Nelimarkka, Matti; Laaksonen, Salla Maaria; Semaan, Bryan

Social media is polarized, social media is polarized

Published in:
DIS 2018 - Proceedings of the 2018 Designing Interactive Systems Conference

DOI:
10.1145/3196709.3196764

Published: 08/06/2018

Please cite the original version:
Social Media Is Polarized, Social Media Is Polarized: Towards a New Design Agenda for Mitigating Polarization

Matti Nelimarkka Aalto University, HIIT
Salla-Maaria Laaksonen University of Helsinki
Bryan Semaan Syracuse University

ABSTRACT
Social media platforms have often been described as online spaces supporting political discourse. However, online discussions are often polarized; people tend to commune with those who are ideologically similar to them. The HCI response to this phenomenon has been to purposefully expose people to diverse viewpoints. This common design agenda is supported through analysis of link sharing, yet little attention has been paid to how users discuss these links. Therefore, the common design agenda may not mitigate polarization. We study the emergent discourse in 10 Finnish migration-related Facebook groups and examine how the same links are shared and discussed across anti- and pro-migration camps. Qualitative analysis of the posts and comments revealed that shared media links do not bridge polarized groups with regard to worldviews and opinions. We then demonstrate alternative design opportunities to resolve this issue and begin to develop a new design agenda to mitigate polarization.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
polarization; selective exposure; echo chambers; filter bubbles

INTRODUCTION
It is well established that in the context of online political activities, certain communicative acts invite the formation of an ad hoc public sphere – an online space where any individual is free to exchange political information and ideas with others [60]. For example, users of social media services such as Facebook and Twitter can participate by joining political groups, exchanging information (e.g., links to media articles) and opinions, or engaging in political discourse on various topics [6, 4, 61, 60]. This leads to questions about the impact of social media services in terms of their ability to facilitate information exchange and political discourse.

As evidenced by the 2016 United States (US) Presidential election or Brexit – wherein the United Kingdom where the citizens were strongly divided about the vote to leave the European Union (EU) – Western society has been plagued by increased polarization. That is, people tend to commune with those ideologically similar to them, while sharing media and engaging in discourse that predominantly supports their own perspectives. The impacts of this ideological distance include a divergence of attitudes, toward extremes, and a lack of common ground [35]. Scholars attribute the increase in polarization to online technologies, since they allow like-minded people to find each other irrespective of geographical distance, form groups [8], state opinions that get spread by their networks and hence strengthen the biases thereof [56], and algorithmically curate content in line with participants’ interests – hence potentially leaving out alternative perspectives [21]. The research community has framed this phenomenon as selective exposure [1, 37, 39], echo chambers [27, 66], and filter bubbles [50] in online spaces.

The polarization phenomenon has led researchers to explore design opportunities for balancing the use of media and, in turn, exposing people to diverse perspectives, with particular emphasis on news media [10, 40, 52]. Given how much effort is put to investigate this line of design, we call it the Common Design Agenda. However, link and network analysis [1, 27, 39], which has motivated many of the systems-design efforts to balance media, does not capture how people receive the information linked to or react to it. The affordances of social media allow users to not only share links but engage in discourse around links through reflection, agreement, and disagreement. We argue that more attention must be paid to the ways in which users react to links. This should yield more insightful perspectives on traditional link analysis and describe how people encounter links – that is, such analysis will help to create valuable insight for design of interactive systems to support the dissemination of and conversation around varied information and opinions.

To empirically investigate the Common Design Agenda, we explored linking and communication practices in Finnish Facebook groups related to immigration and the emergent refugee crisis. The refugee crisis has affected Finland, as it has the rest of Europe. Finland is a unique context for examining polarization for three critical reasons. Firstly, it is characterized by Nordic values manifested in high societal trust (e.g., high trust among citizens and between citizens and the government [9]). Secondly, few issues have polarized the Finnish public at national level historically, with immigration and asylum-seekers being rare instances of this emergent phenomenon [3, 71]. Lastly, the majority of scholarship exploring polarization has been conducted in the US political context [1, 35, 56]; this serves as an important point of departure for understanding online political communication around polarizing issues.
Following our empirical investigation, we develop alternatives to the Common Design Agenda. Our work presents different design opportunities and begins to map out a framework which can be used to mitigate polarization. Importantly, the design agenda we develop is not meant to serve as a final vision, but already demonstrates several opportunities available for designers. Thus, it questions why the focus within human-computer interaction has been on the Common Design Agenda.

In the sections that follow, we firstly outline relevant work within new media research on polarization. Following this, we present the empirical investigation challenging the Common Design Agenda. We describe our case, data, and methods in detail, then present results. Next, we discuss opportunities outside the current Common Design Agenda. Finally, we discuss the implications of our work for shifting towards a new design agenda for mitigating polarization, and we outline limitations of our study.

BACKGROUND

The Social Implications of Polarization

Polarization is increasing of ideological distance between individuals or groups [35]. It typically happens when people become divided into groups with opposing perspectives – for example, people who are “pro-choice” (i.e., who believe in a right to have an abortion) and “pro-life” (i.e., who believe abortion should be illegal). In the US, political polarization between political parties [35] and in media [56] has grown, but research regarding ordinary citizens shows more mixed results. It is not clear whether citizens are becoming more polarized or polarization is only taking place among extremists [15, 35, 56]. While more of the research has been done in the US, Europe is not an exception for the effects of polarization [12, 73, 71].

The general consensus is that polarization is harmful to society, and researchers have identified three main elements in support of this assertion. Firstly, at political party level, polarization can lead to emergent challenges for policy-making, since finding suitable alternatives that appease people across polarized groups can be difficult. Also, polarization may decrease interest and trust in political decision-making [35]. Thirdly, great disagreement on issues (and polarization) makes it difficult to accept alternative policy decisions even if such decisions have come about through a fair process [74]. Therefore, social scientists who have studied polarization have warned of the effects of polarization and the emergence of like-minded groups whose members engage only with one another. Moreover, scholars have increasingly expressed worries about the effects and social implications of the Internet and social media with regard to polarization [56].

Sunstein [66], drawing on Negroponte’s work, warned of potential negative ramifications of Internet technologies in terms of polarization. He hypothesized that Internet use would lead to increased polarization among the public as it makes it easier for people to seek information and opinions, and form camps, with like-minded others. In channeling Sunstein’s hypothesis, several scholars have examined how to mitigate polarization [52, 40, 42, 41, 60, 18, 17]. In the sections that follow, we first present empirical research on online communities and polarization, discuss the challenges and limits of this literature, and finally, review how the common design agenda is based on only a partial perspective on polarization in online communities.

Online Communities and Polarization

Studies of the effects of Internet technologies on polarization have focused heavily on network structures and linking behavior, or how users network with other users and share or consume news media in the form of hyperlinks. Among the first of these studies, Adamic and Glance’s work [1] showed how political blogs tend to link to political blogs ideologically aligned with them. For example, they found that in the US, liberal blogs linked to other liberal blogs and conservative blogs linked to other conservative blogs. This study led to the emergence of research focused on polarizing effects of social media and online news reading practices. Hence, scholars have since identified the emergence of echo chambers [28] and filter bubbles [51]. Here, the main claim is that, through the use of social and new media, people are coalescing with other like-minded people and circulating and reinforcing the same views and opinions, which, in turn, never get contested or questioned.

Other scholars have focused on the impact of algorithms and algorithmic curation for how information moves through social media and on how this can lead to polarization. Pariser [51] argues that echo chambers are reinforced by algorithms governing the flow of content in social-networking services. Recent empirical work has demonstrated the effect of algorithms in contributing to the creation of echo chambers and filter bubbles. For example, in a study of online link-sharing in the context of climate change, Itkonen [26] showed that Facebook users tend to have online friends who share their concerns and that Facebook’s algorithm often displays information that reinforces the same viewpoint. Similarly, using Twitter data from 3.8 million users, Barberá et al. [2] showed that echo-chamber effects are present particularly in connection with political issues. Overall, empirical analysis shows that people who share ideological beliefs are more likely to exchange information around political topics with like-minded others.

These observations can be seen in the common design agenda as scholars have focused on mitigating these issues by balancing news reading and breaking the echo chambers caused by polarization. However, the evidence for selective exposure – that is, of a tendency for users to consume media compatible with already held attitudes – is mixed. For example, research has found that both Twitter and Facebook users link to sites with various ideologies [39, 27, 60], whereas in political blogs people stick more to a particular ideology [37].

More recent scholarship has aimed to move beyond the classic study of polarization in linking behavior. Researchers have started to explore the ways in which polarized groups consume information, and they have found that, while groups differ in views, their consumption of media is similar. Meraz et al. [37] observed that the domains shared between two groups – which constituted 21% of all domains – were mostly
traditional media organizations. They hence suggested these media organizations to have a significant role in bridging various communities. Jacobson et al. [27] found that comments on the Facebook pages of two partisan television channels shared similar domains for over 45% of the links in the two communities. This indicates that, while media use might be polarized, people still use the same sources of information to make their arguments and claims in the comments. These findings challenge the focus of the common design agenda as this behavior might occur naturally.

Beyond Analysis of Linking: Discussing Links
Consumption of news media through social media is not purely based on link-sharing. Rather, social media platforms afford two significant features that move beyond simply sharing information. Firstly, the platforms allow for framing and expressing the links in one’s preferred way and for discussions based on a combination of those links with personal perspectives. Secondly, the links shared through social media stem from individuals’ decision to share them via social media [29]. Thus, link-sharing is a more strategic action; that is, an individual wishes to share this particular link with a particular community.

Therefore, the research on media exposure and on polarization must accommodate the affordances provided by social media services. However, the literary on online polarization does not currently cover these aspects well covered as the analysis has focused on the use of links.

For example, research on news sites and those sites’ comment systems has shown the importance of user comments. Comments are often read for gaining insights into the community’s reaction to the news [13]. Furthermore, it has been found that “user-generated comments accompanying news stories can significantly alter other participants’ beliefs about what other members of society think” [36]. Furthermore, an experimental investigation by Messing and Westwood [38] demonstrated that social endorsements alter sharing and consumption of news online. Using an experimental setting that simulated a Facebook newsfeed, they argued that, while social networking services expose people to a wider variety of news via their social contacts, social endorsements by contacts affect what one chooses to read.

However, such studies proceed from the premise that news consumption takes place in the individuals’ news feeds and do not take into account the social effect of online groups, nor do they investigate the discussions that take place in relation to the news. For example, Gilbert and colleagues, examining blog comments, found that political blogs isolate readers from dissenting opinions [20]. Therefore, the common design agenda must further take into account the social context of information processing, not just the ability to account for the balance of links shared.

How HCI has aimed to mitigate Polarization?
Since polarization is considered harmful for society, research in HCI and related disciplines has examined how to mitigate it through design endeavors [52, 40, 42, 41]. While there is research indicating that echo chambers are not always a result of underlying technology design decisions or algorithmic bias but rather on group processes [46, 24, 11], the work has been focused on “nudging” the users and applying behavioral change techniques.

The NewsCube [52] study showed readers several aspects of a given story by collecting various news items and automatically extracting aspects of them. It also helped readers see how stories they had already read had different aspects and thereby informed readers about the existence of media bias. It was shown that the design led readers to read more stories and, therefore, broaden their perspectives. The NewsCube design has also fostered more extensive system development, including a browser widget with similar goals [10].

The Balancer system [40] was designed to visualize the overall balance of media consumption—the number of liberal and conservative Web sites the user had visited. The system was able to convince participants to visit other sites to some degree, but the authors warned about over-generalizing from these findings, because of the biased sample of study participants.

Furthermore, Munson et al. [41, 42] examined methods to expose additional viewpoints in news aggregation, particularly methods to show content the user does not agree with in the interface. They showed that those with a preference for diversity-seeking preferred systems that showed 60% agreeable content, whereas non-diversity-seeking users favored situations featuring only agreeable content most. Furthermore, they showed that the alternative methods for presenting non-agreeable content, sorting the content such that agreeable content is first or highlighting agreeable content, did not increase satisfaction. Therefore, they concluded that there are ways to improve exposure but that there is a cost in user satisfaction among people who do not seek diverse opinions.

Beyond changes to media consumption patterns, only a few proposals have been made and they seem to also drive from the analysis on lack of different perspectives. Semaan et al. [59, 60] showed that some people are interested in using social media for discussion beyond their own camps. In particular, they studied users who used multiple social media sites to reach different audiences and seek other opinions. They concluded that such uses should be further supported via development of systems that facilitate finding discussion from several social media systems. Similarly, polarization-aware recommendation systems have been discussed. The goal with such work is to develop recommendations as to how to diversify a user’s network [18, 17]. The aim is to recommend that two people, with different opinions, discuss their perspectives, choosing these people from only mildly polarized populations.

The tools on the common design agenda have been focused on providing more information to users. However, these efforts have not accounted for the social interaction and framing carried out in the online communities. Those focusing on social interaction, meanwhile, have aimed to increase interaction between the polarized lines. However, the Common Design Agenda has not addressed the question of how to decrease
We have chosen to study Finland in response to the extensive Agenda. Based on the critique presented above, we draw on an empirical case to answer two questions. First, to establish polarization we examine the link sharing practices in politically opposing Facebook groups and study

Research question 1: To what degree cross-camp link sharing occurs naturally and what type of content is common between these groups?

Following this and to examine the sustainability of the Common Design Agenda we analyze comment threads on these links. The idea of Common Design Agenda is to expose people to news stories and information sources outside their own perspective. This can be seen through Facebook posts with same link shared to both camps and the discussion emerging in those posts. Therefore, we sought to answer this open-ended question:

Research question 2: How is the discussion on the common links different from links shared only within the camp?

This research setup mimics closely the gist of the Common Design Agenda: ensuring that people read the perspectives of opposing camp and reflect on it. This allows us to gauge the common design agenda in a naturalistic settings, seeing what might take place when both camps are proposed similar links. To our knowledge, the analysis on emergent discourse on polarized topics that surrounds links on social media platforms is a novel methodological perspective. The novel methodological perspective provides insights into problems of the common design agenda in a natural setting and therefore, make a contribution to the ongoing debate on online polarization.

The Research Setting: The Finnish Context

We have chosen to study Finland in response to the extensive research conducted in the US. The political context for our study is one of emerging polarization. The topic of migration has been less prominent in Finnish society and, therefore, has less historical context that could enter in than does the United States with its marked polarization across party lines. It has been argued that before 2015, there was little political polarization in Finnish society [3, 71]. Furthermore, the topic examined has high societal importance both in Finland and in Europe at large.

Since 2015, Finland has, not unlike other European countries, experienced a rise in the number of forced migrants, or refugees, from the Middle East and Africa who travel across the continent and often seek asylum or permanent residence [58]. In total, EU member states received over 1.2 million asylum applications in 2015 alone, which represents a twofold increase from the previous doubling, in 2014 [14]. According to the Finnish Immigration Authority, Finland alone received almost 32,500 asylum-seekers in 2015, up from 3,651 in 2014. The rising number of refugees seeking asylum has prompted much media coverage, citizen activities, demonstrations, and debate on various platforms (including social media, especially in the Facebook service). Therefore, the Finnish public has mobilized both in support of and in resistance to the influx of refugees to the country.

In its politics, Finland is a multi-party nation with a total of nine parties in the parliament. As a Nordic state, Finland has historically been characterized by a lack of contentious politics whereby few issues have polarized the public. Today, there exist only four political issues that serve to divide the public: support for traditional values, cultural and social diversity, economic freedom, and economic equality [72]. The economic downturn in the late 2000s led to long-term recession. However, these issues do not typically form strong dividing lines, since citizens can at the same time support economic freedom and economic equality [72, 48, 65, 64].

In this country with high societal trust [9], the most frequently used sources for political information are television news (considered a highly important or important politics-information source by 58% of citizens) and the newspaper (44%). The major news sources aim to provide high-quality journalistic products and have agreed to follow ethics guidelines for journalism. Strongly biased media sites are not part of Finland’s mainstream media, nor do the main media outlets openly espouse any political stances. Social media channels serve as a major venue for political information for approximately 15% of the citizens [23]. The Internet is used on a daily basis by 79% of the population, 74% of people use it to read news, 47% use social media services daily, and 11% use this channel for political or civic purposes. Social media use is somewhat biased towards people under 55, the more educated, and people in urban areas [63]. This level of use makes the study of social media and polarization sustainable endeavour.

Community Selection and Collection of Data

To explore the relationships between polarization, link-sharing, and link-commenting behavior, we compiled a dataset from 10 Finnish Facebook communities. This decision reflects the online communication practices of the Finnish population: Facebook is much more used than Twitter in Finland [57]. Within Facebook, there are two types of online communities: groups and pages. The Facebook groups create a “news feed” for only that group, in which all group members can post and comment on content. Facebook groups can be publicly open or hidden from the public; for reasons of research ethics, we focus on public groups only. The Facebook pages are “profile pages” for a particular cause or for an organizational actor. While pages are driven more by the administration team, the groups are community-driven.

The 10 communities we selected for analysis were the largest ones on expert-sourced lists of pro-immigration and anti-immigration communities, five for each “camp” (see Table 1) (we will henceforth use the term “camp” to refer to a set of
We used a random sample of posts that included a link common to all communities for analysis. Of the posts with no common URL, we selected those identified were classified. In total, this method captures 22% of all domains and 97% of the links present in the dataset, because the link frequencies follow a power law. We consider this coverage sufficient for further analysis. During the coding, when in doubt about the content, we checked the site via Wikipedia, Web of Trust (https://www.mywot.com/), and Media Bias Fact Check (https://mediabiasfactcheck.com/) to help assign the site to a category. Two researchers independently conducted analysis of the whole dataset to ensure the validity of the classification. Disagreements were discussed and agreed upon case-specifically.

We answered RQ1 by examining the differences in link-sharing in our data. We examined the common links (links featured in both camps) in comparison to links shared only within a single camp. For both of these questions, we first conducted a $\chi^2$ test to examine whether differences exist in the distribution of news sources. To examine detailed analysis on what causes these differences, we conducted further analysis with a two-sample (proportional) z-test with Bonferroni correction to account for the large number of comparisons made.

To understand the discussion (RQ2), we classified contributions in line with the relationship to the URL and the type of statements made (see Table 3). The type-of-statements classification was adapted from the literature [62] but expanded to separate out also factual claims based on referenced facts, anecdotal evidence, personal experience, and threat scenarios (marked in the table with *). We tested and developed the categorization via two of the authors classifying a sample set of data simultaneously. Again, on account of the large total number of posts, we could not analyze all the posts. Instead, we used a random sample of posts that included a link common to the two camps and another random sample of posts with camp-specific URLs. Of the posts that had a common URL, we selected 2% of the content (961 posts or comments) for analysis. Of the posts with no common URL, we selected at random, in total, 559 posts or comments for analysis. For the relationship to the URL, we achieved a low but acceptable value in reliability measurements (κ = 0.58, agreement 72%). For statement types, we performed comparison for each statement separately (since the choices there were non-exclusive) and achieved κ values of 0.55–0.67 and agreement between 86% and 95%. Since κ is considered a conservative measurement, it has been argued that even κ > 0.4 can be applied [68]. Values around κ = 0.6 indicate moderate to substantial agreement [34], and we argue that these values are sufficient to warrant further analysis.

Our strategy for answering RQ2 followed the same analysis approach as that for RQ1: we examined differences in the communities by utilizing $\chi^2$ testing and then a two-sample z-test. Our further analysis focuses on links common between the camps; motivated by RQ1, we were interested in examining whether opinions on these links diverge or converge. To ascertain this, we determined the position taken most often on the link in the pro-immigration and in the anti-immigration camp. Discussion of the link on both camps was aggregated only to the majority position. Next, we compared the types of

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
<th>Total posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Anti-immigration movement that aims to close borders to asylum-seekers. Born in response to the refugee crisis in fall 2015</td>
<td>25,584</td>
</tr>
<tr>
<td>A2</td>
<td>Conservative group that promotes Finland as being a place for Finns only.</td>
<td>14,576</td>
</tr>
<tr>
<td>A3</td>
<td>Anti-immigration movement aiming to achieve the status of registered political party. Born in response to the refugee crisis in fall 2015.</td>
<td>5,231</td>
</tr>
<tr>
<td>A4</td>
<td>Anti-immigrant street-patrol group founded in October 2015 and registered as an association in January 2016.</td>
<td>1,478</td>
</tr>
<tr>
<td>A5</td>
<td>Most pro-immigration-focused sub-group of A3.</td>
<td>1,411</td>
</tr>
<tr>
<td>P1</td>
<td>Network-like association that works against racism and xenophobia.</td>
<td>19,867</td>
</tr>
<tr>
<td>P2</td>
<td>Social movement organizing refugee-related activism and concrete help, from accommodation to volunteer translation.</td>
<td>4,833</td>
</tr>
<tr>
<td>P3</td>
<td>Facebook group promoting an international Finland and opposing to racism, led by a non-profit organization. Founded in 2010</td>
<td>5,063</td>
</tr>
<tr>
<td>P4</td>
<td>Facebook group that objects to hate speech.</td>
<td>1,561</td>
</tr>
<tr>
<td>P5</td>
<td>Facebook group that objects to racism in Finland.</td>
<td>1,580</td>
</tr>
</tbody>
</table>

Table 1: Descriptions of the communities studied

The dataset is composed of 73,041 posts within these 10 communities, spanning April 2010 to February 2017. Each post has accompanying details such as the “likes” and other reactions it garnered, all comments, comments on comments, and the “likes” for those, alongside the text content and links in the original post. The data were collected from the Facebook API by means of a custom-built tool. With regard to research ethics, we acknowledge that not all that is public on the Web should be made available for research [5], particularly in studying a politically heated topic such as immigration. Therefore, we do not analyze any individuals in the data.

We used a mixed-methods approach encompassing both qualitative and quantitative elements. Our research is data-driven research [31] since we did not test given hypotheses, instead changing our analysis approach between the use of explanatory and confirmatory tools during the research process. Overall, we conducted classification of the data in line with three main criteria, related to link types (RQ1), characteristics of the message (RQ2), and the relationship between the message and the link (RQ2).

RQ1 required classification of the links shared. Firstly, we conducted open coding on a sample of the domain content to produce coding categories (see Table 2). We then compared our proposed categories to those used in previous work [37] and aligned our categories with those when this was necessary and possible. After this, we proceeded to code a larger set of domains. Since this stage required extensive manual work, we chose to classify only all domains that were used by both camps ($n = 539$) and the half of the domains with the highest number of links ($n = 244$) (five or more links). After this, the rest of the domains were skimmed by the researchers and

1See https://github.com/HIIT/hybra-someloader.
We opted to use manual content classification instead of computational analysis for two main reasons. Firstly, there are no validated computational tools for the Finnish language. We acknowledge that many elements we explored in the qualitative phase could have been examined automatically in other languages, such as English (for which sentiment-analysis tools exist). Furthermore, we conclude that, from this perspective, the media environment is highly fragmented. The total number of shared links in common was small (1.9%, \( n = 852 \)). Both camps were equally active in linking to common sources.

Our analysis revealed that only a few links were common between the camps, and most of those were to user-generated content and unclassified sources, followed by items from national media. This distribution differed greatly from that for links shared only within one camp. Instead, the camps shared different news stories even in an environment as small as Finland.

Findings
From the coding categories (see Table 2) we identified all links in all the posts in the pro-immigration and anti-immigration communities separately, and we extracted the domains used in both camps. As Figure 1 shows, many of the domains are unique to one particular group only. In fact, 74% of domains were not used beyond a single group, let alone between the camps. This supports the core idea of the Common Design Agenda: the groups, even within the same camp, have different media linking behaviors.

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media</td>
<td>Sites from journalistic organizations that have salaried staff. The Web site is often complementary to other forms of distributing news, such as a newspaper or broadcast media (e.g., radio), but this is not mandatory. This corresponds to Facebook’s “mainstream media” [27].</td>
<td>Nokia.com, Apple.com</td>
</tr>
<tr>
<td></td>
<td>Non-governmental organizations: Sites run by non-governmental or non-profit organizations. This corresponds to Facebook’s “NGO” [27] with the exception of including political organizations also.</td>
<td>Save the Children Finland, Finnish Red Cross</td>
</tr>
<tr>
<td></td>
<td>Governmental organizations: Any local, national, international, or multilevel government-operated organizations. This corresponds to Facebook’s “governmental organization” [27].</td>
<td>National Coalition Party, Republican Party</td>
</tr>
<tr>
<td></td>
<td>Person Company</td>
<td>Finnish government organizations operating mainly within Finland and affiliated party organizations.</td>
</tr>
<tr>
<td></td>
<td>Organization</td>
<td>International organization Non-profit organizations operating mainly outside Finland and international political party. Political parties in some other country.</td>
</tr>
<tr>
<td></td>
<td>Non-governmental organization</td>
<td>Finnish social movements Social movements and non-based, “organizations” not organized as well as the non-governmental organizations. This class also involves causes, often related to the topic of migration.</td>
</tr>
<tr>
<td></td>
<td>Governmental organization</td>
<td>International governmental organization Organizations similar to Finnish governmental organizations but run by countries other than Finland.</td>
</tr>
<tr>
<td></td>
<td>Unclassifiable</td>
<td>Governmental organizations operating in Finland or providing the Finnish government with support services to combat cyber crime.</td>
</tr>
<tr>
<td></td>
<td>News-like site</td>
<td>Sites that publish news-like content in a format that resembles media but where the factuality of the content is questionable. This includes clickbait sites, media sites that provide non-factual information, “alternative media” sites, and “fake news” sites. It partially corresponds to “informational” [27].</td>
</tr>
<tr>
<td></td>
<td>User-generated content</td>
<td>Sites that primarily display content produced by their users. This corresponds to Facebook’s “forum/social network” and “portal” [27].</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Information-presenting sites that cannot be classified in any other category at domain level.</td>
</tr>
</tbody>
</table>

Table 2: Coding categories for domains

Figure 1: Selected communities and domains where they link RQ1: Common Links Are Different from Single Camp Links
There are differences also between common links and what was shared only within a camp (see Table 4, where significant results are marked with \( \oplus \) and \( \ominus \)). In both communities, the links shared were more often to user-generated content and unclassified sources. They were less often from media sources, news-like sites, Finnish GOs, NGOs, SMs and parties, persons, companies, or organizations. Therefore, the national media, even while being the major source of information in both camps (accounting for 39% of all links examined), did not “connect” these camps. Instead, the camps shared different news stories even in an environment as small as Finland.

Table 3: Coding categories for messages

<table>
<thead>
<tr>
<th>Type of statement</th>
<th>Author refers to</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Author offers practical solutions to problems raised during the discussion.</td>
<td>&quot;Proposition: Author offers practical solutions to problems raised during the discussion.&quot;</td>
</tr>
<tr>
<td>Negative</td>
<td>Author presents a statement and supports it with personal experiences.</td>
<td>&quot;Personal experience: Author presents a statement and supports it with personal experiences.&quot;</td>
</tr>
<tr>
<td>Neutral</td>
<td>Author makes a call to action and mobilization to support the cause.</td>
<td>&quot;Mobilization: Author makes a call to action and mobilization to support the cause.&quot;</td>
</tr>
</tbody>
</table>

* Examples of content by category shown next page.
We will next examine the discourse that takes place in the anti- and the pro-immigration camp. The overall odds ratio (1.2) is not enormous, owing to the opinion-driven nature of both discussion styles. There are further differences in the type of statements made, with anecdotal and sourced statements being used less often in the non-common link and common link (%).

Table 4: Links separated into common (shared in both camps) and shared only in the “pro” or “anti” camp (%).

<table>
<thead>
<tr>
<th>Category</th>
<th>Common (%)</th>
<th>Non-common (%)</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>National media</td>
<td>28.82</td>
<td>36.77</td>
<td>9.78</td>
</tr>
<tr>
<td>Local media</td>
<td>5.10</td>
<td>1.60</td>
<td>0.10</td>
</tr>
<tr>
<td>Organization media</td>
<td>0.17</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Political party media</td>
<td>3.11</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>International media</td>
<td>2.71</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>News-like site</td>
<td>1.34</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Finnish NGO</td>
<td>0.42</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Finnish SM</td>
<td>0.91</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>Finnish party</td>
<td>0.43</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>International NGO</td>
<td>0.30</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>International SM</td>
<td>0.20</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>International party</td>
<td>0.01</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Finnish GO</td>
<td>0.61</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Finnish GO</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>International GO</td>
<td>0.10</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Multinational GO</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Person</td>
<td>2.89</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Company</td>
<td>0.31</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td>0.10</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>UGC</td>
<td>22.68</td>
<td>50.86</td>
<td></td>
</tr>
<tr>
<td>Unclassified</td>
<td>5.29</td>
<td>34.35</td>
<td></td>
</tr>
<tr>
<td>Unclassified</td>
<td>100.01</td>
<td>100.00</td>
<td>100.02</td>
</tr>
</tbody>
</table>

Non-governmental organization, NG = social movement, GO = governmental organization, UGC = user-generated content.

Since each message could belong to more than one class, the figures do not add up to 100%. A significant difference of \( \chi^2 \) posts and comments with a common URL from posts and comments with a non-common URL at \( p < 0.01 \) with Bonferroni corrections.

Table 5: Classification of messages in conversations with a non-common link and common link (%).

<table>
<thead>
<tr>
<th>Category</th>
<th>Common (%)</th>
<th>Not common (%)</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposition</td>
<td>6.0</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>Opinion</td>
<td>61.3</td>
<td>49.4</td>
<td>1.2</td>
</tr>
<tr>
<td>Mobilization</td>
<td>0.3</td>
<td>6.9</td>
<td>19</td>
</tr>
<tr>
<td>Question</td>
<td>11.3</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td>Anecdotal statement</td>
<td>15.7</td>
<td>19.4</td>
<td></td>
</tr>
<tr>
<td>Personal experience</td>
<td>3.6</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>Sourced statement</td>
<td>3.6</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Threat</td>
<td>1.6</td>
<td>4.2</td>
<td></td>
</tr>
</tbody>
</table>

Phew, hell!!!!!! But it’s good if kids are taken away from people like that... and the women also... I pity them!!!!

OMG, look the imagination and propaganda they have. Those who speak shit like that or publish it must be idiots.

These opinions could be backed up in various ways. We extended our analysis to separate anecdotal statements – statements presented as facts but with a source or other evidence to back up the arguments – from statements grounded in personal experience or referenced facts. The anecdotal experience was the most commonly used strategy to support the claims made in the text.

I’ve heard that in Spain there are areas where you can manage with the Finnish language only.

People who walk like this do have a severe identity problem or at least an unconscious inferiority complex.

Sourced statements were presented with links to provide additional support for the claims made. Sometimes these also included quoting a source, such as the law, directly inline in the comment.

As far as I understand, there are no (genetic) races within the human species, but it is a matter of social construction; i.e., biologically we are all “carved from the same tree.” [link]

Another type of justification relied on participants’ first-hand experience. Unlike in anecdotal evidence, commenters here clearly articulated how the evidence cited was collected, even though anecdotal.

I didn’t see any [pejorative for immigrants] in the care home.

We are particularly interested in these types of justification that participants gave during the discussion. The great extent of anecdotal statements highlights the lack of more concrete rationale for the cause. Rather, it often seemed to present a particular reality felt by the posters that, because of its anecdotal nature, may be hard to challenge.

The rest of the categories (propositions, questions, mobilization, and threat scenarios) were present in only a few posts and could not be directly compared to other categories and across groups. Therefore, we do not go into these in this brief elaboration on our observations.

RQ2: Common Links Do Not Create a Common Ground

In light of the opinion-driven nature of the links, it is interesting to examine how the comments reacted to common links. Therefore, we analyzed in detail the common links and classi-
We then summarized the thread on the basis of its largest count: we found a single case wherein both camps supported content equally: if two views were equally prominent, a value of 0.5 was used and when all three views were “tied,” 0.3 was used. Table 6 provides a summary of these findings; for example, we found a single case wherein both camps supported content justified by the same link.

The main finding, however, is that if one of the camps supported a link, the other camp disagreed with it. In Table 6, of the 16 cases examined, 11 showed this pattern (while one discussion had equal amounts of support, neutrality, and opposition) (see the cells marked with ⋆). This indicates that mere exposure to the other camp’s perspective was not sufficient to adjust the opinion of the majority of participants to be more in line with the other participants. If this was the case, we would have seen more cases where both camps support or oppose the link.

The establishment of a common ground and appreciation to other perspectives and the other camp through shared news sources has however been one of the driving motivations behind the Common Design Agenda. Based on our qualitative analysis in RQ2, this is not supported by this empirical case. First, the discussion on the common links was more opinionated and often had hostility towards the other groups. Second, the common links did not create a shared understanding between these polarized groups, but rather lead to display of disagreement. Therefore, alternative design opportunities can be developed to mitigate polarization.

### CHALLENGING THE COMMON DESIGN AGENDA:

#### RETHINKING THE DESIGN OPPORTUNITIES

The common design agenda has been to propose “more balanced” news-reading among users or support social interaction over the camps. Our empirical observations, especially the observations on comments, showed that the common links do not bridge between people with opposing perspectives. They did not only fail to create common ground but often lead to display of high animosity towards the other camp. Therefore, designing socio-technical systems to decrease polarization must extend beyond balanced-information recommendations.

In this part of the work, we begin mapping the design opportunities for interventions which may decrease polarization. The aim here is to offer not fully formed design implications and the proposed opportunities do not relate to the empirical findings in the previous section. Rather, they aim to demonstrate that several alternative approaches can be used to address political polarization. This section should, therefore, be considered more as the initiation of a design framework to mitigate polarization, displaying that system design can go beyond balanced news-reading or supporting cross-camp social illustration.

### Re-implementation of Common Design Agenda

Several systems developed focused on recommending “more balanced” news-reading [52, 10, 40]. The form of these interventions could be redeveloped in various ways, as elaborated below and seen on sketches in Figure 2. These ideas build upon the Common Design Agenda and its traditional implementation (see Figure Figure 2a). We have varied the format of the interventions to show the potential opportunities to reconsider how this intervention is done. Furthermore, we envision that these interventions – like the Common Design Agenda interventions – must be implemented so that no extra effort is required from the user to be exposed to them.

#### Aiming the intervention at the consumer or the poster

Balanced news-reading systems are aimed at influencing the user’s behavior during the consumption of media. However, the consumption phase in social media already includes framing for the given link. Therefore, suggesting alternative perspectives on a particular link will hardly break the framing supplied for it. When comparing discussion of the same stories between communities, we observed that discussion was uncivil and often questioned the opposing perspectives directly. Furthermore, we discussed that social media activity can be seen as strategic action, including the framing of shared news. Therefore, to mitigate polarization, also the poster who frames the link can be influenced. For example, the link-poster could be prompted with alternative views and news sites, thereby being given an opportunity to frame the story in a more nuanced manner (see Figure 2b). This is different from providing the reader with additional sources of information, which follows the Common Design Agenda (see Figure 2a).

#### The form of interaction: algorithmic or human-like

The Common Design Agenda provides algorithmic intervention wherein the system aims to suggest alternative sources to users (see Figure 2a). For our case, we assume that the common links were not recommended by a system but picked up by users. Therefore, the users had already taken steps to break selective exposure, but the reactions were rather uncivil and even antisocial. We suggest various mechanisms to humanize the opposing camp through design practices should be investigated as tools to mitigate the most severe forms of targeted hate speech. One way to approach this could be to leverage the weak ties and show someone – a friend, relative, or friend of a friend – who has an opposite view, then summarize that person’s perspective on the current issue. Naturally, this should be done in a manner respecting the platform’s privacy settings and ensuring that the person used does not experience harm due to participation in this intervention. The intervention would give a face to, or humanize, “the other” (see Figure 2c).

#### Intervention content: links or discussions

We consistently observed that the mere alternative perspectives provided through the links did not change posters’ or commenters’ ideological positions, which follows the Common Design Agenda (see Figure 2a). Intervention could focus also on showing emerging...
require facilitation, such as prompts which invite participants to consider different perspectives.

Furthermore, the spiral of silence-theory proposes that the people can avoid confrontation and not present perspectives which they believe are different from group majority viewpoints [44, 19, 70]. However, it may be that even within a single camp, opinions differ but are not expressed. Anonymity has been shown to address these kind of social concerns well (e.g., [45, 43]), but its negative implications for online discussion have also been established (e.g. [30, 49, 54, 55]). The discourse architecture can be developed in a way which allows the emergence of different viewpoints, for example allow expression of dislike of the post in an anonymous manner to make it visible to the group, allow building support to confrontation perspective within a smaller subgroup of each group, or present the estimated group level aggregate of support to particular perspectives based on latent features.

To conclude, we aim not to provide extensively designed insights but, rather, to suggest alternative ways to mitigate polarization and show that the community can do more than, following a core premise of previous work, i.e., merely give the reader an alternative set of news sites. Here, we must acknowledge that these designs present potentially sensitive data, positions on “hot” ideological topics, to people whose beliefs differ strongly. The privacy settings and the degree of political openness could be used to gauge which users might participate in such interventions. Furthermore, on a social media platform where commenting can take place, it is possible that these interventions would invite an attack on anyone that the system shows to disagree with the targeted individual. This means that the intervention should not allow direct interaction if it could be expected to create conflicts.

After these considerations, we remind that our aim with this section was to show, both in terms of rethinking the intervention or moving beyond and focus on the discourse architectures, that the there are several opportunities to aim mitigating polarization. While this is the case, the systems presented in human-computer interaction scholarship have focused on the intervention shown in Figure 2a. This indicates stagnation in terms of the design work which may limit the impact our community has to mitigate polarization.

**DISCUSSION**

Our work was motivated by a desire to rethink the design agenda for mitigating polarization in online social media. The common design agenda has involved increasing access to varied sources of information [52, 10, 40, 41, 42]. This approach has been prompted by link-sharing network analysis, which has consistently shown polarization taking place (e.g., in news sources between camps). Before engaging with the design agenda further, we summarize the findings presented above.

Firstly, we conducted an analysis not only of the domains but also of the media sources used by the pro- and anti-immigration camp via Facebook. We observed that the “pro” camp linked more actively to governmental organizations, social movements, and Finnish non-governmental organizations, and also companies, organizations, international media, and organization-media sites. Instead, the “anti” camp linked more
to news-like sites and sites where the actual content is unclear from the domain (as with Dropbox content). These findings demonstrate that differences exist not only in the sites used—as is well established—but also in the types of sites used.

To examine the common design agenda, we restricted our analysis to links shared in both camps. Only 2% of the full sample featured common links of this sort.\(^2\) We showed that the common links were different from links shared in either of the communities separately. The common links were more often to user-generated content (such as meme sites) or unclassified content. Our qualitative analysis revealed that the discussion threads on common links were more opinion-driven than discussions taking place on threads with a link shared only within a camp. Furthermore, we showed that the conversation on common links diverged: when one camp supported the goals behind the link, the other camp opposed them. Often the threads also led to antisocial behavior, attacks, and distrust for the other camp. Similar findings on the impact of biases have been made recently, in such contexts as raw-information processing when polarization is expected [47, 69]. The research indicates that personal biases are extremely strong and can distort decision-making.

Following we presented ways to reconsider the common design agenda to avoid stagnation in this critical area. We highlight that these should not be considered as fully articulated design ideas, but rather aim to show design opportunities which can in the future work be developed towards a design framework. Due to this, the design work does not “emerge” from the empirical findings. The empirical findings have informed us about challenges such as anti-social behavior and even de-humanization (and therefore, leads to proposals of human-like or discussion focused interventions), the goal with the design phase was to show that even minor changes to common design agenda lead to different forms of interventions. This is to demonstrate the stagnation on the common design agenda and make a call for reconsiderations.

**Future Work and Limitations**

Our work has been able to describe the polarization phenomenon and, through the qualitative work, provide critical insights for examining political communication via social media. However, that work is descriptive in nature and does not address the mechanisms that cause this behavior. For example, it has been shown repeatedly that between camps even reading scientific results becomes biased [69]. Similarly, the mechanisms may be related to in-group and out-group classification phenomena, wherein members of one group must demonstrate that they differ from the other group [7]. Understanding the mechanisms would allow researchers to further examine design interventions. However, this cannot be achieved via an observational study such as ours and demands, for example, interviews. For instance, we could follow the approach that Semaan and colleagues [60] take, to understand what motivates participation.

\(^2\) Regrettably, this number cannot be compared to international studies’, since those have operated at a domain and not an URL level. The domain-level overlap is 15.5%, which seems to be on par with Jacobson’s study [27] but is significantly lower than that in the elite-blog study by Meraz [37]. Hence, the case we examine seems not to be an outlier but supported by international studies.

To develop the reconsidered design agenda further, experimental work must be used to investigate whether the suggested design opportunities would actually yield the intended effects, and with what parameters. With this type of work, the design proposals can be developed to a comprehensive design framework.

The findings are limited also by the scale of the data and context of the study. A simple suggestion would be to apply automated tools for data analysis. However, there have been recent concerns about the validity of automated tools to analyze politeness [25]. Therefore, we have chosen not to conduct this as using manual coding with acknowledging its limitations. Furthermore, the sample was limited for reason of concerns about the researchers’ mental wellbeing. The total number of posts and comments was rather small (1,520), and the analysis of support and opposition pairs by URL (see Table 6) included only 16 unique URL pairs (each having two or more discussion threads, with a total of 961 posts and comments). As with our research subjects, we aimed to limit exposure to harmful content on the researchers’ part, and, with sufficient demonstration speaking to the results, we did not extend the coding beyond the 2% sample of common-link posts and comments.

Finally, the empirical case was clearly an extreme context of polarization. It may be that discussions with moderate participants may have lead to less anti social behavior and development of the common ground. The case, while being able to demonstrate potential harms of naive recommendation systems, does not generalize across all potential discussions.

**CONCLUSION**

Researchers have proposed various content and person recommendation systems to mitigate effects of polarization and echo chambers. We aim to challenge this common design agenda both empirically and through design proposals. We have empirically examined the common design agenda by analyzing the links used in pro-immigration and anti-immigration communities and the discussions that emerged upon the posting of those links. We observed that on the domain level both camps most commonly used various news-media sources but overlap at the level of individual links were small. The overlap – links used by both pro-migration and anti-migration posters – is formed mainly of user-generated content, such as material on image-sharing sites and various social media platforms. We then turned to how links were discussed. We observed that much of the discussion was based on opinions and supported by anecdotal examples. We found that people had rather strong perspectives on the links and that they framed and discussed the links accordingly. Following this, we question the success of the common design agenda focused on recommendation systems. We demonstrate that the intervention can be designed in several ways and some of the proposals of which may succeed better to avoid the emergence of anti-social behavior – more empirical work on this is required. Furthermore, the recommendation systems are not the only way polarization can be mitigated and we developed proposals to engage with a reconsideration of the design architecture. Based on the empirical work and the presentation of design proposals, we call for a reconsideration of the common design agenda.
REFERENCES


69. Anthony N. Washburn and Linda J. Skitka. 0. Science Denial Across the Political Divide. Social Psychological and Personality Science 0, 0 (0), 1948550617731500. DOI: http://dx.doi.org/10.1177/1948550617731500


