced in the next section.

Variable	Definition	Train	Test	Obs.
Design Topic	The subject of design, e.g., ride sharing	0.85	0.89	359
Talk about the design problem				
Constraint	Message includes a design constraint or a barrier, raising a problem but not offering a proposed solution.	0.88	0.93	191
Proposal	Message offers a proposed solution, such that the statement could directly translate into a product or service.	0.92	0.86	148
Question	Message includes a question about the design topic, proposals, or constraints.	0.80	0.91	102
Talk with each other				
Agreement	Message provides agreement to a prior response (e.g., "I agree" "yes" "good idea")	0.89	0.93	44
Disagreement	Messages provides disagreement with existing (e.g., "I don't think that's the right idea"), even implicit disagreement ("I agree, but I think ...")	1.00	1.00	13
Social-talk	Acknowledgements, greetings and salutations, thanking, other non-design focused social interaction	0.91	0.85	202

Table 5. Communication coding definitions and inter-rater reliability.

4.4 Statistical Models and Interpretation

In Section 5.3, we examine how the manner in which participants talk about the design topics affects participant interaction. We fit a set of mixed-effects logistic regressions to model specific forms of participant interaction (i.e., initiating a reply thread, messaging within the main vs thread

space, replying with a topically coherent message). We applied mixed-effect rather than standard logistic regressions to account for the non-independence of both participants and the channels, as some participants posted more messages than others' and some channels were more active than other channels. The models evaluate the following response variables:

- *Initiated Thread*: A binary variable identifying whether a message posted to the main discussion of a channel initiated a reply-thread ("TRUE") or not ("FALSE").
- *Thread vs Main*: A binary variable identifying whether a message was posted as part of a reply-thread or to the main discussion for a channel.
- *Topic Coherence*: A binary variable capturing whether the *next* message, either chronologically in the main discussion channel or the first post initiating a reply-thread, addresses the same design topics as the previous message (see Section 4.3).

Since characteristics other than *talk about the design problem* and *talk with each other* may affect participant interaction, our models include a set of control characteristics related to the message. Comments posted during more active periods of the competition were likely to receive more attention than others. To account for this dynamic, we coded a variable capturing the amount of messaging during the 24-hour period before a message was posted (called "Recent Activity"), which we scaled so that a one-unit increase in the variable reflects a one standard deviation shift in the volume of messages.

Messages that are longer (called "Post Length") or contain links to external references (called "Links") might also attract more participant interaction than shorter messages without references. The Slack interface also treats messages with links differently than messages without, adding an image and title text to make the external material prominent in the discussion interface. For the Topic Coherence model, we also include an indicator variable to reflect whether the topic coherent response to a message was the first in a reply-thread or the next chronologically in the main channel (called "Thread").

In Section 5.3 Table 8, we present the three mixed-effects logistic regression models described above. We exponentiated the logit estimate for each coefficient to report the odds ratio for each response variable. An odds ratio can be interpreted as the change in the response expected by a one-unit increase to an independent variable, holding all others constant. We evaluated the "fitness" of each model to the data with the following procedure: We estimated a model for each response variable with the message characteristics, and then used the log-likelihood ratio test to compare a model that also includes independent variables for how participants talk about the design problem and with each other. This process provides a higher threshold for comparison, as many of the message characteristics significantly relate to the response.

4.5 Post-event survey

Following the awards ceremony at the design summit, we surveyed participants' about their experience and for feedback on the competition organization and communication technology, such as "In a few words, what in the Slack experience worked well for you? What did not?"

Fifty-six participants filled out the survey. Responses from the post-event survey were linked to the comments that survey respondents added to the online discussion system by the same assigned non-identifiable numbers to respondents ranging from 1-252.

5 FINDINGS

5.1 Competition structure

5.1.1 Discussion activity occurred in bursts around events. The depiction of message volume by channel type over time (see Figure 1) captures how sparse the discussion was during early periods

of the competition. Though the discussion spiked around events and deadlines, such as the online Ideation Sprints, the two day hackathon, and the design studio sessions. Participants reported that these events were useful as they worked through the competition: "I was expecting to learn a lot in a short amount of time [during the design sprints]. Having more people help allocate resources really helped kick off our design process," (P26). Both student and public participants took part in the general discussions, though student participants made more use of the private team channels than public participants (see Table 3).

				Hours to initiate a thread by initiator		
	Period	Description	Thrds	Public *M (SD)*	Students *M (SD)*	Moderator *M (SD)*
A.	8/25-9/20	Soft online launch	8	304.5 (259.9)	—	4.36 (6.5)
B.	9/21-9/23	Kickoff and hackathon	9	215.7 (—)	—	4.77 (8.1)
C.	9/24-10/2	Hackathon to design studio	24	6.3 (7.5)	13.8 (23.6)	0.4 (0.5)
D.	10/3-10/10	Design studio to registration	26	12.3 (15.5)	10.1 (18.4)	—
E.	10/11-10/22	Registration to submission	7	—	—	0.9 (0.8)
F.	10/23-10/26	Submission to conference	0	—	—	—

Table 6. Amount of time (in hours) to initiate a thread from a main channel post, by the type of initiator (i.e., public, student, moderator) and by period. A Value less than 1 indicates less than 1 hour.

Participants shared and critiqued ideas as they familiarized themselves with the four mobility challenges. For instance, responding to another in the Commuter UX channel, a participant (P246)

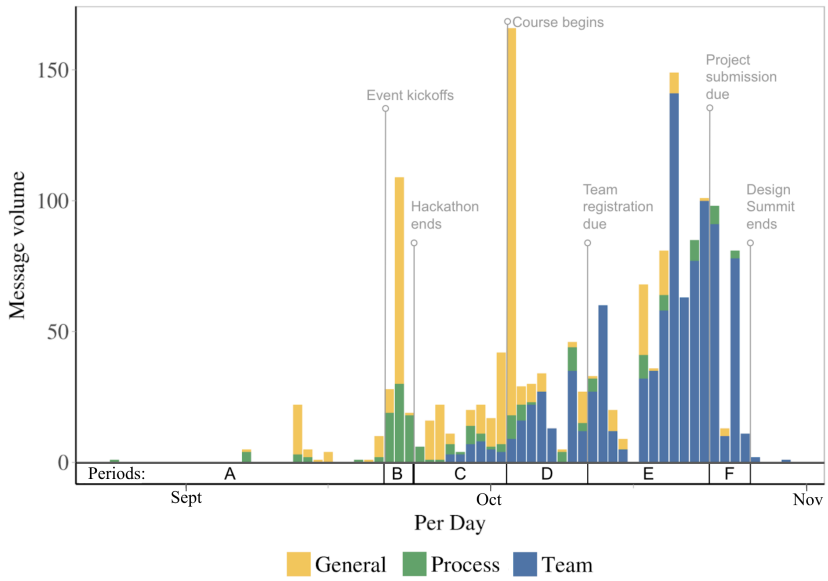

Fig. 1. Post volume over time, by channel type. The online discussions included several public and private channels: four to discuss the topics, three to address process questions about the competition, and fifteen private team channels.

asked, "(@P206) Interesting. How would bike share programs help to solve the last mile problem?"[8] To which a third participant chimed in to disagree with bike sharing as a feasible solution (P105):

> "While I love bikes and think they're a great way to get around, I don't think they are the solution for [our city]. There are too many hills and it's generally too warm for many people to ride comfortably. Additionally, a bike sharing program would mean that bikes would sit for long periods of time in lower density residential areas. Perhaps we could have a fleet of golf carts dropping people off."

This response expresses disagreement, but also adds design constraints to the last-mile problem, such as (1) biking up hills, (2) biking on hot days, and (3) low levels of use in low density areas.

Participants also shared design proposals during periods C & D. In the Walking & Biking channel, a message describing the problem of cars approaching an intersection and (P2) "[...] creeping into the bike lane maybe 50-200ft before the turn," generated a particularly long thread of responses. The reply-thread echoed with agreement and included a few design proposals, such as (P12) "adding lane dividers or bumps [to] help [limit] the behavior of drivers [or] widening the bike lane near intersections so that there is a special passing zone," and (P2) changing the law to allow bikes to ride on the sidewalk in business districts—an idea that generated some disagreement (P206).

5.1.2 Participant engagement with Helper Bot was mixed. We used Helper Bot to initiate Ideation Sprints as a way to trigger bursts of discussion activity. On day 31 (Sept 25th), we launched a 24 hr Ideation Sprint to the Walking & Biking channel, with the general call: "Let's generate as long of a list as we can of ways to detect that an intersection is becoming dangerous." Nine participants took part in the sprint and generated a list of twelve ideas with links to relevant articles, stretching the conversation from day 31 to 33. However, the first round of Ideation Sprints had varying levels of success, generating no response in the Accessibility channel, though a few messages in the Commuter UX and Autonomous Vehicles channels.

As we noticed that the general discussion was starting to dwindle, we launched a second round of Ideation Sprints on day 52 (Oct 16th), asking participants to vote on a forgotten idea or topic from the discussion that we should revisit. To facilitate the voting, we developed a message interface with several buttons (see Figure 2). We had also programmed Helper Bot to immediately respond to a button press by directly challenging the participant's vote with a message like "*@participant* don't you think that **Live Data** is too challenging?" The in-situ and in real-time challenge from Helper Bot elicited immediate participant reactions. In the Autonomous Vehicles channel, Helper Bot's message "*@participant* don't you think that **Ethics** is too challenging?" led to an extensive thread discussion (19 messages) about ethical issues, including the runaway trolley problem, but also less popularized dilemmas, such as:

- Should AVs always follow the rules, even when it slows down traffic
- Should AVs get access to HOV lanes
- How should curb-space access be prioritized in urban environments
- Should AVs be restricted to specific services (i.e., to protect some human jobs or industry)?
- After the AV selects an optimal route, do the passengers have say over when/where they get dropped off?

5.1.3 Discussion activity decreased following the team registration. Figure 1 also captures the transition from general discussion to team driven activity as the competition progressed. The volume of messages to the General channels decreased following team registration (period D) (see Table 6). During periods A & B, moderators initiated nearly all of the thread responses, but

[8] In transportation and planning policy, the *last-mile* describes the movement of people from a transportation hub (e.g., bus stop, trolley station) to their final destination.

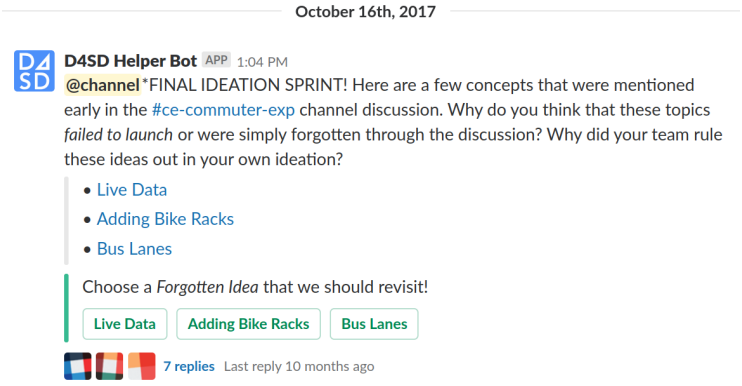

Fig. 2. Ideation Sprint challenge to the Commuter UX channel on Oct 16th

participants initiated much of the interaction during periods C & D. Table 6 also indicates that responsiveness was high during periods C-D as threads were initiated within a few hours of a post. That changed in the run-up to the team submission deadline (periods E & F), as participation dropped off in the general channels and picked up in the private team channels.

In their post-survey comments, participants expressed that they felt the pressure to compete made some teams less willing to share. As one participant explained it (P12):

> "[the online discussion system] was effective in showing the problem space in each topic area and helping guide further research into problems to be solved. However, due to the competitive nature of the challenge, people stopped sharing their further research which might have reduced the quality of all projects. So, at that point, we forgot about Slack as no one was willing to help other projects at expense of their own"

Another respondent hoped for more from the Ideation Sprints, that "[...] other people would also share their ideas and not hold back. Was hoping people weren't trying to keep their ideas secret to themselves because it's a competition," (P40).

5.2 Discussion system

5.2.1 A majority of the messages added to interactional design topics. Some topics generated a lot of messages within a channel, yet some topics spanned multiple channels (see Table 4). In the post-survey participants described the channels as a shared space to collect relevant information (e.g., news clippings), but also as an opportunity to vet possible solutions. Describing their use of the discussion system, "channels were used to gain a big picture idea of the scope and range of topics within each topic area, and even provide research that could help teams narrow in on specific problem areas," (P12). Other participants valued the opportunity for critique, "I thought it was great to be able to bounce ideas off of other people and see how other people were interpreting the prompt" (P28).

Not all participants were as positive about the usefulness of the channels, "I liked that it was broken into separate challenge conversations but they seemed almost too free-flow" (P167), "[...] sometimes navigating through Slack was a bit confusing," (P39) and "[...] the channels were chaotic and hard to follow," (P24). These frustrations are not surprising, as organizational conversation systems, like Slack, are subject to topical shifts with each new message [66]. Table 7 indicates that less than half of the discussion messages were on the topics structured for the discussion channels

(e.g., Accessibility). Participants introduced new foci for design as they talked with each other about the design challenges (e.g., public transit, bike sharing, etc.).

	General Channels				
Structured Topics	Accessibility	Autonomous Veh.	Commuter UX	Walking & Biking	Total
Accessibility	9	0	0	0	9
Autonomous Veh.	1	80	2	0	83
Commuter UX	0	0	12	0	12
Walking & Biking	0	0	0	45	45
Total	10	80	14	45	149
Interactional Topics					
Driving Culture	5	1	15	17	38
Personal Automobiles	6	3	18	8	35
Physical Environment	8	1	3	40	52
Public Transit	11	0	60	0	71
Ride Sharing	8	0	6	0	14
Total	38	5	102	65	210

Table 7. Structured and Interactional Topics by Channel. Not surprisingly many of the post topics corresponded with the topics structured for the channel (i.e., 80 posts to the #autonomous-cars channel were able autonomous vehicles); however, a majority of the posts introduced other interactional topics. Unlike the structured topics, the emergent topics spanned multiple channels (i.e., ride sharing was discussed in both the #acc-accessibility and #ce-commuter-exp channels)

5.2.2 Missed opportunities to connect discussions across channels. As new topics formed through participant interaction within a channel, participants may not have noticed similar conversations in other channels. For example, a common problem is how pedestrians and drivers communicate with each other. In the Accessibility channel, participants suggested that the (P10) "steering wheel or a seat [might] vibrate," or a message might (P29) "override the driver's stereo system [...] in order to further grab the[ir] attention," about the presence of a pedestrian. One participant in the Walking & Biking channel proposed incentives and systems to monitor signalling behavior, (P173) "maybe insurance would be a little lower for nice drivers." The problem is somewhat different for autonomous vehicles, as a participant explained, (P71) "[...] for a walker or biker in particular, knowing that they've been *seen* by the AV might contribute to their perception/feelings that it's now their turn to cross the road." These messages address a similar design problem, yet were not connected across channels.

"Driving Culture" was another important missed connection. In the Walking & Biking channel, participants debated whether car-cyclist collisions could be addressed by design, or if the problems stem from willfully disobeying the laws and a lack of respect for others on the road. Feeling strongly that the problem is a lack of respect for cyclists, (P13) responded with "Rather than the problem being the markings on the ground or signage, the culture is to not treat cyclists with the same amount of respect as a vehicle," to which another participant responded, (P4) "[...] However, thinking

about your point regarding respect (@P13), I agree that drivers must recognize and give bikers just as much respect on the road, if not more. That being said, I have often seen bikers do things to avoid obeying regular traffic rules." The debate went on in Walking & Biking, while in other channels participants discussed how urban planning could promote better car-cyclist interaction, drawing examples from other countries (e.g., Denmark, Japan). Though similar in topic, the conversations were not connected across the channels.

5.3 Participant interactions

5.3.1 Participants responded to messages that added constraints. When participants responded to each other with constraints, proposals and questions they added insight to the collective understanding of the mobility issues. For this reason, it was important to understand what types of messages led to more responses, and to responses that built on existing topics. We found that reply-thread responses were an important factor in continuing the discussion about a topic.

To consider how characteristics of the content in a message relate to initiating a reply-thread, we fit the mixed-effects logistic regression reported in Table 8 *Initiate Thread*. The timing of a message matters. Messages posted during periods with a higher level of recent activity were more likely to initiate a reply-thread. This mirrors our observation that some stages of the competition were much more interactive than others (e.g., periods C & D).

In addition to being well timed, offering insight about the design problem contributed to a response. Messages that described a design constraint were 2.59 times as likely to initiate a reply-thread than messages that did not. Similarly including a link increased the likelihood of a reply-thread by 2.26 times. Participants responded to messages that added substance, in terms of design constraints and links, to the collective knowledge of a mobility issue.

	Initiate Thread	*Thread vs Main*	*Topic Coherence*
Post characteristics	OR (Std Dev)	OR (Std Dev)	OR (Std Dev)
(Intercept)	0.27 (0.45) **	0.39 (0.50) .	0.23 (0.94)
Thread	—	—	6.26 (0.49) ***
Recent Activity (scaled)	0.47 (0.24) **	1.58 (0.26) .	1.21 (0.31)
Post length (scaled)	1.37 (0.19) .	0.55 (0.22) **	0.69 (0.24)
Link	2.26 (0.36) *	0.16 (0.37) ***	0.55 (0.51)
Talk about the design problem			
Proposal	0.92 (0.36)	1.06 (0.32)	3.88 (0.48) **
Constraint	2.59 (0.38) *	0.84 (0.34)	4.71 (0.50) **
Question	1.06 (0.43)	0.46 (0.38) *	1.88 (0.57)
Talk with each other			
Agreement	0.53 (0.90)	2.72 (0.48) *	0.66 (1.06)
Disagreement	12.25 (1.26) *	1.89 (0.90)	1.16 (1.63)
Social-talk	0.56 (0.39)	2.66 (0.31) **	0.32 (0.53) *
Goodness-of-fit:	$\chi^2(6, 264)=14.54$ *	$\chi^2(6, 447)=19.83$ **	$\chi^2(6, 264)=24.50$ ***

Table 8. Mixed-effects regression models nested on *author* and *channel name*.

Only a few design disagreements were identified in the analysis (13 total), though disagreements were significantly more likely to receive a response. For example, the following message argues that more public transit may not be the solution, and calls out to another participant (P36) for support in the disagreement, "(@P42) I feel that even if we create a cheap public transportation

system, it wouldn't do much to encourage people to use it instead of just driving. This point is noted by (@P39) by how she would much rather drive since it was much more effective [...]"

5.3.2 Messages within a reply-thread were more social. To understand how messages posted to a reply-thread might differ from those posted to the main channel, outside of a reply-thread, we fit a mixed-effects logistic regression reported in Table 8 *Thread vs Main*. Reply-threads tended to be shorter and contain fewer links than posts to the main channel. To our surprise, reply-threads were also less likely to contain questions, by a factor of 0.45 times. This may mean that when participants added questions, it was more often to introduce a new point to the channel than to build on (or examine) points already posed by an existing message.

Thread messages involved more social-talk, such as greetings and appreciation, and agreement to a design idea or observation than messages to the main channel, e.g., (P147) "All great ideas! I agree these would be important to collect. Could tie into the city's Get It Done app?" The implication is that reply-threads might be a good place for participants to look for possible team members. Though the results also suggest that participants might need additional support asking design questions about the proposals that they generate through a reply-thread.

However, some messages posted to a reply-thread did not build on topics in the original message. For example, an initial message about public transit, (P181) "I like the idea of the buses as it takes tons of car[s] out of circulation, but I feel like leaving each person at their doorstep may be not efficient (thinking on time and fuel)," shifted to a new mode of transportation in the next response, (P241) "Autonomous vehicles could play a similar role for commuter in general, moving people to areas of higher concentration and back again." Then the thread shifted again to a personal question, (P241) "Will you be at the event tonight? I'll be arriving a bit later." Such seemingly minor shifts affect the topic coherence of a discussion. When a design discussion is regularly off-topic, or moves away too quickly it cannot deeply consider a problem-space [52, 66].

5.3.3 Reply-thread messages were more topic coherent. To follow the topic coherent responses, whether to a reply-thread or within the main channel, we fit the mixed-effects logistic regression reported in Table 8 *Topic Coherence*. As the model predicts a coherent response, the coefficient "Thread" indicates that reply-thread responses were significantly more likely to be topic coherent than main channel responses to a message. The model also indicates that messages with design constraints and proposals contributed to topic coherent responses. Messages that contained social-talk (e.g., greetings, appreciation) were less likely to lead to a topic coherent response. It is interesting to note that including links and post length had no effect on the likelihood of a coherent response. The takeaway is that messages with a proposal to address some specific constraints more often led to reply-threads that built on the initial topic, effectively adding depth to the problem-space exploration.

Figure 3 presents a thread of messages that address the initial post about the design constraints of traveling with a bike on public transit. Messages posted in response share a similar experience as the initiating message, extend the topic to consider the broader cultural issues facing the city, and provide specific facts about bikes on the trolley system. While the thread does not build toward a design proposal for this situation, the conversation adds experience, facts and insights to the problem-space.

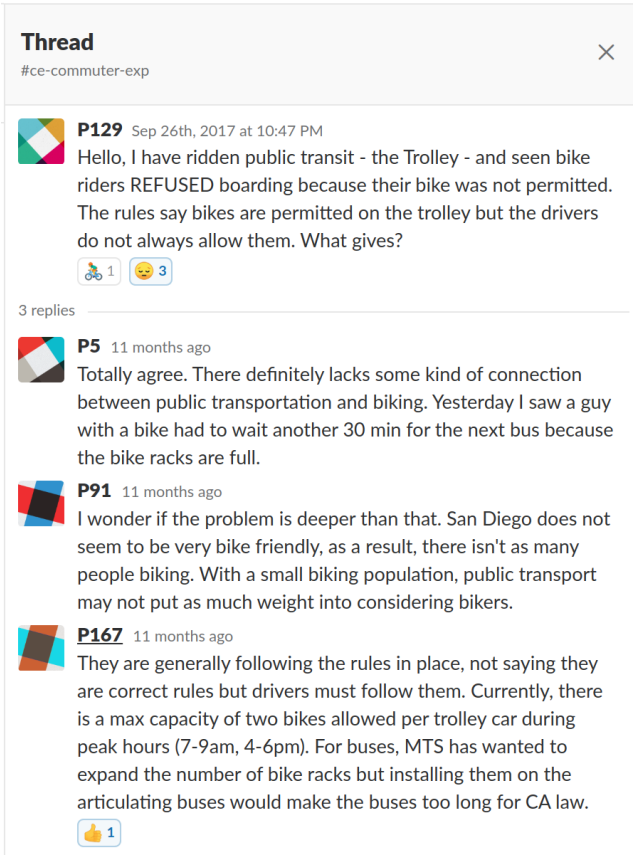

Fig. 3. Topic coherent thread about the design constraints to traveling with a bike on public transit. To conceal the identities of the participants, the user names have been changed to the P#'s used throughout the paper.

6 DISCUSSION

On day 62, the competition judges had selected eight finalists from the 23 team submissions. The winning teams proposed novel sensing systems to improve the visibility of bicycles to vehicles, developed educational and assistance services to introduce autonomous vehicles onto city streets, and created a plan to provide WiFi access on all metropolitan transit shuttles. These ideas were pitched to an audience of designers, entrepreneurs and city officials at a local design summit. Beyond the specific ideas, the discussions throughout the competition enriched our collective understanding of the transportation problem-space.

To better understand how factors of the competition structure and the online discussion forum affected the way that participants contribute to the shared understanding of a problem-space, we structured the analysis around our three research questions and offer implications that could be valuable toward future civic design competitions. We draw from this case, several open questions about how to create effective discussion systems for collective problem-solving.

6.1 How did competition structure affect the general discussion of the problem space?

Discussion activity spiked during online and in-person events. We had planned the competition timeline expecting some early discussion of the issues before the Kickoff event and Hackathon (Sept 21-23rd), but instead the in-person events spurred online activity. This meant that most participants joined the discussion around day 30 and found a team 7-10 days later, just in time to meet the team registration deadline on day 46 (Oct 10th). This abbreviated period made for a swift transition from a collaborative issue-driven discussion to focused team work. We observed the volume of discussion shift from public to private channels (see Figure 1), participant-to-participant interaction drop (see Table 4), and through the post-survey, participants talked about feelings of competitiveness that precluded them from sharing.

6.1.1 Curb the cost (or boost the benefit) to sharing. Future civic design competitions might include nudges that encourage teams to share some of their findings with other participants. As an example, rather than focus the competition on team-based rewards, there could be a set of rewards for contributions to the collective: e.g., for sharing research about community stakeholders, for crafting discussion prompts that spark a topic coherent conversation. Such incentives might help carry the general topic conversations past the team registration deadline. Competition organizers could also consider creating opportunities for team-to-team design critique. Several teams working through a similar topic might be willing to share their design constraints with each other as a way to talk through some of their common challenges, without divulging a specific proposal. We think that such measures would make it possible for the community at large to continue to benefit even after competitiveness sets in.

6.1.2 Experiment with the Ideation Sprint paradigm. Online events, such as Ideation Sprints, also triggered a burst of activity. The first wave of Ideation Sprints, launched in the days following the hackathon, met with with mixed success. An Ideation Sprint in the Walking & Biking channel generated numerous ideas, yet the sprint in the Accessibility channel failed to start. The second wave was similar, with more activity in the Autonomous Vehicle and Commuter UX channels than others. Future HCI research might experiment with this paradigm in a controlled setting, such as with a large University level course, where it might be possible to try various topics and designs repeatedly to understand how factors like the question type, topic and bot involvement with the team(s) affect performance (e.g., count, quality, or creativity of ideas).

6.2 How does the discussion system factor into exploring the breadth of topics?

We found that more than half of the messages added new design topics to the discussion. By tracking what the messages actually contributed to the design topics, we noticed that participants often added topics to a channel that paralleled similar discussions in other channels. This made for a missed opportunity to bridge the channels along broader issues, such as driving culture and pedestrian-to-vehicle communication.

6.2.1 Visualize and synthesize topic discussions. It was clear from the post-survey feedback, that some participants found the discussion in the channels chaotic and hard to follow. Visual representations of the topics and their shift over time might help. The connections between topics and across channels, could be made more salient by visualizing the design topic network, as proposed in other HCI research [14, 66]. While visuals can help to identify the latent topics, we also could have set a discussion agenda with designated periods for each issue and a formal process, *á la* Robert's Rules [58, 68], for participants to propose (or even lead) a discussion.

To make the topic discussions more useful we might look to integrate discussion summary tools, like *Wikum* [75]. For example, a Wikum-like tool might be particularly valuable after an Ideation

Sprint, to quickly summarize the insights and references into a succinct paragraph or when a channel discussion shifts to a new sub-topic. The specific implementation might integrate Wikum as part of the Slack "readline" feature, as a way to make better use of the time and cognitive effort that people spend "catching up" on the discussion.

6.3 How does participant interaction affect the coherence of a topic discussion?

Messages that added substance to the design topics, in terms of constraints and links, increased the likelihood of a reply-thread. While the subsequent messages posted to a reply-thread were often topically coherent (66.2%, N=49, see Table 4), we also found that about a third of the main channel posts that did not initiate a reply-thread still led to a topically coherent response (33.5%, N=67). What we take this to mean is that participant's added breadth to the channels by contributing messages that built on existing topics, but also added depth to specific topics, by initiating or continuing a topically coherent reply-thread.

6.3.1 Consider bot messaging to elicit design questions. Posts to a reply-thread were less likely to include design questions than posts to the main channel. From the perspective of a competition organizer, this is a concern: Messages to a reply-thread often include agreement, but lack the design questions that might critique an idea or request clarification. In future efforts we might consider a similar in-situ and in real-time response to reply-thread messages that include statements like, "I agree." A bot might listen for this sentiment and immediately reply with a message *only visible to the user* that offers a few design questions for them to consider. Or, a *snarky*-bot might play the devil's advocate and directly challenge a participant who simply states their agreement.

6.3.2 Design reply-threads to add to the knowledge structure. Reply-threads mattered more to the design topic discussions than we had anticipated in planning the civic design competition (at one point we actually debated about whether people would use this feature). For that reason several of our recommendations propose ways to improve the conversation within the reply-thread. However, in Slack, the reply-thread has a somewhat ephemeral existence. When a reply-thread is initiated it provides a space for people who contribute to the thread to continue talking with each other, even though the main channel discussion has moved well past the initial thread message. As the link to a reply-thread moves further into the channel history, that side discussion can be hard for new people to join. Future HCI research should consider how to index the insights from a thread discussion within the knowledge structure for a discussion forum.

6.4 Knowledge Structures in Discussion Systems

The analysis details several organizational learning gaps in the collective understanding of the transportation problem-space that emerged during the civic design competition. For one, similar design topics were introduced in separate channels. This reflects a missed opportunity to strengthen the collective understanding around these topics. For instance, a discussion system might recognize a similar topic and send a message to the group of participants involved in the parallel conversations, calling their attention to the common topic. While we developed a hand-coding procedure for recognizing similar topics across the channels, latent topic analysis and other natural language processing techniques might be a natural next step.

We also found that the content of a message matters when eliciting a response. Specifically messages that talked about the design problem, by offering constraints or proposals, more often yield a topic coherent response. This draws attention to the fundamental role of "talk" when exploring a problem-space. When people struggle to articulate their thoughts or observations of a design problem, others' may be less able or willing to talk with them. The result is a missed opportunity to share insights, but also for people to meet and to mentor each other. Existing

systems developed to crowdsource design critique might be adapted to help people articulate their understanding of a design problem [38, 45, 46], while active listening systems, like *Reflect* [37], might be adapted to provide conversational support for mentoring.

7 LIMITATIONS

An underlying tension implicit in this case study was how to navigate our joint responsibilities as researchers, educators and competition organizers. As researchers, we chose to present a case study of how people talk with each other online about civic topics. Unlike a controlled experiment, a case study provides an opportunity to deeply consider a situation, a group of people, and how they interact with each other in a naturalistic setting. If instead we had chosen to conduct an experiment, teams exposed to one condition may have had an unfair advantage in the competition over others assigned to other conditions. For this reason, we present our findings as potential opportunities for competition organizers and system designers.

As educators, we developed a design studio course and other in-person events to teach participants human-centered design concepts and technical skills. Some of the online discussion activity was driven by the fact that the design studio course prompted students to post on the discussion channel during the first day of class. Despite this, the studio sessions provided a valuable service to competitors and an authentic glimpse into the interactions between students and non-student participants.

As the competition organizers, it was important to encourage teams to adopt the technology and strategies that worked best for them. We chose Slack as the official online discussion platform because of its stability and availability on multiple devices and operating systems. However, our post-event survey showed that it was not the dominant channel for internal team communication. Many respondents used Facebook Messenger, while other teams reported using in-person meetings and text messages as their primary communication methods. We did not coordinate or monitor team-level activity outside of Slack and the scheduled in-person events, and thus we may not have a full picture of how the competition participants navigated the problem-space.

The competition also took place in one city and on one major problem-space, emphasizing four broad, yet specific topics. It remains to be seen if similar patterns would emerge in other places and on other topics. For example, had we focused the competition on a more controversial topic, such as land-use or public school improvements, the nature of the discussion might have included more conflict or even direct personal attacks [11, 47]. To preempt such behavior, we setup the discussion forum so that participant identities would not be anonymous [22, 26].

Finally, we reported on online discussions that took place during a civic design competition, but we do not have enough data on how or whether the discussion behavior affected actual performance in the design challenge.

8 CONCLUSION

The paper presents a mixed-method case study of the online discussions surrounding a civic design competition about transportation issues facing a large U.S. metropolitan area. 92 participants contributed to the online discussion during a 62-day period, generating over 1,500 messages overall. The analysis focuses on a subset of 476 messages posted to the public general channels about four major challenges in transportation (i.e., Accessibility, Autonomous Vehicles, Commuter UX, Walking & Biking). We discuss how factors of the competition structure, discussion system and content of messages affected the ways participants interacted with the design topics, as well as with each other. The competition structure seemed to create spikes in discussion activity around online and in-person events, such as the Hackathon and Ideation Sprints. As the competition progressed,

competitiveness among the teams led to lower levels of public participant interaction about the design topics.

We also found that the conversation about design topics was somewhat divided by the technological barrier imposed by the discussion structure. Specifically, several topics emerged through participant interaction that were common across the channels, though these parallel conversations were never linked together (or went unnoticed). Through an analysis of the types of message content that led to participant interaction, we found that: (a) messages that include design constraints, links or disagreements more often yielded responses, (b) messages within a reply-thread were more social, included more agreement, yet fewer design questions, and (c) messages containing a design proposal or constraint more often led to a coherent reply-thread response than messages that did not contain these features. These observations are useful for civic design competition organizers and provide implications for building knowledge through discussion systems.

ACKNOWLEDGMENTS

The 2017 D4SD mobility challenge was a community effort that involved local entrepreneurs, academics and regional government leaders. A large team of people helped to organize the competition, though we want to specifically acknowledge the contributions of Don Norman, Michele Morris, Narges Mahyar, Yujin Cho, Brandon Hong, Joseph Le, Lauren Liu, and Gabriel Amoako. The data analysis would not have been possible without the expert assistance of Claire Han, Simran Shinh and Varun Talwar. Thanks to our partner organizations the Design Forward Alliance and ScaleSD as well as the City of San Diego, Port of San Diego, and SANDAG. Our research was supported by the National Science Foundation (Grants #1122206 and #1122320).

REFERENCES

[1] Jaime Arguello, Brian S. Butler, Elisabeth Joyce, Robert Kraut, Kimberly S. Ling, Carolyn Rosé, and Xiaoqing Wang. 2006. Talk to me: foundations for successful individual-group interactions in online communities. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 959–968.

[2] Michael S. Bernstein, Joel Brandt, Robert C. Miller, and David R. Karger. 2011. Crowds in two seconds. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. ACM Press, New York, New York, USA, 33. https://doi.org/10.1145/2047196.2047201

[3] Osvald Bjelland and Robert Wood. 2008. An inside view of IBM's Innovation Jam's. *MIT Sloan Management Review* 50, 1 (2008), 32.

[4] Laura W. Black. 2012. Blog, chat, edit, text, or tweet? Using online tools to advance adult civic engagement. *New Directions for Adult and Continuing Education* 2012, 135 (2012), 71–79.

[5] Laura W. Black, Howard T. Welser, Dan Cosley, and Jocelyn M. DeGroot. 2011. Self-Governance Through Group Discussion in Wikipedia. *Small Group Research* 42, 5 (10 2011), 595–634. https://doi.org/10.1177/1046496411406137

[6] Moira Burke and Robert Kraut. 2008. Mind your Ps and Qs: the impact of politeness and rudeness in online communities. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 281–284.

[7] Moira Burke, Cameron Marlow, and Thomas Lento. 2009. Feed me: motivating newcomer contribution in social network sites. *Proceedings of the SIGCHI conference on* (2009). http://dl.acm.org/citation.cfm?id=1518847

[8] Stephanie Burkhalter, John Gastil, and Todd Kelshaw. 2002. A conceptual definition and theoretical model of public deliberation in small face-to-face groups. *Communication Theory* 12, 4 (2002), 398–422.

[9] Robert Cavalier, Miso Kim, and Zachary S. Zaiss. 2009. Deliberative democracy, online discussion, and project PICOLA (Public informed citizen online assembly). *Online deliberation: Design, research, and practice/Eds. Davies T., Gangadharan SP Stanford, CA: Center for the study of language and information* (2009), 71.

[10] Boreum Choi, Kira Alexander, Robert E. Kraut, and John M. Levine. 2010. Socialization tactics in wikipedia and their effects. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10*. ACM Press, New York, New York, USA, 107. https://doi.org/10.1145/1718918.1718940

[11] Kevin Coe, Kate Kenski, and Stephen A. Rains. 2014. Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments. *Journal of Communication* 64, 4 (8 2014), 658–679. https://doi.org/10.1111/jcom.12104

[12] Stephen Coleman and John Gøtze. 2001. *Bowling together: Online public engagement in policy deliberation*. Hansard Society London.

[13] Jeff Conklin and Michael L. Begeman. 1988. gIBIS: a hypertext tool for exploratory policy discussion. *ACM Transactions on Information Systems* 6, 4 (10 1988), 303–331. https://doi.org/10.1145/58566.59297
[14] Kushal Dave, Martin Wattenberg, and Michael Muller. 2004. Flash forums and forumReader: navigating a new kind of large-scale online discussion. In *Computer Supported Cooperative Work*. ACM, 232–241.
[15] Todd Davies and Reid Chandler. 2011. Online Deilberation Design: Choices, Criteria, and Evidence. *Democracy in motion: Evaluating the practice and impact of deliberative civic engagement* (2011), 103–131.
[16] Richard Davis. 1999. *The web of politics: The Internet's impact on the American political system*. Oxford University Press.
[17] Fiorella De Cindio and Cristian Peraboni. 2010. Design issues for building deliberative digital habitats. *Online Deliberation* (2010), 41.
[18] Fiorella De Cindio, Cristian Peraboni, and Leonardo Sonnante. 2008. A two-room e-deliberation environment. *Proceedings of Tools for Participation: Collaboration, deliberation, and decision support* (2008), 47–59.
[19] Carl DiSalvo. 2009. Design and the Construction of Publics. *Design issues* 25, 1 (2009), 48–63.
[20] Steve M. Easterbrook, Eevi E. Beck, James S. Goodlet, Lydia Plowman, Mike Sharples, and Charles C. Wood. 1993. A survey of empirical studies of conflict. In *CSCW: Cooperation or Conflict?* Springer, 1–68.
[21] Dmitry Epstein and Gilly Leshed. 2016. The Magic Sauce: Practices of Facilitation in Online Policy Deliberation. *Journal of Public Deliberation* 12, 1 (2016).
[22] Rolf Fredheim, Alfred Moore, and John Naughton. 2015. Anonymity and Online Commenting: The Broken Windows Effect and the End of Drive-by Commenting. In *Proceedings of the ACM Web Science Conference*. ACM, Oxford, United Kingdom, 11. https://doi.org/10.1145/2786451.2786459
[23] Deen Freelon. 2015. Discourse architecture, ideology, and democratic norms in online political discussion. *New Media & Society* 17, 5 (2015), 772–791.
[24] Elizabeth Gerber. 2007. Improvisation principles and techniques for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1069–1072. https://dl.acm.org/citation.cfm?id=1240786
[25] Aspa Gospodini. 2002. European Cities in Competition and the New ' Uses ' of Urban Design. *Journal of Urban Design* 7, 1 (2 2002), 59–73. https://doi.org/10.1080/13574800220129231
[26] Daniel Halpern and Jennifer Gibbs. 2013. Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior* 29, 3 (2013), 1159–1168. https://doi.org/10.1016/j.chb.2012.10.008
[27] Katja Hutter, Julia Hautz, Johann Füller, Julia Mueller, and Kurt Matzler. 2011. Communitition: The Tension between Competition and Collaboration in Community-Based Design Contests. *Creativity and Innovation Management* 20, 1 (3 2011), 3–21. https://doi.org/10.1111/j.1467-8691.2011.00589.x
[28] Luca Iandoli, Ivana Quinto, Anna De Liddo, and Simon Shum. 2014. Socially augmented argumentation tools: Rationale, design and evaluation of a debate dashboard. *International Journal of Human-Computer Studies* 72, 3 (2014), 298–319.
[29] Elisabeth Joyce and Robert E. Kraut. 2006. Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication* 11, 3 (2006), 723–747.
[30] Christopher F. Karpowitz, Chad Raphael, and Allen S. Hammond. 2009. Deliberative Democracy and Inequality: Two Cheers for Enclave Deliberation among the Disempowered. *Politics & Society* 37, 4 (12 2009), 576–615. https://doi.org/10.1177/0032329209349226
[31] Aniket Kittur and Robert E. Kraut. 2008. Harnessing the wisdom of crowds in wikipedia. In *Proceedings of the ACM 2008 conference on Computer supported cooperative work - CSCW '08*. ACM Press, New York, New York, USA, 37. https://doi.org/10.1145/1460563.1460572
[32] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1301–1318.
[33] Mark Klein. 2011. How to harvest collective wisdom on complex problems: An introduction to the mit deliberatorium. *Center for Collective Intelligence working paper* (2011).
[34] Robert E. Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. 2012. *Building successful online communities: Evidence-based social design*. Mit Press.
[35] Travis Kriplean, Caitlin Bonnar, Alan Borning, Bo Kinney, and Brian Gill. 2014. Integrating on-demand fact-checking with public dialogue. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*. ACM Press, New York, New York, USA, 1188–1199. https://doi.org/10.1145/2531602.2531677
[36] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. 265–274.
[37] Travis Kriplean, Michael Toomim, Jonathan Morgan, Alan Borning, and Andrew Ko. 2012. Is this what you meant?: promoting listening on the web with reflect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1559–1568.

[38] Chinmay Kulkarni, Koh Pang Wei, Huy Le, Daniel Chia, Kathryn Papadopoulos, Justin Cheng, Daphne Koller, and Scott R. Klemmer. 2013. Peer and self assessment in massive online classes. *ACM Transactions on Computer-Human Interaction* 20, 6 (12 2013), 1–31. https://doi.org/10.1145/2505057

[39] Werner Kunz and Horst Rittel. 1970. Issues as elements of information systems. (1970).

[40] Karim R. Lakhani, David A. Garvin, and Eric Lonstein. 2010. *TopCoder (A): Developing Software through Crowdsourcing*. Technical Report. Harvard Business School. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2002884

[41] Christopher Le Dantec and Keith Edwards. 2008. Designs on dignity: perceptions of technology among the homeless. *Proceedings of the SIGCHI conference on* (2008). http://dl.acm.org/citation.cfm?id=1357155

[42] Ann Light and Yoko Akama. 2012. The human touch. In *Proceedings of the 12th Participatory Design Conference on Research Papers: Volume 1 - PDC '12*. ACM Press, New York, New York, USA, 61. https://doi.org/10.1145/2347635.2347645

[43] Ann Light and Ann. 2010. The unit of analysis in understanding the politics of participatory practice. In *Proceedings of the 11th Biennial Participatory Design Conference on - PDC '10*. ACM Press, New York, New York, USA, 183. https://doi.org/10.1145/1900441.1900473

[44] Yong Lu and Dan Yang. 2011. Information exchange in virtual communities under extreme disaster conditions. *Decision Support Systems* 50, 2 (1 2011), 529–538. https://doi.org/10.1016/J.DSS.2010.11.011

[45] Kurt Luther, Amy Pavel, Wei Wu, Jari-lee Tolentino, Maneesh Agrawala, Björn Hartmann, and Steven P Dow. 2014. CrowdCrit: crowdsourcing and aggregating visual design critique. In *Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing*. 21–24.

[46] Kurt Luther, Jari-Lee Tolentino, Wei Wu, Amy Pavel, Brian P. Bailey, Maneesh Agrawala, Björn Hartmann, and Steven P. Dow. 2015. Structuring, Aggregating, and Evaluating Crowdsourced Design Critique. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*. ACM Press, New York, New York, USA, 473–485. https://doi.org/10.1145/2675133.2675283

[47] Rousiley C. M. Maia and Thaiane A. S. Rezende. 2016. Respect and Disrespect in Deliberation Across the Networked Media Environment: Examining Multiple Paths of Political Talk. *Journal of Computer-Mediated Communication* 21, 2 (3 2016), 121–139. https://doi.org/10.1111/jcc4.12155

[48] Thomas W. Malone and Kevin Crowston. 1990. What is coordination theory and how can it help design cooperative work systems?. In *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*. 357–370.

[49] Lena Mamykina, Drashko Nakikj, and Noemie Elhadad. 2015. Collective Sensemaking in Online Health Forums. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 3217–3226. https://doi.org/10.1145/2702123.2702566

[50] Jane Mansbridge, Janette Hartz-Karp, Matthew Amengual, and John Gastil. 2006. Norms of deliberation: An inductive study. (2006).

[51] James G. March. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science* 2, 1 (2 1991), 71–87. https://doi.org/10.1287/orsc.2.1.71

[52] Brian McInnis, Dan Cosley, Eric Baumer, and Gilly Leshed. 2018. Effects of Comment Curation and Opposition on Coherence in Online Policy Discussion. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork*. ACM, Sanibel Island, Florida, 347–358.

[53] Brian McInnis, Dan Cosley, Chaebong Nam, and Gilly Leshed. 2016. Taking a HIT: Designing around Rejection, Mistrust, Risk, and Workers' Experiences in Amazon Mechanical Turk. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2271–2282.

[54] John Parkinson. 2006. Holistic democracy and physical public space. In *British Journal of Political Science Conference*. 1–17.

[55] Jennifer Preece and Ben Shneiderman. 2009. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer* (2009).

[56] Cynthia R. Farina, Dmitry Epstein, Josiah B. Heidt, and Mary J. Newhart. 2013. RegulationRoom: Getting more, better civic participation in complex government policymaking. *Transforming Government: People, Process and Policy* 7, 4 (2013), 501–516.

[57] Horst W. Rittel and Melvin M. Webber. 1973. 2.3 Planning problems are wicked. *Polity* 4 (1973), 155–169.

[58] Henry M. Robert. 2011. *Robert's rules of order newly revised.*

[59] Lynn M. Sanders. 1997. Against deliberation. *Political theory* 25, 3 (1997), 347–376.

[60] Douglas Schuler. 2009. Online civic deliberation with e-Liberate. *Online deliberation: Design, research, and practice* (2009), 293–302.

[61] Buckingham Shum and Simon Buckingham Shum. 2008. Cohere: Towards Web 2.0 Argumentation How to cite: Cohere: Towards Web 2.0 Argumentation. (2008), 97–108. http://www.irit.fr/comma08/accepted.html

[62] David Sibbet. 2002. Principles of facilitation: the purpose and potential of leading group process. (2002).

[63] Neeraj Sonalkar, Ade Mabogunje, and Larry Leifer. 2013. Developing a visual representation to characterize moment-to-moment concept generation in design teams. *International Journal of Design Creativity and Innovation* 1, 2 (4 2013),

93–108. https://doi.org/10.1080/21650349.2013.773117
[64] Jennifer Stromer-Galley. 2007. Measuring deliberation's content: A coding scheme. *Journal of public deliberation* 3, 1 (2007).
[65] Jennifer Stromer-Galley, Lauren Bryant, and Bruce Bimber. 2015. Context and Medium Matter: Expressing Disagreements Online and Face-to-Face in Political Deliberations. *Journal of Public Deliberation* 11, 1 (2015).
[66] Jennifer Stromer-Galley and Anna M. Martinson. 2009. Coherence in political computer-mediated communication: analyzing topic relevance and drift in chat. *Discourse & Communication* 3, 2 (5 2009), 195–216. https://doi.org/10.1177/1750481309102452
[67] Colleen Stuart, Laura Dabbish, Sara Kiesler, Peter Kinnaird, and Ruogu Kang. 2012. Social transparency in networked information exchange: a theoretical framework. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. 451–460.
[68] W. Ben Towne and James D. Herbsleb. 2012. Design considerations for online deliberation systems. *Journal of Information Technology & Politics* 9, 1 (2012), 97–115.
[69] Sonja Utz, Friederike Schultz, and Sandra Glocka. 2013. Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster. *Public Relations Review* 39, 1 (3 2013), 40–46. https://doi.org/10.1016/J.PUBREV.2012.09.010
[70] Richard G. Weaver and John D. Farrell. 1999. Managers as Facilitators: A Practical Guide to Getting Work Done in a Changing Workplace. (1999). http://cds.cern.ch/record/1486032
[71] Karl E. Weick. 1995. *Sensemaking in organizations*. Vol. 3. Sage.
[72] Scott Wright and John Street. 2007. Democracy, deliberation and design: the case of online discussion forums. *New media & society* 9, 5 (2007), 849–869.
[73] Fulong Wu. 2007. Re-orientation of the city plan: Strategic planning and design competition in China. *Geoforum* 38, 2 (3 2007), 379–392. https://doi.org/10.1016/J.GEOFORUM.2006.05.011
[74] Yan Qu, P.F. Wu, and Xiaoqing Wang. 2009. Online Community Response to Major Disaster: A Study of Tianya Forum in the 2008 Sichuan Earthquake. In *2009 42nd Hawaii International Conference on System Sciences*. IEEE, 1–11. https://doi.org/10.1109/HICSS.2009.330
[75] Amy X Zhang, Lea Verou, and David Karger. 2017. Wikum: Bridging Discussion Forums and Wikis Using Recursive Summarization. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 2082–2096.
[76] Haoqi Zhang, Paul André, Lydia B. Chilton, Juho Kim, Steven P. Dow, Robert C. Miller, Wendy E. Mackay, and Michel Beaudouin-Lafon. 2013. Cobi: Communitysourcing large-scale conference scheduling. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*. ACM Press, New York, New York, USA, 3011. https://doi.org/10.1145/2468356.2479597

Received April 2018; revised July 2018; accepted September 2018