Sabine Ylönen
University of Jyväskylä

Social media as the fifth estate: Strategic organisation of German and Finnish counterspeech campaigns on Facebook

Social media, also called the fifth estate, form a new control body and may thus fulfil democratic functions. On the other hand, social media have enabled the proliferation of hateful debates in which, in the name of freedom of speech, previously tabooed content and language are becoming de-tabooed and accepted. As a response to fake news, direct and indirect hate speech, however, forms of organised counterspeech have emerged that counter the normalisation of aggressive and hateful speech. To influence the discourse in the comment sections of social media in terms of the fifth estate, counterspeech has to be visible also quantitatively. In this contrastive study, I analyse the activities of the German and Finnish Facebook groups of the network *#iamhere international*. Special attention goes to how the groups organise their activities strategically and how much the conventionalisation of digital genres influences the intensity and continuity of their activities.

Keywords: counterspeech, social media, fifth estate, strategic organisation of activities
Asiasanat: vastapuhe, sosiaalinen media, viides valta, aktiviteettien strateginen organisaatio

1   Social media as the fifth estate

Social media are central arenas of private and public opinion formation (Saresma 2017: 219). They allow users to connect via digital platforms at high speed and with wide outreach, thus creating a new digital public sphere. This digital public sphere has become a new control body that is also called the fifth estate. As Bunz puts it, "more eyes see more", which is seen as an advantage because it counteracts the risk of too benevolent reporting on politicians, that is, the tendency for the political public sphere and the media public to become increasingly congruent (Bunz 2012: 165, 166). She concludes that social media serve a democratic function as the fifth estate:

> A new separation of powers also serves a **democratic function**. So let's welcome a second, digital public, operated by the **nimble fingers of the citizens**, who, armed with keyboards, computers and supported by algorithms, have stepped alongside the first public – as the **fifth estate**. (Bunz 2012: 166, my translation and emphasis)

Social media increasingly influence public discourses. As Burkhard (2015) puts it, journalism's monopoly on discourse in the mass media is changing as a result of (lay) communicators in the heterogeneous public spheres of the social web. As examples he mentions the resignation of prominent German politicians (minister of defence Karl-Theodor zu Guttenberg resigned from all political positions in March 2011, Federal President Christian Wulff resigned in February 2012) preluded by social media initiatives revealing plagiarism and corruption, respectively. Public discourses are increasingly influenced by social media as the fifth estate. Journalists gladly pick up high attention-receiving and especially scandalising discussions in social media. In public alliance, social and journalistic media are able to turn up the scandalous spiral of public indignation that led to the resignation of prominent German politicians, as mentioned above (Burkhard 2015: 122). The power of social media also lies in the possibility to serve as direct communication channels for politicians who, consequently, are no longer entirely dependent on journalistic media reporting. Everybody can engage in social media discussions, insofar as there are no control mechanisms comparable to those in journalistic media. Undoubtedly, social media have the potential to promote participation and democracy.

On the other side, social media also constitute a risk for democratic societies, as the spread of hateful discussion cultures and disinformation has shown (Dumbrava 2021; Heinrich Böll Stiftung 2021; Knuutila et al. 2019; Schwarz & Holnburger 2018: 35; Wesley & Gradon 2020). The best example for the power of fake news and conspiracy theories we saw on 6 January 2021 when a violent mob, inspired by President Donald Trump's claims of electoral fraud, attacked the US Capitol in Washington, D.C., where Congress was in session to formalise Joe Biden's victory in the 2020 presidential election. Even more worrying is that "social media can also give ordinary citizens the power to generate false and inaccurate information" (Mejias & Vokuev

2017: 1028), such as by consuming and disseminating Russia's ongoing propaganda in the conflict with Ukraine, which on 24 February 2022 resulted in the "special military operation" to "demilitarise and denazify" Ukraine.

In the name of freedom of speech, previously tabooed topics and language use are de-tabooed and accepted. In this connection, Wodak (2021) uses the expression "shameless normalisation" (see also Wodak et al. 2021). Zweig et al. (2017) and Montag (2018) have suggested that, based on their algorithms, social media favour the formation of filter bubbles in which users are offered content that is adapted to their search history and likes, and which can give the wrong impression of the prevailing mood in society. In addition, they see that social media ease the organisation of like-minded people in so-called echo chambers and enable the distribution of de-tabooed content through digital networks. Because right-wing populistic politicians and their followers use social media over-proportionally (Lucht et al. 2017; Knuutila 2019: 4), one could get the impression that their views are shared by the majority of the citizens. In addition, troll farms and social bots may influence elections and votes, as shown in the presidential elections in the USA or in voting for Brexit, both conducted in 2016 (Egli & Rechsteiner 2017: 250; Kreißel et al. 2018).

To counter the normalisation of aggressive and hateful speech, forms of organised counterspeech have been developing in social media. Laubenstein and Urban (2018: 55) point out that "only a few organisations appear to have created an online strategy or addressed it significantly on social media". The development of online strategies and the presence of counterspeech campaigns naturally depend on what resources the action groups (whose members volunteer) and especially their administrators and moderators have.

To influence the discourse in the comment sections of social media, counterspeech has to be **visible** also **quantitatively**. The aim of this contrastive study is to analyse the strategic organisation of counterspeech in the German and Finnish Facebook groups of the network *#iamhere international*, in order to measure its impact on the visibility of group activities.

2 Counterspeech

The aim of the counterspeech groups in the network *#iamhere international* is to influence public discourse in factful and empathetic ways so that the perception of the prevailing mood in society is not completely biased. Generally speaking, counterspeech is an attempt by civil society to **refute hate speech** and extremism with the aim to influence public discourses. It is usually understood to be **reactive**, comprising any response from members of the civil society to hateful, harmful or dangerous speech (Benesch et al. 2016: 2; Bojarska 2018: 15). However, counterspeech

may also be **proactive**, in the case of advocacy for members of discriminated groups, for example (Warnke 2021).

There is no set definition for hate speech but various definitions depending on the context (cf. Ylönen 2021). Thus, hate speech can be seen from a **language-critical perspective** as characterised by speech acts such as *denigrating, harassing, insulting, negatively stereotyping, stigmatising* or *intimidating* groups or members of groups (ECRI 2016: 3), who are attributed negative qualities that supposedly belong to this group collectively, universally and unchangeably (Scharloth 2017: 97). With a focus on content such as *racism, fake news* or *conspiracy theories*, and on its visibility in public discourses, hate speech can also be defined from a **discourse-critical perspective** as threatening democracy. Accordingly, counterspeech can also be viewed from a language- and discourse-critical perspective. In fact, the question of what is more important, form (factful and empathetic language) or content, was discussed intensively in the German *#ichbinhier* group. For example, not every group member approved an action defending Alexander Gauland, leader of the far-right Alternative for Germany party, when he became a target of scorn and derision after his clothes were stolen while he was swimming in a lake outside Berlin in 2018, which forced him to walk to the police station wearing only his bathing shorts. However, such actions defending right-wing politicians were not common. On the contrary, actions focused mostly on topics such as racism and discrimination. However, an analysis of the trigger discourses and strategies for counterspeech is a subject for another study.

In this article, I focus on factors that may possibly influence the visibility of counterspeech campaigns in social media and thus have an effect on protecting democracy, namely on the strategic organisation of counterspeech in the German and Finnish Facebook groups of the network *#iamhere international*, the most famous and perhaps unique network for counterspeech. This network consists of 16 private Facebook groups, acting in 20 countries and in 12 languages and has altogether over 150,000 members (iamhere international n.d. a). The initiative came from Mina Dennert who founded the Swedish Facebook group *#jagärhär* on the 13 May 2016 which has over 71,000 members (31 January 2022). Following Mina Dennert's initiative, the German group *#ichbinhier* was founded by Hannes Ley on 18 December 2016, and the Finnish group *#olentäällä* by Tarita Memonen on 19 February 2017. The German group has over 42,000 members and the Finnish group over 2,000 members (31 January 2022). Compared to May 2019, the Swedish and German groups both lost some 2,000 members whereas the Finnish group gained about 1,000 new members.

The main aims of the network *#iamhere international* are, among others, to "oppose hate speech, trolling, misinformation, intolerance, racism, homophobia and other discrimination", to "defend human rights, democracy and freedom of speech" in a "respectful, open, empathetic, polite and factful" way (iamhere international n.d. b). Although the groups from different countries share common goals and values,

they act with varying degrees of intensity and continuity that influence their visibility. The research question for this study is: To what extent do differences in intensity and continuity of activities depend on the level of strategic organisation in the Finnish and German group of the network *#iamhere international*?

3   Material and methods

The starting point for the contrastive analysis at hand were observations of the activities in the German and the Finnish Facebook groups of the network *#iamhere international*, over a long period. For the systematic analysis, all postings from the *#ichbinhier* group were collected manually from June until August 2020, and from the *#olentäällä* group from January until August 2020 (altogether 737 postings). The collecting periods in both groups differ because the activities in the German group were relatively high and steady whereas those in the Finnish group were lower and varied over time. Systematic documentation included recording the type (action or other), the date and the title of the posting in an Excel document. For actions, the medium (ZDF, Bild, Helsingin Sanomat, Yle etc.), the URL of the group's posting as well as its number of comments were recorded. More explicit descriptions, including the introductory texts of the groups' postings and some comments under the postings, were added later to ensure the postings identification, in case the posted content was no longer available later on. Until now, June 2022, only a few postings were no longer online, especially those actions that were targeted at the Facebook pages of *RT Deutsch*, obviously because the EU banned the distribution of Russian state media in February 2022.

To explore the **intensity** of the groups' activities, the number of postings by both groups for the investigation period was broken down according to their type (*counterspeech actions* and *other postings*) and set in relation to the average monthly activity. In order to obtain a more accurate picture of the **continuity** of activities in both groups, a monthly scope from January or June to August 2020 was examined (see section 4.1).

In order to examine the role of network strategies for the intensity and continuity of group activities, the groups' **strategic organisation of the activities** was analysed and compared. In particular, the types of *actions* and *other postings* and the degree of conventionalised group rules for these types were analysed (see section 4.2). All translations from German and Finnish into English were done by the author.

4 Results

4.1 Intensity and continuity of the activities of Facebook action groups #ichbinhier and #olentäällä

Overall, there were twice as many postings from #ichbinhier in just three months as from #olentäällä in eight months (see Table 1).

TABLE 1. Number of activities in the German and Finnish Facebook groups of the network #iamhere international during the period of investigation.

	#ichbinhier (June–August 2020)	#olentäällä (January–August 2020)
Actions	269	152
Member actions (bonfire/fire extinguisher)	82	11
Other postings	147	76
Total	**498**	**239**

Activities can be broadly divided into *actions*, *member actions*, and *other postings*. *Actions* are posted or approved by the moderators and draw attention to the location, that is, to the comment sections on social media, where counterspeech is needed. *Member actions* are — as the name suggests — actions that are suggested by members. For this purpose, the moderators posted calls to the group members to submit their suggestions for actions. These calls were marked with a gif for a bonfire (Lagerfeuer) in the German group and with an icon of a fire extinguisher (palosammutin) in the Finnish group (see Figure 1).

Guten Morgen liebe Gruppe,

wir zünden jetzt das tägliche Lagerfeuer an. Ihr könnt hier den ganzen T... Mehr anzeigen

🏷 Angehängte Themen #lagerfeuer

27   346 Kommentare

#olentäällä
olentäällä · 26.6.2020 · 🖼

#OLENTÄÄLLÄ -PALOSAMMUTIN

Kun kohtaat verkossa epäasiallista, vihapuhetta sisältävää kom... Mehr anzeigen

#olentäällä

PALOSAMMUTIN

20   2 Kommentare

FIGURE 1.   Member actions: *Lagerfeuer* (*bonfire*) in the German group, illustrated by a gif, and *Palosammutin* (*fire extinguisher*) in the Finnish group, illustrated by an icon.

Under these postings, members could make suggestions for counterspeech actions in comment sections where they needed help. As shown in Figure 1, the amount of comments differed greatly between the German *bonfire* and the Finnish *fire extinguisher*. There were also *other postings*, such as information and moderation notes or member discussions. The main difference between actions and member actions was that actions gained higher visibility because they stood alone in the timeline whereas member actions were posted as comments under the bonfire and fire extinguisher postings.

As expected, *actions* were posted most often. A simple division of the postings by the number of months showed that an average of three *actions* and an additional *bonfire* were posted on #ichbinhier every day while on #olentäällä *actions* were not posted every day and only around one *fire extinguisher* per month (see Figure 2). The number of *other postings* also differed greatly in both groups: It was five times as high in the *#ichbinhier* group in just three months as in the *#olentäällä* group in eight months.

Activities of counterspeech campaigns
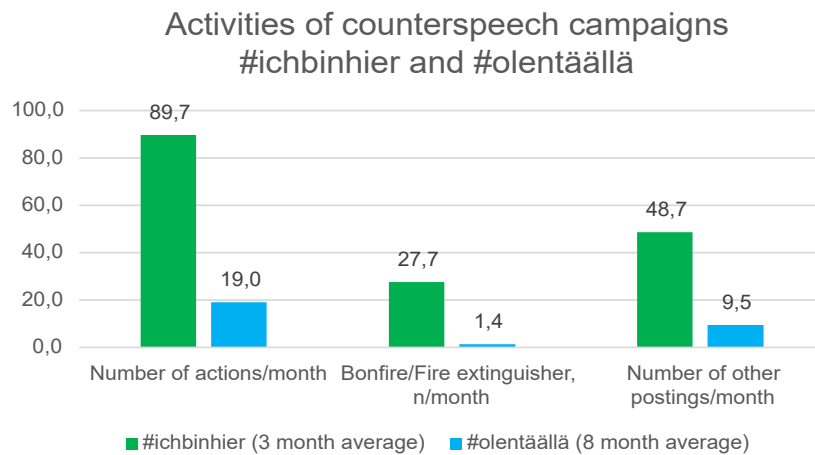#ichbinhier and #olentäällä

FIGURE 2.   Number of activities per month on average.

The number of *member actions* differed the most in both groups: *bonfire* postings
amounted to 22 times that of *fire extinguisher* postings. Among the 82 *bonfire* post-
ings for the three months of *#ichbinhier* examined, there were a total of 22,301 com-
ments, which is an average comment number of 272 per *bonfire*. The first comments,
usually ranging from 2 to 7, consisted of introductory tips from *bonfire fairies* and
elves (who moderated the *bonfires*), such as about the *tool* or the *Paten* (*godpar-
ents*). The *tool* was an app that was used by the *bonfire fairies* and *elves* to collect
the links to members' counterspeech comments into one place. The tool showed
200 links at a time, which were continually updated (when a new link was added, an
old one disappeared). This *tool* enabled optimised liking of group member's com-
ments, but Facebook inhibited its use in 2021, a move which was met with great
regret in the group. The aim of this *tool* was to promote the comments of the group
members with many likes, so that the Facebook algorithms would push them up the
timeline and thus give them greater visibility. The *Paten* (*godparents*) were one of
the 12 teams of the *#ichbinhier* group consisting of volunteers who offered help to
new members to orient themselves within the group and participate in its activities.
Following the introductory tips, group members could add links to media postings
where they needed help countering hate speech. These links were usually accom-
panied by short explanatory texts about the medium (*Medium*), the time (*Uhrzeit*)
of the posting and its topic (*Thema*), which together form the acronym *MUT* (the
German word for courage), see example 1.

(1)   medium: *RTL aktuell*, time: *9:45*, topic: *Five years 'We can do it', many Merkel-
        haters*, link.

One member action proposal usually had an average of five sub-comments, which contained the links to the group members' counterspeech comments. These links were either posted by the commentors themselves or alternatively (in case they had commented from a mobile device) they could ask for linking by the bonfire fairies and elves. The number of suggested actions posted by members of the German *#ichbinhier* group on a daily basis was enormous (samples showed 40 to 90 actions per day).

The eleven *fire extinguisher* calls on *#olentäällä* over a period of eight months received a total of only 46 comments, which resulted in an average of 4.2 comments per *fire extinguisher* call. These calls regularly included tips from the moderators, but they were only rarely used by members to suggest actions. On the other hand, suggestions for *member actions* were found in the comments to *actions* posted or released by the administrators, and these suggestions concerned other media houses that dealt with the same topic. For example, when an *action* was posted concerning a report about Black Lives Matter demonstrations on the Facebook page of the tabloid newspaper *Iltalehti*, members posted links to the Facebook pages of the country's other major tabloid newspaper *Iltasanomat* as well as to *Yle TV* where the same topic was reported.

There were also major differences in the **continuity** of both groups' activities. While the German *#ichbinhier* group had various postings on a daily basis, the number of activities in the Finnish *#olentäällä* group varied greatly (see Figure 3). From January to mid-March, there was one posting or more almost daily.

One explanation for why activities in the Finnish *#olentäällä* group were relatively high at the beginning of 2020 is that the group had received many new members at the beginning of the year via the "herring movement" (*silakkaliike*), founded on 25 December 2019 in response to racially motivated hunts for children who were transferred from the Syrian al-Hol camp to Finland shortly before Christmas. Such racially motivated malice triggered a wave of outrage, after which many Finns engaged in counterspeech. The sharp drop in activities from mid-March is due to the consequences of the COVID-19 pandemic. In July and August, there were hardly any activities left, for which, in addition to the pandemic, the vacation period is also responsible.

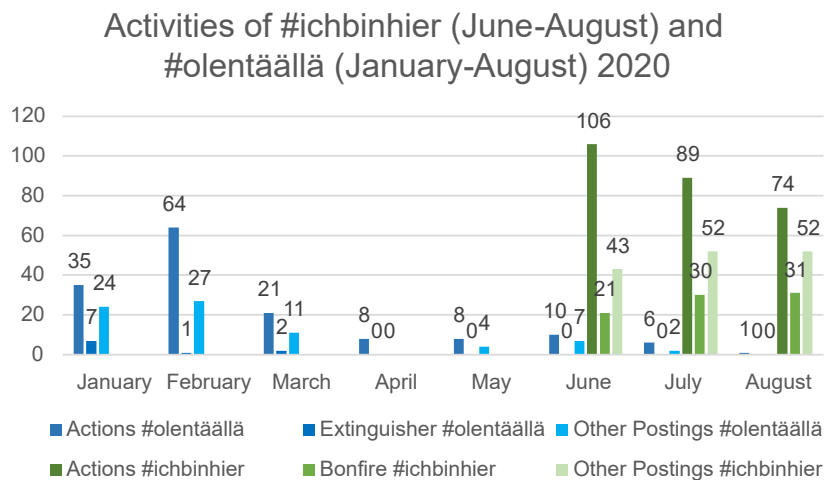## Activities of #ichbinhier (June-August) and #olentäällä (January-August) 2020

FIGURE 3. Actual distribution of activities from January until August 2020
(#olentäällä) and June until August 2020 (#ichbinhier).

The German #ichbinhier group also experienced a slight decline in activity from June to August, which is also linked to the holiday season here. From 23 to 29 July, the group went on company holidays, meaning no *actions* were posted during this time and the *bonfire* was burning on a low flame. Based on the three months examined here, no reliable statements can be made about the continuity of activities over a longer period of time. However, I can say from experience that activities of the *#ichbinhier* group were not subject to major fluctuations, even though the number of daily counterspeech *actions* (which are the "core business" of these action groups) varied from one to eight.

4.2   Strategic organisation of activities in the #ichbinhier and #olentäällä groups

In the following, I will analyse how *actions*, *member actions* and *other postings* were organised, that is, whether and what kind of rules had been developed in the German and Finnish Facebook groups of the network *#iamhere international* and to what extent these rules had been conventionalised.

4.2.1 Actions
Concerning *actions*, I will focus on who posted calls, what kind of locations were the actions aimed at, to what extent the actions were categorised further, and how far introductory texts and member comments were conventionalised (see Table 2).

TABLE 2.   Strategic organisation of *actions.*

	#ichbinhier	*#olentäällä*
Continuity and intensity	Daily 1–8 (except for holidays: 7 days in July 2020)	Sporadically since mid-March 2020
Authors of action postings	Moderators	Moderators and members (approved by moderators)
Target location	Comment sections of big media houses' Facebook pages with more than 10 000 followers	Comment sections of big and small media houses' Facebook pages, Facebook pages of other groups and politicians, Twitter
Further categorisation	Explicitly with headings: *Action 1, 2, 3 …, Hotspot, Solidarity action, Action International*	No headings except for international actions: *kansainvälinen tehtävä!*
Introductory texts	Conventionalised, detailed, reactive	Variable, short, reactive and proactive
Member comments	Links to member comments *TLC* and *SLC* (top and sub level comments) or requests for linking, no content-related comments	*T, K, I* (L, C, R – Like, Comment, Report), content-related comments and discussions until the beginning of June 2020 as well as further suggestions for actions made by members (mostly related to the same topic on other media houses' Facebook pages)

In the German group, up to eight *actions* were posted daily by the moderators, except for the one-week holiday in July 2020. They focused on the comment sections of big media houses' Facebook pages with more than 10,000 followers. The *#ichbinhier* Group's daily actions are numbered (*action 1, action 2, …*) or labelled with headings according to the location chosen for an action (*hot spot, solidarity action*). The heading *hot spot* (*Brennpunkt*) refers to media sites with particularly high hate speech potential in comment sections, such as *RT Deutsch*, where counterspeakers were advised to participate only if they had strong nerves, and after they had controlled their private settings first in order to protect themselves from attacks. *Solidarity actions* (*Solidaritätsaktionen*) were launched on the pages of public figures, institutions, initiatives or enterprises, who became targets of online harassment, such as the politician Karl Lauterbach, the German rock band BAP, the police Berlin, UNICEF Deutschland, Knorr and so on.

In the Finnish group, *actions* were posted by moderators or members (the latter released by moderators with some delay) and targeted not only big media's Facebook pages, but also the comment sections of other Facebook groups, pages

of political parties or politicians and government, smaller enterprises and even Twitter accounts. At the beginning of June 2020, a guiding notice was published, announcing that actions would be started in comment sections of public media and corresponding institutions, but not on individual people's profile pages. A similar categorisation of actions as in the German group, with headings such as *action 1, action 2, … hot spot, solidarity action*, was not found in the #*olentäällä* group. After all, there was rarely more than one action per day, and after mid-March only a few actions, which is why it was not necessary to count them.

Both groups participated in a manageable number of *international actions* of the #*iamhere international* network, which were marked as *Aktion international / Kansainvälinen tehtävä*, on the Facebook pages of the United Nations, the European parliament or UNICEF. During the period examined, #*ichbinhier* participated in four such actions and #*olentäällä* in six.

The **introductory texts** to *actions* were highly conventionalised in the German group (example 2), whereas they varied greatly in the Finnish group (examples 3–4).

(2)   *Liebe Gruppe, […] unsere sechste Aktion heute führt uns zur Seite der FAZ (Achtung Wortfilter) […] Lasst uns gerade hier respektvoll, empathisch, sachlich und differenziert in den Diskurs einsteigen. […] Da es in letzter Zeit doch sehr emotional zuging, hier der Hinweis: Wir wollen eine anständige Diskussionskultur fördern, aber niemanden belehren oder gar beleidigen. Bitte diskutiert sachlich und lasst euch nicht provozieren. […]*
Dear group, […] our sixth action today leads us to the FAZ page (Attention word filters!) […] Let us, **especially here**, enter the discourse in a respectful, empathetic, factual and differentiated manner. […] Since things have been very emotional **lately**, here's the advice: We want to promote a decent culture of discussion, but don't want to lecture or even insult anyone. Please discuss factually and don't let yourself be provoked. […] (my emphasis)

(3)   *Olkaa siellä ennen trolleja.*
Be there before the trolls.

(4)   *Täällä tarvitaan kommenttiapua.*
Commenting help is needed here.

The German introductions were relatively long and made use of schematic formulations that were found in all action postings, with only a few variations, if at all. The schematic use of expressions such as "especially here" and "lately" (see example 2) loses persuasiveness to some extent because of the repetition in each posting over a long time. On the contrary, Finnish introductions were usually formulated briefly

and individually varying (see examples 3 and 4), and often also proactively, as in example 3.

There were strict rules for member comments under *actions* in *#ichbinhier*, while certain rules seemed to be developing slowly in *#olentäällä*. The comments under *#ichbinhier* actions were given by the group members and contained the links to their counterspeech comments (e.g., www.facebook.com/bild/posts/10159627575125730..., with the dots indicating the address to the group members' comments). In case a group member commented from a mobile device and could not provide a link, a request for linking was made that mentioned, for example, that a top level comment (TLC) was posted (*TLK mobil*). Subsequently, the moderators of the respective action provided the links to these comments in sublevel comments to the request. These links helped to find the comments of *#ichbinhier* group members in the comment sections more quickly so that they could like or comment on them. There was no discussion here. In the comment sections under actions by *#olentäällä*, on the other hand, there were often extensive (also controversial) discussions. These discussions were either about the topics covered in the media reports (sometimes with links to sources that should support their own arguments) or about the kind of comments under these media reports (see example 5).

(5)   *Huhhuijjakkaa. Tämä käy kyllä välillä ihan työstä. Miksihän nuo mediatalot ei vaan*
        pistä tuonen sivuilleen automaattisia "roskasuodattimia" päälle?
        Ohoho. That really turns into real work sometimes. Why do these media companies
        simply not put automatic "spam filters" on their pages?.

Comments on the impact of the group's engagement were also found here, often combined with thanks to the colleagues for their efforts (see example 6).

(6)   *Kiitos ja kumarrus teille, jotka olette kommentoineet asiallisesti.* 🏹 🐚
        Kommenttiosiot ovat alkaneet kummasti siistiytyä.
        Thanks and a bow to you for commenting factually. 🏹 🐚
        The comment sections have become amazingly clean.

The *#olentäällä* moderators only urged members to enter a *T* for *tykkäys* (L for like) and/or *K* for *kommentti* (C for comment) if they had liked a comment or had commented on one (or more). Some commenters also informed the members they had reported hate speech to Facebook *I* for *ilmianto* (R for report). In other words, on the Finnish site they were not expected to post the link (in contrast to the German site).

4.2.2 Member actions
Different organisational strategies were also identified for member actions (*Bonfire* and *Fire extinguisher* postings, see Table 3). As noted in section 4.1, *#ichbinhier*'s

Bonfires were usually lit daily, whereas the *#olentäällä* group had only 11 calls for member actions in eight months: seven in January, one in February, two in March and one in June.

TABLE 3.   Strategic organisation of *member actions.*

	#ichbinhier	**#olentäällä**
Label	*Bonfire (Lagerfeuer)*	Varying, *Extinguisher (Feuer-löscher)* since end of January
Continuity/intensity	Daily	Sporadically
Location	Facebook pages of so-called mass media of any shade	Not specified
Introduction to member actions	Regular specification of explicit and consistent rules	Sporadic specification of varying rules

In the Finnish group, these calls for member proposals were only referred to as *Fire extinguishers* from the sixth call at the end of January. Before that, a no-parking sign was used consisting of a head with a speech bubble containing hate speech ("! @$% "&") coming from its mouth, crossed out by the bar of the prohibition sign. The third call in January was called *Ilmiantoketju* (Reporting chain). For the first time (in the material examined here) there was the advice to post suggestions either as a comment under this post or directly in the group's timeline. Since the seventh call at the end of January, members have been encouraged to post their proposals directly to the timeline as an action. In the ninth and tenth calls in March, it was once again added that suggested actions can also be posted as comments on the *fire extinguisher* when speed matters (because the actions posted by members in the chronicle must first be approved by the moderators, which can lead to delays). The eleventh *fire extinguisher* in June did not mention any place to submit proposals. There were no clear rules for posting proposed actions during the time investigated, and hints that could help members find their way around were sometimes contradictory and difficult to find. Only the *Ilmiantoketju* (Reporting chain) post (the third call in January) could be found with less effort under the button "Topics in this group" (on computers) or "Topics" (on mobile phones).

In the German group, there were clear rules for the bonfire, which were repeated in the postings every day. These relate, among other things, to the form of action, as in example 7:

(7)   *Guten Morgen liebe Gruppe, wir zünden jetzt das tägliche Lagerfeuer an. Ihr könnt
        hier den ganzen Tag Links zu Beiträgen posten, in denen ihr auf besonders viele*

> *unsachliche, nicht zielführende Kommentare gestoßen seid und Unterstützung benötigt.*
>
> Good morning, dear group, we are now lighting the daily bonfire. You can post links here all day long to posts in which you have come across of inappropriate, ineffective comments and need support.

In addition, there were rules for the places of action (see example 8) and for content details (see example 9):

(8)   *Denkt bitte daran, dass wir auf folgenden Seiten nicht aktiv sind. Entsprechende Verlinkungen dazu werden von uns kommentarlos gelöscht:1. Parteiseiten, 2. Private Seiten. Kein Problem sind Facebookseiten sogenannter Massenmedien jeglicher Couleur.*
        Please remember that we are not active on the following pages. Corresponding links will be deleted by us without comment: 1. Party pages, 2. Private pages. Facebook pages of so-called mass media of all shades are no problem.

(9)   *Bitte gebt neben dem Link das Medium, das Thema, sowie den Startzeitpunkt des Artikels an und vermeidet visuelle Vorschau des Artikels ("Vorschau entfernen"). […] Bitte führt hier auch keine inhaltlichen Diskussionen, damit der Thread übersichtlich bleibt.*
        Please specify the medium, topic and start time of the article next to the link and avoid visual preview of the article ("Remove preview"). […] Please do not have any content discussions here either, so that the thread remains clear.

There was also a link to the "tool", an app in which all links to member comments were listed, providing the possibility to like them in a time-saving manner (without having to visit the comment sections of the media houses). However, since 2021, this tool is no longer available due to technical changes on the Facebook platform (see also section 4.1).

4.2.3 Other postings

Other postings included discussions, information, and moderation notes. In the #ichbinhier group, they were divided into certain categories by headings (*Nightcap*, *#ichtalkhier*, *Boot camp*, *Company holidays* and *Welcoming new members*), while the posts in the #olentäällä group were usually not explicitly identified by headlines – yet could be assigned to specific categories (member discussions, moderation notes and information, as well as welcoming new members and birthday wishes). The nature and scope of such posts is shown in Figure 4.

Other postings of #ichbinhier and #olentäällä

Nightcap	28,3
#ichtalkhier/member discussions	2,1 12,3
Info, advice	5,1
Bootcamp	3,3
Welcoming new members	1,1 1,7
Holidays	3,0
Birthday greetings	1,1

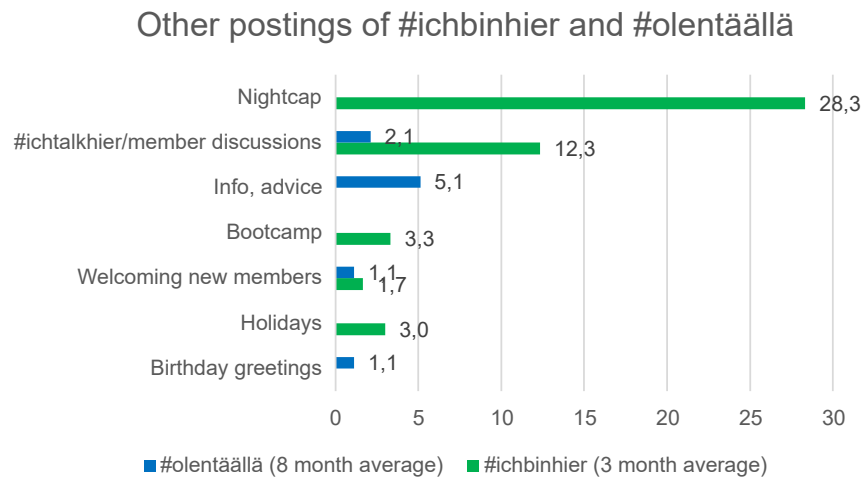■ #olentäällä (8 month average)   ■ #ichbinhier (3 month average)

FIGURE 4.   Average number of other postings per month.

Other categories did not occur during the investigation period. One example is the *fan mail* section at #ichbinhier, which expressed thanks from people and organisations (e.g., from the rock band BAP in September 2020 or from UNICEF in May 2020). Such acknowledgements were gladly received and interpreted as a sign of the success and sustainability of #ichbinhier's counterspeech campaigns.

Discussions in the German group had clearly defined functions and were strictly regulated in terms of time. *Nightcaps* (*Absacker*) served as a way to calm down after a busy day in the comment sections. Moderators and sometimes members posted stories related to pictures, videos, or songs. In the comments posted below, members expressed their thoughts on the chosen topic. Occasionally an *Explanatory bear* (*Erklärbär*) fulfilled the role of nightcap, a discussion round where members could ask questions about the organisation and working methods of the group or the terminology typically used (what are TLK and SLK, how do forms of action differ, etc.) or about technical issues on Facebook. *#ichtalkhier* sessions consisted of thematically focused discussions that were relevant to the work of the group, such as form of address (*Du* or *Sie*), blocking, clear announcement or insult, agitations, networking and joy. Such discussions always took place in the evenings for two hours after which the comment sections were closed. In the Finnish group, member discussions appeared spontaneously and included questions and suggestions (such as how to use the hashtag or collaborate with the herring movement), thanks (e.g., from members of the herring group for supporting their counterspeech campaigns), information, warnings (against fake news media) and calls (to sign petitions).

5   Conclusion and prospect

Research has shown that right-wing politicians and their followers, who use social media over-proportionally (Lucht et al. 2017; Knuutila 2019: 4) and often resort to bots (Egli & Rechsteiner 2017: 250; Kreißel et al. 2018), create the impression that their views reflect the opinion of the majority of a society. Counterspeech as practiced by the groups of the network *#iamhere international* aims to communicate "balanced opinions to silent readers" (Laubenstein & Urban 2018: 54) in a "respectful, open, empathetic, polite and factful" way (iamhere international n.d. b). In other words, counterspeech aims to influence public discourse in factful and empathetic ways so that the perception of the prevailing mood in society is not completely biased. In this study, I focused on a contrastive analysis of the **quantitative** appearance of counterspeech campaigns of the German and Finnish groups of the network *#iamhere international*, which has an influence on the visibility of the campaigns. A contrastive analysis of the groups' qualitative counterspeech characteristics, such as objectivity, respect and empathy, as well as an analysis of trigger discourses for counterspeech, such as racism, migration, and climate change, will be the subject of another study.

Organised counterspeech was more active in Germany (*#ichbinhier*) than in Finland (*#olentäällä*), as shown in the analysis of the intensity and continuity of the groups' activities (section 4.1). These differences obviously result from the existence or absence of a strategic organisation of group activities. Counterspeech was strategically organised in the #ichbinhier group, where conventionalised rules for *actions*, *member actions* and *other postings* were found, while in the *#olentäällä* group a limited number of conventions seemed to be developing slowly (section 4.2). One reason for the strategically better organised work of the German group was apparently the foundation of a non-profit association *ichbinhier e.V.*, registered 28 September 2017, which helped to professionalise the otherwise mainly voluntary engagement in counterspeech to some degree. In the German *#ichbinhier* group, 12 teams were responsible for organising all activities, such as the teams *extended management circle* (*Erweiterter Leitungskreis*), *timeline* (*Chronik*), *bonfire* (*Lagerfeuer*), *members' reception* (*Mitgliederempfang*), *godparents* (*Patinnen*), and so on. These teams were introduced in *introductory nightcaps* (*Vorstellungsabsackern*) from November 2020 until February 2021, that is, after the systematic data collection for this study (but analogue introductions were provided also earlier in 2018 and 2019). In the Finnish *#olentäällä* group, a comparable strategic organisation by teams responsible for specific tasks could not be identified.

Clear rules for communication within and outside groups with thousands of members obviously provide valuable information for participation in counterspeech campaigns. In addition, they facilitate the administration of the groups and the moderation of the activities. On the other hand, missing or inconsistent rules make

orientation difficult and can have a negative impact on the intensity and continuity of group activities.

In order to have a lasting positive influence on social discourse in terms of the fifth estate, to defend human rights, democracy and freedom of expression and to counteract division in society, social media counterspeech campaigns must be quantitatively visible in order to show silent readers that hate and disinformation do not go unchallenged. Strategic organisation of activities can promote this visibility. The positive impact of counterspeech actions on the tone of comment sections was discussed within the Finnish group (see chapter 4.2), and the study by Ziegele et al. showed that the predominantly respectful and discursive comments of the action group *#ichbinhier* could help silent readers to perceive a better overall discussion climate (Ziegele et al. 2019: 7).

However, digital genres change rapidly and some of the conventions found in 2020 have already changed due to Facebook's technical evolution and group dynamics. While activities and the number of members in the German group were somewhat decreasing, and activities even paused from 10 May 2022 on to develop new concepts, they are increasing slightly in the Finnish group. The reasons for such developments could be the subject of further investigations. Preliminary observations suggest that the strategic organisation of activities could profit from defining the relationship between the form (language) and content (discourse) of counterspeech.

Literature

Benesch, S., D. Ruths, K. P. Dillon, H. M. Saleem & L. Wright 2016. Considerations for successful counterspeech. *Dangerous speech project*. https://dangerousspeech.org/wp-content/uploads/2016/10/Considerations-for-Successful-Counterspeech.pdf [accessed 4 Feb. 2022].

Bojarska, K. 2018. The dynamics of hate speech and counter speech in the social media. Summary of scientific research. Europa Universität Viadrina Frankfurt/O.: Centre for the Internet and Human Rights (CIHR). 1–16. https://cihr.eu/wp-content/uploads/2018/10/The-dynamics-of-hate-speech-and-counter-speech-in-the-social-media_English-1.pdf [accessed 25 Aug. 2022].

Bunz, M. 2012. Die fünfte Gewalt: Über digitale Öffentlichkeit und die Neuausrichtung von Journalismus und Politik. In U. Rußmann, A. Beinsteiner, H. Ortner & T. Hug, (eds): *Grenzenlose Enthüllungen. Medien zwischen Öffnung und Schließung*. Innsbruck: Innsbruck University Press 2012 (Medien – Wissen – Bildung 2), 163–167. http://dx.doi.org/10.25969/mediarep/791.

Burkhard, S. 2015. *Medienskandale: Zur moralischen Sprengkraft öffentlicher Diskurse*. 2., überarbeitete und ergänzte Auflage. Köln: Halem.

Dumbrava, C. 2021. *Key social media risks to democracy. Risks from surveillance, personalisation, disinformation, moderation and microtargeting*. European Parliamentary Research Service (EPRS). Brussels: European Union. https://www.europarl.europa.eu/thinktank/en/document/EPRS_IDA(2021)698845. [assessed 17 June 2022].

ECRI 2016. *ECRI General policy recommendation no. 15 on combating hate speech.* Strasbourg: European Commission against Racism and Intolerance (ECRI), Council of Europe. https://rm.coe.int/ecri-general-policy-recommendation-no-15-on-combating-hate-speech/16808b5b01 [accessed 5 Feb. 2022].

Egli, P. & D. Rechsteiner 2017. Social Bots und Meinungsbildung in der Demokratie. *Aktuelle Juristische Praxis,* 2/2017: 249–258.

iamhere international n.d. a). Join us. https://iamhereinternational.com/join-us/ [accessed 5 Feb. 2022].

iamhere international n.d. b). Our story. Our Manifesto. Our Values. https://iamhereinternational.com/who-we-are/ [accessed 5 Feb. 2022].

Knuutila, A. 2019. Närkästyksen kone: Miksi uusoikeiston ääni kuuluu verkossa muita vahvemmin? (Indignation machine: why is the voice of the new right on the net louder than that of others?). *Politiikasta* (From politics) 5 Sep 2019. https://politiikasta.fi/narkastyksen-kone-miksi-uusoikeiston-aani-kuuluu-verkossa-muita-vahvemmin [accessed 29 May 2020].

Knuutila, A., H. Kosonen, T. Saresma, P. Haara & R. Pöyhtäri 2019. Viha vallassa: Vihapuheen vaikutukset yhteiskunnalliseen päätöksentekoon. (Hatred in Power: Influence of Hate Speech on Social Decision-Making). *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja* (Reports and Research of the Government Chancellery) 2019: 57. Helsinki: Valtioneuvoston kanslia. http://urn.fi/URN:ISBN:978-952-287-786-4 [accessed 22 May 2020].

Kreißel, P., J. Ebner, A. Urban & J. Guhl 2018. *Hate at the push of a button. Right-wing troll factories and the ecosystem of coordinated hate campaigns online*. London: Institute for Strategic Dialogue (ISD). https://www.isdglobal.org/wp-content/uploads/2020/04/Hate-at-the-Push-of-a-Button-ISD.pdf [accessed 20 June 2020].

Laubenstein, S. & A. Urban 2018. Case studies: which types of campaign against hate and extremism on the internet work, which do not, and why? In Baldauf, J., J. Ebner & J. Guhl (eds) *Hate Speech and Radicalisation online. The OCCI Research Report (Online Civil Courage Initiative)*. London: Institute for Strategic Dialogue, 50–57. https://www.isdglobal.org/isd-publications/hate-speech-and-radicalisation-online-the-occi-research-report/ [accessed 5 Feb. 2022].

Lucht, J., L. Udris & D. Vogler 2017. *Politische Inszenierungen: Eine Inhalts- und Resonanzanalyse der Facebook-Seiten bundesdeutscher Parteien*. Bonn: Friedrich-Ebert-Stiftung. https://doi.org/10.5167/uzh-147523.

Mejias, U. A. & N. E. Vokuev 2017. Disinformation and the media: the case of Russia and Ukraine. *Media, Culture & Society,* 39(7) 1027–1042. https://dx.doi.org/10.1177/0163443716686672.

Montag, C. 2018. Filter bubbles: how do filter bubbles affect (political) opinion, taking personality into account? In Baldauf, J., J. Ebner & J. Guhl (eds) *Hate Speech and Radicalisation online. The OCCI Research Report (Online Civil Courage Initiative)*. London: Institute for Strategic Dialogue, 27–34. https://www.isdglobal.org/isd-publications/hate-speech-and-radicalisation-online-the-occi-research-report/ [accessed 5 Feb. 2022].

Saresma, T. 2017. Väkivaltafantasiat ja pelon politiikka. Sukupuolitetun ja sukupuolittavan verkkovihapuheen luentaa. (Violence fantasies and the politics of fear. Interpreting gender-specific and gender-specifying online hate speech.) In S. Karkulehto & L.-M. Rossi (eds) 2017. *Sukupuoli ja väkivalta. Lukemisen etiikkaa ja politiikkaa.* (Gender and violence. Reading ethics and politics.) 219–244. Helsinki: Suomalaisen kirjallisuuden seura. https://oa.finlit.fi/site/books/10.21435/skst.1431/ [accessed 28 March 2018].

Scharloth, J. 2017. Editorial. In J. Scharloth (ed) Themenheft: Hate Speech/Hassrede. *Aptum –
Zeitschrift für Sprachkritik und Sprachkultur,* 02/2017, Hempen Verlag. 97–99.

Schwarz, K. & J. Holnburger 2018. Disinformation: what role does disinformation play for
hate speech and extremism on the internet and what measures have social media
companies taken to combat it? In Baldauf, J., J. Ebner & J. Guhl (eds) *Hate Speech
and Radicalisation online. The OCCI Research Report (Online Civil Courage Initiative)*.
London: Institute for Strategic Dialogue, 35–43. https://www.isdglobal.org/isd-
publications/hate-speech-and-radicalisation-online-the-occi-research-report/
[accessed 5 Feb. 2022].

Warnke, I. 2021. Widerrede als Fürsprache? Überlegungen zur Doppelgesichtigkeit von
Empörung. *Fachtagung "Pragmatik der Gegenrede"*. Universität Greifswald, 25-26 Feb.
2021. Keynote. https://www.wiko-greifswald.de/mediathek/beitrag/n/widerrede-
als-fuersprache-ueberlegungen-zur-doppelgesichtigkeit-von-empoerung-87847
[accessed 15 June 2022].

Wesley, M. & K. Gradon, 2020. COVID-19 Effects and Russian Disinformation. *Homeland
Security Affairs* 16, Article 8. https://www.hsaj.org/articles/16533. [accessed 17 June
2022].

Wodak, R. 2021. *The Politics of Fear. The Shameless Normalization of Far-right Discourse*. 2nd
revised and extended edn. London: Sage.

Wodak, R., J. Culpeper & E. Semino 2021. Shameless normalisation of impoliteness:
Berlusconi's and Trump's press conferences. *Discourse & Society,* 32(3): 369–393.
https://doi.org/10.1177/0957926520977217.

Ylönen, S. 2021. Soziale Medien als fünfte Gewalt: Strategische Organisation deutscher und
finnischer Gegenredekampagnen auf Facebook. Discussion Paper. *jfml Journal for
Media Linguistics*. https://dp.jfml.org/2021/opr-ylonen-soziale-medien-als-funfte-
gewalt-strategische-organisation-deutscher-und-finnischer-gegenredekampagnen-
auf-facebook/ [accessed 25. August 2022].

Ziegele, M., P. Jost, D. Frieß & T. Naab 2019. Aufräumen im Trollhaus. Zum Einfluss von
Community-Managern und Aktionsgruppen in Kommentarspalten. *DIID-Précis
04/2019*. Düsseldorf: Institute for internet and democracy. https://dx.doi.org/
10.36200/0012.

Zweig, K., O. Deussen & T. Krafft 2017. Algorithmen und Meinungsbildung. *Informatik
Spektrum,* 40, 318–326. https://dx.doi.org/10.1007/s00287-017-1050-5.