

counter-speech on Facebook

Jamie Bartlett
Alex Krasodonski-Jones

September 2016



Open Access. Some rights reserved.

As the publisher of this work, Demos wants to encourage the circulation of our work as widely as possible while retaining the copyright. We therefore have an open access policy which enables anyone to access our content online without charge. Anyone can download, save, perform or distribute this work in any format, including translation, without written permission. This is subject to the terms of the Demos licence found at the back of this publication. Its main conditions are:

- Demos and the author(s) are credited
- this summary and the address www.demos.co.uk are displayed
- the text is not altered and is used in full
- the work is not resold
- a copy of the work or link to its use online is sent to Demos

You are welcome to ask for permission to use this work for purposes other than those covered by the licence. Demos gratefully acknowledges the work of Creative Commons in inspiring our approach to copyright. To find out more go to www.creativecommons.org



Published by Demos September 2016
© Demos. Some rights reserved.

Unit 1, Lloyds Wharf
2-3 Mill Street
London
SE1 2BD

hello@demos.co.uk
www.demos.co.uk

This paper was supported by Facebook. The views expressed are those of the authors and do not necessarily reflect those of Facebook. All errors and omissions are the authors' own.

CONTENTS

Background	4
Method	5
Case Study 1: France	9
Case Study 2: United Kingdom	13
Case Study 3: Morocco and Tunisia	16
Case Study 4: Indonesia	22
Case Study 5: India	28
Case Study 6: Post-Paris Terror Attacks	32
Overall Conclusions	35

BACKGROUND

Facebook serves more than 1.5 billion people globally. Although the majority of people use the site for positive purposes, there are some who use the platform in negative ways. With that in mind, Facebook has created a set of policies – its [Community Standards](#) – detailing what type of content people can and cannot post. For instance, Facebook prohibits and removes hate speech and does not allow dangerous organisations (defined as groups that engage in terrorist or organised criminal activity) to have a presence on Facebook. Content that supports or promotes those groups is removed. However, sometimes people post content which other users may consider hateful or extreme, but which does not violate Facebook’s policies.

To counter this type of disagreeable or extremist content, Facebook has publicly stated that it believes counter-speech is not only potentially more effective, but also more likely to succeed in the long run.

Counter-speech is a common, crowd-sourced response to extremist or hateful content. Extreme posts are often met with disagreement, derision and counter-campaigns. Combating extremism in this way has some advantages: it is faster, more flexible and responsive, and capable of dealing with extremism from anywhere and in any language; and it retains the principle of free and open public spaces for debate. However, the forms counter-speech takes are as varied as the extremism it argues against. It is also likely that it is not always as effective as it could be; and some types of counter-speech could potentially even be counter-productive.

In the light of its belief in the power of counter-speech and the growing interest in a more rigorous and evidence-led approach to understanding it better, Facebook commissioned Demos to undertake this research report, examining the extent to which different types of counter-speech are produced and shared on Facebook.

In October 2015 we published a [report](#), supported by Facebook, which examined the activity of counter-speech and populist right-wing groups on Facebook and made recommendations for how counter-speech groups could more effectively diffuse their messages. This present report sets out the findings of phase II of this project, examining how speech which challenges extreme Islamist narratives in six countries – France, India, Indonesia, Morocco, Tunisia and the UK – is produced and shared.

IS propaganda

In recent months there has been increased concern about how various extremist groups (in particular so-called Islamic State or IS) are using social media to share propaganda and recruit. How to best respond to this remains a pressing policy question: particularly in the case of content which, although extreme, might not technically break any laws or contravene Facebook's policies.

There is a growing consensus that, alongside efforts to remove certain types of content, counter-speech and counter-messaging must be part of any response. For example, David Fidler's work at the Council on Foreign Relations argues that any online response should be based on counter-speech and challenging extreme content in a 'marketplace of ideas'. Similarly, a White Paper by the Quilliam Foundation describes 'censorship and filtering initiatives' as 'ineffective', and emphasises the critical role of counter-speech in 'challenging the sources of extremism and terrorist material online'.¹ At the 2015 White House Summit on Countering Violent Extremism, challenging extremist narratives online was one of the three key programmes in defeating IS.

At Demos we believe it is important that the principle of internet freedom should be maintained; and that the internet should be a place where people feel they can speak their minds openly and freely. We therefore believe that debate, disagreement and challenge is nearly always preferable to censorship and removal of content, including when dealing with extreme or radical content, whatever its origin. However, we also believe that this can and should be put on an empirical basis to help us better understand the phenomenon and how to respond. This research series is an attempt to do that.

METHOD

For every country included in this report we collected data from public Facebook Pages using an iterative process. Researchers from Demos identified Pages liked by individuals who had reported extremist content, and manually marked up those Pages where prima facie counter-speech was taking place. We defined this as content

¹ David Fidler, 'Countering Islamic State exploitation of the internet', www.cfr.org/cybersecurity/countering-islamic-state-exploitation-internet/p36644 (accessed Sept 2016). See also Erin Marie Saltman and Jonathan Russell, 'White Paper – the role of Prevent in countering online extremism', www.quilliamfoundation.org/wp/wp-content/uploads/publications/free/white-paper-the-role-of-prevent-in-countering-online-extremism.pdf (accessed Sept 2016); and Scott Beattie, *Community, Space and Online Censorship: Regulating pornotopia* (Routledge, 2009).

which criticised, confronted, disagreed with or presented an alternative to IS. We then identified other similar Pages to those already identified, and again manually marked up Pages that appeared to include counter-speech and removed those which did not. This process usually took three iterations to complete before we felt confident we had a range of relevant Pages. However, we cannot claim this is a comprehensive set of all relevant Pages.

We did not attempt to collect or use any personal information about individuals; nor did we attempt to identify any individuals. We did not collect any data from groups' or from individuals' Pages; and we did not collect any data from closed or secret Pages. Throughout, only data from Pages that were public and viewable by everyone were used. In order to further protect individual privacy, we have not quoted or republished any specific posts that might identify individuals.

Across the different countries we collected and manually marked up a total of 1,425 Pages. Using Facebook's public 'API' (Application Programming Interface), we collected public posts and interaction data from these Pages. We used 'R', an open source software that allows researchers to access publicly available data from the public API. 'Posts' in this sense refers to updates that were made on the Page by the administrator(s) of that Page. In addition to posts, we collected all the interactions that were associated with the posts. Interactions refers to 'likes', 'shares' and 'comments' on those posts. Interaction data can be used to estimate the reach of content, because each time a user interacts with a piece of content it will appear in their friends' timeline (depending on the privacy settings applied).

Over a period of 13 months, using this method we identified Pages, posts and interactions in each country as shown in table 1.

Table 1: Data collected

Country	Pages	Posts	Interactions (rounded)	Date of data collection
UK	355	6,482	677,000	24/6/15–24/7/15
France	229	10,523	626,000	1/4/15–12/4/15
Morocco and Tunisia ²	147	9,436	149,000	13/11/15–20/11/15
Indonesia	382	8,572	280,000	1/3/16–31/3/16
India	312	4,866	42,000	17/5/16–28/5/16
Total	1,425	38,879	1,774,000	

² After reviewing the available data, given the relatively small volume for each country and the overlap between themes we decided to combine Tunisia and Morocco into a single category.

Each country is slightly different in terms of the date range, volume and interactivity of the Pages we identified. Because of the way the data were collected, we do not claim this is a comprehensive set of Pages where counter-speech might occur.

Following the Paris attacks of 13 November 2015, we collected the data again from the same Pages between 13 and 21 November, in the UK and France, in order to calculate any changes in activity.

We conducted a series of analyses. This included:

- a) calculating average interactions per Page and per post using automated API results
- b) calculating the format of the most popular types of data using automated API results
- c) calculating the type and style of the most popular types of content through manual human analysis
- d) calculating the types of speech occurring on different Pages using manual human analysis
- e) calculating the way different types of content was shared on Pages vis-à-vis users' own newsfeeds using automated analysis

In every country this analysis was conducted in the language of the country in question. In France, we analysed French Pages and posts; in Tunisia and Morocco, Arabic (and a small number of French) Pages and posts; in India, Hindi; in Indonesia, Indonesian; in the UK, English.

We found that specific types of content were very different in each country, which meant that the categories used to define Pages and specific posts varied significantly. Therefore, we did not combine data sets from different countries. Furthermore, available data varied slightly per country, which meant there were small variations in the types of analysis carried out.

The purpose of these modes of analysis was to better understand the scale and nature of counter-speech content on Facebook and to identify what types of content were most likely to be engaged with by users. However, it is important to stress that these are in many cases quite novel methodologies. There are no firmly established 'best practice' methods to collect and analyse data of this nature – social media research is an emerging academic discipline. It is not possible to set out, for example, the total volume of counter-speech that is taking place, and this report does not aim to do that; rather it aims to illustrate the nature and popularity of different sorts of counter-speech through a series of country-specific case studies.

Therefore this report should be understood as a scoping study rather than a comprehensive analysis of counter-speech, and findings should be read as such.

The report is structured as follows. First we set out the results from each case study, in the order they were carried out: France, UK, Morocco and Tunisia, Indonesia, and India. We also include the results from an analysis conducted on French and UK Pages in the immediate aftermath of the November 2015 terrorist attacks in Paris. We then set out the overall conclusions from these case studies, with some lessons and proposals for groups working in the field of counter-speech efforts.

CASE STUDY 1: FRANCE

Overall data

Using the process above gave us a total of 931 public Pages, from which we used the API to collect all public posts and interactions, and some other pieces of demographic data as available. On further analysis we found that there were 229 Pages that we were able to access, and that had posts on them from the last six months. Following a manual coding exercise, Demos researchers divided these Pages into the following broad groups, which best reflected the types of Pages we identified:

1. Pages which were generally about Islam
2. Pages which were anti-Islamic
3. Pages which were potentially counter-speech (for example, Pages about progressive Islam, anti-IS humour, interfaith Pages, French Muslim identity)

Table 2 Broad groups of Pages in France

Page category	Pages (total)	Page likes (average)	Posts (average)	Interactions on these Pages (average)
General Islam	191	669,178 (3,504)	6,112 (32)	92,991 (15)
Counter-speech Pages	36	398,737 (11,076)	2,039 (57)	370,473 (182)
Anti-Islam Pages	2	16,209 (8,105)	2,372 (1,186)	162,742 (69)
Total	229	1,084,124	10,523	626,206

Post content

In order to better understand the nature of posts on these Pages, Demos researchers manually read and marked up the most popular 624 posts made in our counter-speech category. From this we found 246 posts, or 39 per cent of the total, which were marked up as a form of counter-speech.

Using a system of manual coding, we determined that these 246 posts could be broken down into six broad categories of counter-speech, as per table 3 below.

Table 3: Types of counter-speech post

Content type	Description	No. of posts (/624)	% across data set	Extrapolated across total data set
National solidarity/	Expression about not letting extremists divide France	70	11%	1,157

Charlie Hebdo				
Differentiation	Stressing the importance of not confusing Islam with extremism (often quoting the Qur'an on tolerance)	13	2%	210
How to respond	Discussion about how to respond to terrorism around the world, including in France	31	5%	523
Humour/parody	Mocking or exposing the absurdity of extremism	36	6%	631
Risks of over-reaction	Raising awareness about the danger of over-reaction and Islamophobia	59	9%	947
Exposing IS	Active content showing the true character of IS	37	6%	631

'Charlie Hebdo' was selected as a category because, following the attack on the Charlie Hebdo offices, there was a significant amount of posts and Pages which referenced or used the magazine.

If extrapolated, this would give a total of 4,099 counter-speech posts across the 229 Pages we identified, over the last six months. (We assume there are many more on Facebook as a whole.)

Interactions with posts

In order to better understand popularity, we looked at which type of content was the most popular. Charlie Hebdo posts were the most shared and liked in terms of volume. However, when calculated as an average, we found that posts about how the French government should respond were more shared and liked.

Table 4: Interactions with counter-speech posts (averages in brackets)

Content type	Posts	Likes	Comments	Shares
National solidarity/ Charlie Hebdo	70	22,094 (315)	1,061 (15)	19,226 (275)
Differentiation	59	21,419 (363)	502 (9)	18,289 (310)
How to respond	37	21,111 (571)	2,359 (64)	11,795 (319)
Humour/parody	36	14,257 (396)	777 (22)	4,820 (134)
Risks of over-reaction	31	12,911 (416)	1,369 (44)	3,101 (100)
Exposing IS	13	5,457 (420)	1,178 (91)	3,748 (288)

We examined who interacted with these counter-speech posts. We found that the overwhelming majority of users interacting with this content were French users, and

with a relatively equal gender split. Similarly, the age categories of users suggested a relatively broad cross-section of age groups – overall 44 per cent were under 34 and around one in four were under 24.

Popular posts

We took the top 20 posts in each category of counter-speech and, in order to better illustrate the specific types of posts, defined for each the type of content the tone of the post and the origin of the post.

1. The ‘type of content’ category was broken down into: argument, news, denouncing violence, Hadith, news from Muslim News Network, exposing IS’s arguments and the aftermath of the Charlie Hebdo attacks.
2. The ‘tone of content’ category was broken down into: constructive, tolerant/solidarity, aggressive, humour, outrage and celebratory.³

On this analysis, we found that posts on Charlie Hebdo Pages against extremism (1,236) and anti-Islamophobic Pages (956) had the highest number of average interactions per post. As social media is heavily event-driven, it is unsurprising that the attacks on Charlie Hebdo were a topic of interest.

In terms of the tone of posts, tolerance/solidarity (1,833) posts had the highest number of average interactions per post. In terms of the content of posts, exposing IS arguments (1,514) and argument (1,501) had the highest number of average interactions per post.

Reach

We were also able to determine whether or not an interaction was made by a user who had liked the Page where the original post was posted. This shows how far content can travel beyond the people who like Pages. This suggests that content relating to how to respond to extreme Islamism, humour/parody and exposing IS are the most likely types of content to go ‘beyond’ Page followers.

Table 5: Who interacted with content?

Content type	Follower	Not a follower
National solidarity/Charlie Hebdo	72%	28%
Differentiation	67%	33%
How to respond	54%	46%

³ These categories were determined by a French analyst based on a review of the available data. A coding system was used based on ‘grounded theory’, whereby an analyst would mark up data into categories of meaning which were created based on what kind of data was found. The categories were revised iteratively, until further data no longer resulted in any further categories being created.

Humour/parody	58%	42%
Risks of over-reaction	66%	34%
Exposing IS	60%	40%
Overall	62%	38%

We also examined the average interactions from people who were not followers of a Page, in order to see what type of content ‘travelled well’.

Posts from Muslim news groups and groups extremely critical of Islam were the most likely to have interactions (likes and comments, since the API does not include shares in this measure) from people who had not liked the original Page. (In respect of the groups that are very critical of Islam, we think this reflects the fact that they are often very good at producing content which appeals to a wide audience.)⁴

In terms of the tone, celebratory/pride and tolerance/solidarity posts were the most likely to have interactions from people who had not liked the original Page. In terms of content, argument and posts exposing IS arguments were the most likely to have interactions from people who had not liked the original Page.

⁴ The ‘origin of content’ category was broken down into: Muslim faith group, Muslim News Network, anti-islamophobia group, Anti-Islam group, Charlie Against Fundamentalism group, and multifaith group. These categories were determined by a French analyst based on a review of the available data.

CASE STUDY 2: UNITED KINGDOM

Overall

We ran the same data collection effort as in the French study above, which gave a total of 355 Pages, from which we scraped all public posts and interactions, and some other pieces of demographic data as available between 24 June and 24 July 2015.

These Pages were manually marked up into ten groups, based on categories chosen by the researchers following a coding exercise, using the same method as for the France data above. The difference from the French Pages demonstrates that each country has its own pattern of counter-speech.

Table 6: UK Pages and interactions

Page category	Pages (total)	Posts (average)	Interactions on these Pages (average)	Average interactions per post, per Page
Charity	1	2 (2)	2 (1)	1
Individual/public figure	35	0	28,546 (31)	31
Local community group	14	185 (13)	1,540 (8)	8
Mosque/Islamic centre	7	89 (13)	837 (9)	9
Muslim educational organisation	15	115 (8)	1,033 (9)	9
News network	12	4 (0.3)	31,903 (48)	48
Non-religious political or social (based abroad)	47	1,246 (27)	148,022 (119)	119
Non-religious political or social (based in the UK)	60	1,474 (25)	378,195 (257)	257
Positive religious campaign	154	3,171 (21)	187,126 (59)	59
Religious group with no explicit/clear agenda	10	196 (20)	2,496 (13)	13
Total	355	6,482	779,700	

We found that there were 677,000 unique interactions on the Pages in our data and around 1.78 million Page likes overall.

Through our findings we saw that positive religious campaign Pages were the most numerous, with 43 per cent of the total. By examining how many posts had been posted on each Page, we found that positive religious campaign Pages were also the most active. These Pages posted 13,100 posts during the period (35 each, or just over one a day).

In terms of average interactions per Page, we found that non-religious political or social group Pages were by some margin the most popular, followed by individual/public figure Pages and positive religious campaign Pages.

Audience reach

We examined what type of people interacted with these Pages. Users currently in work (but ‘late’ in their careers) were overwhelmingly the most active when it came to interacting with the content across all the Pages. Non-religious political or social group Pages (based abroad) were disproportionately highly interacted with by people belonging to a high-school age group. In contrast, Mosques and Muslim educational organisations are failing to reach young people, despite being able to reach college-age Facebook users. Overall, male users were more active on these Pages. Only in Muslim educational organisations and charities were women more active.

Post content

In order to do some more detailed analysis of posts, we manually marked up the 500 posts most interacted with in the data set, from any category. In total, we found that 127, or 25 per cent of the total, were examples of counter-speech. Following a manual coding exercise, these were broken down into the categories set out in table 7.

Table 7: Types of counter-speech post

Content type	Description	No. of posts (/500)	% across total data set	Extrapolated across total data set
Exposé of IS	Specific efforts to expose IS narrative or behaviour	44	8.8%	2,123
Community event	Advertising or sharing stories about positive	17	3.4%	820

	Muslim events in the UK			
Counter-extremism campaign	Specific campaigns relating to countering extremism	21	4.2%	1,013
Moderate media	Muslim media outlets posting content	22	4.4%	1,061
Moderate religious voice	Scholars or other discussing why Islam rejects extremism/violence	13	2.6%	627
News article	Sharing stories about moderate Muslim leaders	10	1.2%	289

This found that exposés of IS were the most commonly shared type of post, followed by moderate media content and specific campaigns relating to countering extremism. This analysis shows that, if extrapolated across our data set of posts, there would be approximately 2,123 posts exposing IS on our Pages during the time period for which data were collected (24 June–24 July 2015).

Interactions with posts

In terms of the interactions with this content, exposés of IS were by some margin the most interacted-with type of posts both in volume and in average interactions per post. English-language exposés of IS require further analysis to determine the extent to which this kind of counter-speech is productive. Interestingly, counter-extremism campaigns and moderate religious voices were the next most popular types of content.

Table 8: Interactions with posts

Content type	Posts	Total interactions	Average interactions
Exposé of IS	44	574,702	13,061.4
Community event	17	11,998	749.9
Counter-extremism campaign	21	99,968	4,760.4
Moderate media	22	13,076	653.8
Moderate religious voice	13	30,794	2,368.8
News article	10	5,104	510.4
Total	127	735,642	

CASE STUDY 3: MOROCCO AND TUNISIA

Tunisia and Morocco constitute the first and second largest contributors (respectively) of foreign fighters to Syria outside of the Middle East. Moreover, IS videos and statements have repeatedly targeted the overthrow of the Tunisian and Moroccan political systems. Thus, the form and content of IS counter-speech in these two countries provides an interesting case, potentially illuminating for strategies to limit the reach of IS internationally.

Tunisia and Morocco have followed different political trajectories, especially since 2011, which has led to significant differences in political discourse and the place of IS counter-speech in it. Tunisia's experience of political opening since the overthrow of secularist autocrat Zine El-Abidine Ben Ali has led to much national-level debate over the role of Islam in society and politics. This process has paralleled – and been shaped by – an uptick in domestic terrorism, some with affinities to IS. This has included assassinations of prominent secularist figures, growing audacity of anti-state armed groups, and attacks in Tunis and elsewhere.

Tunisia's vigorous and vitriolic debate over Islam in politics contrasts starkly with mainstream discourse in Morocco. Despite Morocco's numerous currents of religious thought, the Moroccan state practises a near zero-tolerance policy on questioning the religious legitimacy of the king, which includes IS propaganda. In 2014, prominent Moroccan Islamists received direct threats from IS after they condemned the organisation as un-Islamic. At a policy level, the Moroccan state has taken prominent positions against IS, joining the international coalition fighting it and encouraging social and religious programmes to combat extremism. Overall, Morocco's relative insulation from terrorism (or at least that perception) has meant a more detached view of IS and the stakes involved in countering it.

Overall data

Using the process above gave us a total of 427 Pages where we considered that counter-speech had occurred, from which we scraped all public posts and interactions. Based on a qualitative analysis of these Pages, we found that 147 of these 427 Pages appeared to be places where counter-speech had taken place.

These Pages were manually marked up into the groups shown in table 9 by a North African Arabic-language specialist. Based on a review of the Pages themselves, we decided to divide them into these broad groups, which we felt best reflected the types of Pages we found.

Table 9: Pages in Morocco and Tunisia

Page category	Pages	Total likes	Average Page likes
Comedy/satire	16	10,987	647
News/commentary	90	93,796	1,043
Organisation Page	9	2,752	305
Political	300	229,587	765
Other	12	5,546	462
Total	427	342,669	

North Africa narratives

As table 9 shows, counter-speech may be found most often, by a large margin, in political Facebook Pages. These Pages were also the second most liked, after the news/commentary Pages (with which there was often a considerable amount of overlap). Unlike other countries being analysed, political and current affairs Pages were particularly important in Tunisia and Morocco.

The initial analysis focused on the most liked and shared instances of counter-speech. Among these posts, the dominant discourse about IS in both Morocco and Tunisia seemed to rest on two assumptions that affect how IS discourse is countered on Facebook. First, Pages with an explicit religious message tended to begin with the assumption that IS is not Islamic. Secondly, there was a widely held view that IS's growth is the result of international intelligence operations. Nonetheless, such Pages did still appear to be anti-IS in their dismissal of it.⁵

Post content

⁵ In the Tunisian context of recent terrorist attacks (in Sidi Bouzid, in November 2015), categorising Islamist speech or messages as pro-IS is sometimes used as a political tool. Thus, many secular/leftist groups/Pages in the region tend to conflate and implicate all Islamists, ranging from Nahada to slightly off-message imams to al-Qaeda, in extremism. To many, these all come under the banner of 'Daesh'. At the same time, mainstream Islamist groups both tacitly condemn and look to distance themselves from ideas/images we associate with IS, even if they do not condemn them outright. As a result of this local politicisation of IS counter-speech, some links and Pages seem to have contradictory messages – Nahada could be considered both a counter to IS-speech, and the target of counter-speech itself, at the same time.

In order to better understand the nature of posts on these Pages, Demos researchers scraped post data from them. In total, this gave us 9,436 posts taken from the week of 13 November 2015, corresponding to IS-related violence in Paris, Beirut and Sidi Bouzid, Tunisia. We took a random sample of all posts and marked up 500. From this we found 155 posts, or 31 per cent of the total, which were marked up as a form of counter-speech. Using a system of manual coding, we determined that these 155 posts could be broken down into six broad categories of counter-speech, as per table 10. These categories were chosen by the analyst based on a manual review of the posts themselves.

Table 10: Types of counter-speech post

Content type	Description	No. of posts (/500)	% across total data set (9,436)	Extrapolated across total data set
Religious messages	Religious messages condemning IS, domestic terrorism (that usually invokes IS) or violence more abstractly	4	1%	75
News about Paris attacks	Links to news articles about Paris attacks, usually presenting facts without commentary	38	8%	717
News about armed groups in Tunisia	Links to news articles about armed groups (claiming IS links) operating in Tunisia, especially but not exclusively related to the death of Mabrouk Soltani	9	2%	710
Other news	Links to news articles related to IS other than in Tunisia or France, including Morocco, Syria, etc.	10	2%	189
Denouncing violence in Tunisia	Posts that denounce violence done in IS's name in Tunisia, especially but not exclusively related to the death of Mabrouk Soltani	23	5%	434
Denouncing violence abroad	Posts that denounce IS violence, primarily related to Paris attacks	18	4%	340

Anti-Islamist messages	Posts that use IS counter-speech to condemn Tunisian Islamists, conflating the two	7	1%	132
Criticising reaction to Paris attacks/IS	Posts that criticise backlash to attacks as ‘hypocritical’; tends to be a form of counter-counter-speech, but to implicitly (and sometimes explicitly) argue for the irrelevance/un-Islamic nature of IS	36	7%	679
Questioning official story	Posts that argue that IS was not responsible for Paris attacks; also a form of counter-counter-speech that discredits IS’s legitimacy	10	2%	189
Total		155		

The Paris attacks were the most common subject of posts being produced and shared over the one-week period during which the data was collected. News about the attacks was the most popular content, followed by posts criticising the reaction to the attacks – for example that it risked a backlash against Muslims or was far greater than the reaction to the Tunisian attacks. This often implicitly or explicitly argued that IS is counter to Islamic teachings. In total, 9 per cent of the posts denounced violence in Tunisia (23 posts) or abroad (18 posts).

Some of these posts may seem like a form of ‘counter-counter-speech’, in that they were directed as a response to those showing solidarity with France. Yet they also implicitly worked as a form of anti-IS counter-speech, as there was a widespread sentiment that the ‘un-Islamic’ nature of IS is so obvious that it does not merit mention, let alone denunciation. Likewise, posts that question the official story behind IS by claiming conspiratorial links to international governments undermine the legitimacy of IS. These types of posts were much more frequent (and widely liked and shared) reactions to IS narratives than explicit condemnations of its doctrine.

Overall, Morocco had considerably less direct counter-speech than Tunisia. This probably reflects Tunisia’s experience of terrorist violence at the hands of IS imitators, and Morocco’s isolation from it. One Moroccan Facebook post linked to a TV news piece which interviewed Moroccan men in Casablanca, asking them if they

thought Paris-style attacks could be repeated in Morocco. Every respondent said they thought it was impossible or unlikely.

Interactions with posts

In order to better understand popularity, we looked at which type of content was the most popular. We did this by taking the posts above and working out the interactions they received.

Table 11: Interactions with counter-speech posts

Content type	Posts	Total interactions	Average interactions (per post)
Criticising reaction to Paris attacks/IS	36	100,804	2,800
News about Paris attacks	38	28,528	751
Religious message	4	2,128	532
Questioning official story	10	4,384	438
Denouncing violence abroad	18	5,963	331
Anti-Islamist message	7	1,776	254
Denouncing violence in Tunisia	23	3,192	139
News about armed groups in Tunisia	9	892	99
Other news	10	913	91
Total	155	148,580	

In terms of interaction volumes, we found that criticisms about the response to the Paris attacks were the most popular on average. However, a religious message was also popular, although very limited in terms of volume of posts.

Tone of posts

In order to better understand what sort of post was most popular in terms of how it was written, counter-speech posts were also categorised by tone: aggressive, celebrating victory, constructive, foreboding, humour, neutral (such as the posting of facts without comment), outrage at hypocrisy, outrage at violence, and solidarity. As above, these categories were selected by the analyst based on the findings in the data.

Table 12: Interactions by post tone

Content type	Posts	Total interactions	Average interactions (per post)
Aggressive	11	4,909	446
Celebrating victory	5	889	178
Constructive	8	2,595	324
Foreboding	8	2,760	345
Humour	13	51,330	3,948
Neutral	31	22,725	733
Outrage at hypocrisy	28	53,181	1,899
Outrage at violence	36	5,967	166
Solidarity	15	4,224	282
Total	155	148,580	

This analysis finds that humorous content is by some margin the most likely type to receive a high level of interaction per post from other users, receiving twice as many interactions as the second most popular content type when using the same measure – outrage at hypocrisy.

CASE STUDY 4: INDONESIA

Overall data

Using the process outlined above we reviewed 500 separate Pages.

An Indonesian researcher manually read and marked up these Pages. Using a system of manual coding, the researcher determined that the Pages could be divided into six categories.

In total, we judged that 382 Pages were prima facie likely places where counter-speech activity took place. Christian and anti-Islam Pages were the most voluminous. For the purpose of the rest of the analysis we discounted ‘other’ and ‘irrelevant’ Pages.

Table 13: Counter-speech Pages in Indonesia

Page category	Description	Pages
Christian	Pages discussing or promoting Christianity in Indonesia	157
Anti-Islam	Pages primarily designed as a forum for anti-Islamic discussion	72
Religion	Pages promoting or discussing cross-faith relations in Indonesia, or a split between Islamic and Christian topics	51
News Page	News Pages	50
Moderate Islam	Pages promoting a moderate interpretation of Islam	39
Anti-IS	Pages specifically countering IS terrorism (exclusive of a religious counter-argument)	13
Other	Blogs and people Pages not categorised above	41
Irrelevant	Irrelevant Pages (football clubs, popular culture, etc.)	77
Total		500

Table 14: Counter-speech Pages: popularity

Page category	Pages	Total likes	Average Page likes
Moderate Islam	39	166,562	4,271
Anti-IS	13	45,319	3,486

Religion	51	92,315	1,810
News Page	50	80,568	1,611
Anti-Islam	72	30,674	426
Christian	157	58,645	374
Total	382	474,083	

When broken down by total and average Page likes, Pages about moderate Islam and specific anti-IS Pages had the highest number of average Page likes. This suggests they are the most popular Pages on average, although they were less voluminous than all the other Page types. This in turn suggests that there is considerable scope for increasing the number of these Pages.

Christian and anti-Islam Pages, although the most voluminous, were the least popular when measured by average Page likes.

We also examined the average number of interactions on these Pages to get a better sense of engagement levels (see table 15).

Table 15: Average interactions per Page and per post

Page category	Pages	Average Page likes	Total posts	Total interactions	Average interactions per Page	Average interactions per post
Anti-IS	13	3,486	16	4	0.3	0.3
Christian	157	374	1748	6,040	38.5	3.5
Anti-Islam	72	426	669	4,359	60.5	6.5
Moderate Islam	39	4,271	4,358	31,427	805.8	7.2
News Page	50	1,611	1,015	234,351	4,687.0	230.9
Religion	51	1,810	766	3,820	74.9	5.0

In line with other research, this shows that news Page posts tended to receive the highest levels of interaction. This is because they were typically published by well-known media outlets with a wide readership. However, after removing news Pages from the analysis, we find that moderate Islam Pages had the highest number of posts, the most interactions and the highest number of interactions per Page as well as per post. The anti-IS Pages, however, appeared to be quite inactive. Although they had high numbers of likes, they did not post a lot of content and subsequently received few interactions.

Post content

In order to better understand the nature of posts, Demos researchers scraped post data from these 500 Pages. In total, this gave us 10,555 posts over the period 1 to 31 March 2016.

An Indonesian researcher manually read and marked up 500 of these posts. Using a system of manual coding, the researcher determined that the posts could be divided into six broad categories of counter-speech, as per table 16. It should be noted that, for the analysis below, 23 posts were deemed too vague, and the ‘anti-IS’ Pages were removed as the very small number of posts and interactions meant we were not able to retrieve enough data to mark up.

Table 16: Types of counter-speech post

Category	Description	Number of posts	% of posts	Extrapolated across total data set
Religious argument	Religious messages with arguments	240	48%	5,066
News reports	Links to news articles and sharing of news	126	25%	2,660
Anti-IS	Specific anti-IS content	25	5%	528
Anti-Islam	Anti-Islamic posts not engaging in religious discussion	25	5%	528
Anti-violence	Messages denouncing violence in Indonesia and abroad	22	4%	464
Government/politics	Discussions of government and politics	20	4%	422
Satire/parody	Satire and parody	11	2%	232
Other	Other content	8	2%	169
Total		477		

This analysis suggests there is a significant volume of religious discussion taking place on these Facebook Pages. Nearly half were specifically religious messages relating to Islamist radicalism, and 9 per cent were either denouncing IS explicitly or violence more generally.

The average interactions for each type of post were calculated (see table 17) in order to better understand which types of post Facebook users engaged with.

Table 17: Average interactions by post type

Category	Number of posts	Total interactions	Average interactions
Religious argument	240	23,307	97
News reports	126	101,578	806
Anti-IS	25	2,906	116
Anti-Islam	25	1,465	59
Anti-violence	22	28,087	1,277
Government/politics	20	9,249	462
Satire/parody	11	6,582	598
Other	8	1,332	167
Total	47	174,506	

Although far less voluminous than news reports or religious arguments, anti-violence posts had the highest volume of average interactions; once news was removed, this was followed by satire and parody.

To fully understand the way in which our Pages were producing posts and how Facebook users were interacting with that content, researchers cross-tabulated the post data with the Pages producing that content. Because the Pages were quite different, this offered a more nuanced understanding of which type of content was being produced on different Pages (see table 18).

Table 18: Post type broken down by Page type

	Christian	Anti-Islam	Moderate Islam	News Page	Religion
Religious argument	68	74	69	8	21
News reports	10	3	30	59	24
Anti-IS	1	0	0	2	22
Anti-Islam	3	17	0	0	5
Anti-violence	0	0	0	18	4
Government/politics	0	0	0	6	14
Satire/parody	1	2	1	4	3
Other	0	2	0	1	5

This analysis shows that much of the anti-violence and satirical content came from news Pages, rather than moderate Islam or religion Pages. Moderate Islam and religion Pages focused mainly on religious argument and news reports.

The analysis was repeated to understand the total and average interactions per post across both categories of data, Page and post. This allowed us to identify very clearly and precisely what type of content received the most interactions on average.

Table 19: Total interactions per post per Page type

	Christian	Anti-Islam	Moderate Islam	News Page	Religion
Religious argument	1,646	1,771	8,682	10,769	439
News reports	250	227	3,595	96,928	578
Anti-IS	13	0	0	2,402	491
Anti-Islam	93	1,064	0	0	308
Anti-violence	0	0	0	28,029	58
Government/politics	0	0	0	8,775	474
Satire/parody	12	40	190	6,232	108
Other	0	42	0	1,198	92

Table 20: Average interactions per post per Page type

	Christian	Anti-Islam	Moderate Islam	News Page	Religion
Religious argument	20	18	87	110	4
News reports	3	2	36	989	6
Anti-IS	0	0	0	25	5
Anti-Islam	1	11	0	0	3
Anti-violence	0	0	0	286	1
Government/politics	0	0	0	90	5
Satire/parody	0	0	2	64	1
Other	0	0	0	12	1

As above, news content was both voluminous and highly interacted-with (both on average and in total). The (relatively) high number of interactions on moderate Islamic Pages is interesting – and again suggests there is a large and relatively active group of users interested in these issues. Both anti-violence and anti-IS content were popular – but were overwhelmingly posted on news Pages rather than others.

A similar analysis was performed on the tone of posts across the different Page categories to identify those with the highest average number of interactions by the

tone of post. An Indonesian researcher analysed the tone of the posts (neutral, angry or celebratory) to examine the extent to which different tones of posts affected the engagement level with certain posts.

Table 21: Total interactions per post per Page type (tone)

	Christian	Anti-Islam	Moderate Islam	News Page	Religion
Angry	304	1,321	1,054	39,772	1,033
Celebratory	139	264	3,871	58,670	300
Neutral	1,571	1,559	7,542	55,891	1,215

Table 22: Average interactions per post per Page type (tone)

	Christian	Anti-Islam	Moderate Islam	News Page	Religion
Angry	4	13	11	406	11
Celebratory	2	3	39	599	3
Neutral	19	16	75	570	12

Surprisingly, it was neutral, non-emotive content that was the most interacted with in every category. This suggests that, contrary to received wisdom, it is not emotive content that encourages higher engagement levels.

CASE STUDY 5: INDIA

Overall data

Using the process outlined at the beginning of this report, we reviewed 500 separate Pages. After having removed all public posts and interactions we found that, based on the qualitative analysis, 312 of these 500 Pages were locations where counter-speech may have occurred. Using a system of manual coding, an Indian researcher marked up the Pages into eight categories, chosen according to the Pages themselves. Moderate religious and other religious Pages were the two most voluminous categories of Page, with approximately 100 Pages each, followed by fundamentalist Islam Pages. For the purpose of the rest of the analysis we discounted unclear, irrelevant and dead Pages.

Table 23: Groups of counter-speech Pages in India

Page category	Description	Pages
Charity/cause	Pages supporting a charity or cause	12
Education	Pages supporting an educational cause	4
Fundamentalist Islam/fanatic Muslim/Islamist fanatic/radical Islamist	Pages promoting a fundamentalist interpretation of Islam	58
Moderate religious	Pages promoting a moderate interpretation of Islam or a secular cause	100
News	News and media Pages	14
Political	Pages promoting a political cause or representing an arm of the state	13
Religious	Pages promoting a religious cause	97
Satirical/religious criticism	Pages satirical of religion	14
Total		312

Table 24: Total and average likes per Page

	Pages (total)	Total likes	Average Page likes
Charity/cause	12	23,446	1,953.8
Education	4	6,732	1,683.0
Fundamentalist Islam	58	276,770	4,771.9
Moderate religious	100	481,092	4,810.9
News	14	72,596	5,185.4
Political	13	29,346	2,257.4
Religious	97	289,968	2,989.4

Satirical/religious criticism	14	84,436	6,031.1
Total	312	1,264,386	

Based on average Page likes, satirical/religious criticism had the highest average likes per Page, followed by news Pages. However, Pages falling into the moderate religious category were the most voluminous, followed by Pages categorised as religious and fundamentalist Islam/Islamist fanatic.

Post content

In order to better understand the nature of posts, Demos researchers scraped post data from these 312 Pages. In total, this gave us 4,866 posts over the period 17 to 28 May 2016.

An Indian researcher manually read and marked up 500 of these posts. Using a system of manual coding, we determined that these posts could be broken down into six broad categories of counter-speech, as shown in table 25. These categories were determined by the researcher. It should be noted that 73 Pages categorised as ‘other’ were deemed too vague to use and were removed from the data set.

Table 25: Post content totals

Category	Description	Number of posts	% of posts	Extrapolated across total data set
Positive/moderate religious	A positive or moderate religious interpretation or opinion	332	66.4%	3,231
Extreme religious	An extreme religious interpretation, often expressing an aggressive opinion and often offensive	51	10.2%	486
News	Links to news articles and sharing of news related to IS	15	3.0%	146
Charity	Cause/campaign/protest	9	1.8%	88
Religious media	Media posts such as articles posted or shared related to religion	6	1.2%	58
Government/politics	Discussions in relation to government and politics	5	1.0%	49
Anti-IS	Specific anti-IS content	3	0.6%	29
Anti-violence	Messages denouncing violence in India and abroad	3	0.6%	29

Religious education	Religious education discussions online	3	0.6%	29
----------------------------	--	---	------	----

This analysis suggests that the majority of counter-speech Facebook content was considered to fall into the positive/moderate religious category. While not necessarily directly confronting IS, it was promoting a very different message to that proposed by IS. It is also worth noting how religious discourse plays a central role in facilitating counter-speech discussion on Facebook in India.

Interactions with posts

For the analysis below, we removed those categories with five or fewer posts. The average interactions for each type of post were calculated in order to gain a better understanding of the types of post which Facebook users engaged with the most.

Table 26: Average interactions by post type

Category	Number of posts	Total interactions	Average interactions
Positive/moderate religious	332	36,397	109.6
Aggressive/extremist/personal religious posts/opinions	51	2,885	56.6
News	15	1,249	83.3
Charity/cause/campaign/protest	9	1,127	125.2
Religious media	6	284	47.3
Total	413	41,942	

With over 12 times as many interactions as in the other categories (owing to the sheer volume of posts), it is clear that the majority of the discourse on Facebook was focused on positive and moderate religious campaigning. But it was also – along with charities and causes – the most popular on average. Both were significantly more likely to be interacted with than extremist views or religious media.

Post tone

A similar analysis was performed on the tone of posts across the different Page categories to identify those with the highest average number of interactions by the tone of post. We analysed the tone of the posts to examine the extent to which different tones of posts affected the engagement level. Table 27 shows the total number of posts in each category; table 28 shows the total and average numbers of interactions per category.

Table 27: Tone by category per post

Category	Neutral/ motivational	Happy	Angry	Defensive
Positive/moderate religious	233	96	1	2
Aggressive/extremist/personal religious posts/opinions	41	4	3	3
News	13	1	1	0
Charity/cause/campaign/protest	4	1	4	0
Religious media	6	0	0	0

Table 28: Interactions (total and average)

Category	Neutral/ motivational (average)	Happy (average)	Angry (average)	Defensive (average)
Positive/moderate religious	24,098 (103.4)	11,946 (124.4)	105 (105)	248 (124)
Aggressive/extremist/personal religious posts/opinions	1,917 (46.8)	275 (68.8)	253 (84.3)	440 (146.7)
News	1,131 (87)	84 (84)	34 (34)	0 (0)
Charity/cause/campaign/protest	247 (61.8)	613 (613)	267 (66.8)	0 (0)
Religious media	284 (47.3)	0 (0)	0 (0)	0 (0)

The analysis of tone by category per post suggests that those which were positive or happy (where the category is positive/moderate), or angry or defensive (where the category is extremist), were most likely to have a high average number of interactions.

CASE STUDY 6: POST-PARIS TERROR ATTACKS

On 13 November 2015, a series of coordinated terrorist attacks took place in Paris and resulted in the deaths of 130 people. Following the attacks, we once more collected data from the same Pages in both France and the UK in order to calculate any changes in activity. We collected this data from 13 to 21 November 2015. This allowed us to calculate the change in the averages across the two periods for both the number of posts circulated and the number of interactions those posts received.

Post-Paris attack data: France

The France data shows that during the eight days that followed the Paris attacks there was a surge in activity on some counter-speech Pages, with a tenfold increase in the number of posts being shared and a fivefold increase in the number of interactions with that content.

Interestingly, there was not an increase in activity on general Islam Pages, which suggests that there was a specific spike in counter-speech activity, as opposed to a general increase in activity across all Pages. This again owes something to the way social media is driven by current affairs.⁶ In order to gauge relative activity, the figures were averaged to activity per day.

Table 29: Pre- and post-Paris attack activity on counter-speech (France)

	Pre-attack		Post-attack		% change (posts)	% change (interactions)
	Average daily posts	Average daily interactions	Average daily posts	Average daily interactions		
General Islam	34	516.6	26.5	259	- 22%	- 50%
Counter-speech Pages	11.3	2,058.2	128.8	14,122.8	+ 1,010%	+ 568%
Total	45.3	2,574.8	155.3	14,381.8		

As might be expected, across our Pages the top ten most popular posts after the Paris attacks were all about the attacks. They were all posted by the Je Suis Charlie

⁶ None of the original five anti-Islam Pages posted during the period, so comparison is not possible in that category. This was due to Pages being either inactive or suspended by Facebook.

account (a Community Page). The posts were an assortment of images and statements that France should not be fearful or cowed. The most popular post, with 15,000 interactions, stated: ‘They have guns. Screw them. We have champagne!’ In total, these top ten posts received 66,768 interactions. It is not possible to calculate how many people would have seen these posts as a result, but it would certainly have been hundreds of thousands of users.

Post-Paris attack data: UK

We conducted the same analysis as for the French Pages, above.

Table 30: Pre- and post-Paris attack activity on counter-speech (UK)

	Pre-attack		Post-attack		% change (posts)	% change (interactions)
	Average daily posts	Average daily interactions	Average daily posts	Average daily interactions		
Charity	0.6	1.8	4.0	5.3	+ 667%	+ 292%
Individual/ public figure	9.0	20.9	4.8	100.6	- 47%	+ 481%
Local community group	13.0	4.1	6.7	11.5	- 58%	+ 284%
Mosque/ Islamic centre	13.4	5.3	2.3	6.1	- 83%	+ 116%
Muslim educational organisation	10.3	18.2	6.7	209.6	- 45%	+ 1,152%
News network	60.7	21.9	25.0	584.3	- 59%	+ 2,672%
Non-religious political or social (based abroad)	19.1	116.3	9.4	2,607.8	- 51%	+ 2,242%
Non-religious political or social (UK- based)	26.4	102.5	10.1	6,464.5	- 62%	+ 6,309%
Positive religious campaign	34.7	9.8	11.1	2,565.7	- 68%	+ 26,080%
Religious group with no explicit/unclear agenda	20.2	5.0	5.4	32.8	- 73%	+ 635%

Despite a relatively smaller number of posts being circulated over the period (with the exception being in the charity category), the level of interactions was much higher. Positive religious campaigns in particular saw a 260-fold increase in interactions with their content in the week after the Paris attacks, compared to three months earlier. The majority of the top ten posts in the UK in the week after the Paris attacks were in Arabic or Urdu. However, a key use of Facebook was to circulate the hashtag #AMessagetoISIS. Three of the top ten most interacted-with pieces of content referred to the hashtag.

The most interacted-with piece of content in English (3,222 interactions) was the following: ‘Around the world, people are showing their solidarity with a #MessageToISIS through words, art, film and more. What’s your message?’

It linked to an article on mic.com showing artists paying tribute to the victims of the attacks.

OVERALL CONCLUSIONS

Overall, the analysis undertaken suggests various patterns relating to the volume and nature of counter-speech Pages and content published on Facebook, pointing to significant similarities, as well as differences, between the countries.

Certain categories of content were published more frequently, and received greater engagement from users, than other types. Event- and country-specific contexts determine the type of counter-speech created by users, and the type of content that they were most likely to interact with. **This suggests counter-speech must be reactive and responsive to news and current affairs.**

However, it is critical to note that the strength of this causal relationship varies according to each category of counter-speech. Some categories of counter-speech content vary more between countries and across time than others. The following are categories of counter-speech that the analysis revealed to be posted with greater frequency, or as having received higher levels of user interaction, in specific countries:

- In both France and the UK, posts which exposed IS were the most numerous and the most frequently interacted with. As noted above, there is a risk to this category of counter-speech, because, when insensitive, a post is more likely to provoke anger than begin a reasoned debate.
- In Tunisia and Morocco, posts criticising the response to the Paris attacks were on average the most popular according to interaction measures, followed by posts giving news about the Paris attacks. Given the timing and the scale of the Paris attacks – which occurred shortly before we started collecting data – this was to be expected. Following the Paris attacks, it was notable that religious messages were the most popular. This was typically content denouncing IS as un-Islamic.
- In Indonesia, moderate Islam Pages published the greatest volume of posts, and received the most overall interactions, interactions per Page and interactions per post. On the other hand, explicitly anti-IS Pages appeared to be relatively inactive. Despite having high numbers of Page likes, they published content only infrequently and received few interactions.

- In India, moderate religious Pages were the most common type of counter-speech Page, and posts classified as positive/moderate religious received the highest levels of user engagement. More generally, over three-quarters of counter-speech content was deemed to fall into a religion-related category of some kind.

These patterns suggest that counter-speech content operates differently as the context in which it is produced and shared changes. For example, the types of counter-speech posted in Tunisia and Morocco are determined by country-specific contexts, and will be different to those posted in the UK or France. Some of these differences are not surprising, and can be understood by placing the data in the political and historical context of the country from which it was drawn. For example, as former French colonies, both Morocco and Tunisia have significant populations living in France, and reactions on Facebook to the Paris attacks were heavily framed by this contentious political relationship, especially as it related to discussions of IS and Islam. Differences in the way Moroccan and Tunisian Facebook users posted and interacted with counter-speech content can also be explained by the lower levels of terrorist violence suffered in Morocco than in Tunisia in recent years. Another significant difference between the countries was the role played by religious discourse in providing a platform for counter-speech content. Whereas religion-related Pages were relatively few in number and low in user engagement in France and the UK, such platforms were critical in facilitating the spread of counter-speech content in countries such as India and Indonesia.

As well as varying according to country-specific contexts, **the terrorist attacks in Paris illustrate that the volume of counter-speech varies according to event-specific contexts.** It springs into action and increases dramatically following an offline event. Following the attacks in Paris there was a 260-fold increase in posts on Pages relating to positive religious messages in the UK; in France there was a fivefold increase in posts on Pages that were explicitly counter-speech over the same period. This also suggests that following major events there is a good opportunity for groups and individuals to produce content that can connect to large numbers of users.

To observe that the way users create and interact with counter-speech content depends on country- and event-specific factors is not to suggest that none of its tones or categories resonate universally with users across different countries.

- Humorous counter-speech Pages and content received consistently high levels of engagement from Facebook users across each of the countries. In the UK and France humour was the most popular type of tone, and in India it was the most popular Page type in terms of average likes. In Indonesia, although far less voluminous than news reports or religious arguments, satire/parody posts (together with anti-violence posts) similarly received the highest number of average interactions.
- News Pages produced high volumes of counter-speech content, and received high levels of user engagement, in each of the case studies outside Europe. In Indonesia news Pages produced by far the highest number of average interactions per post, in Morocco and Tunisia they received the highest number of average Page likes, and in India they received the highest average number of Page likes.

The main conclusion drawn from this analysis is that despite a few light trends, **Facebook users create and interact with different types of counter-speech Pages and content depending on where in the world they are, and whether or not their interaction follows a terrorist event.** For those interested in expanding the potential of counter-speech on Facebook in the future, understanding, defining and encouraging counter-speech will require a different approach in each country.

There is room for more research into why specific forms of counter-speech succeed in different countries. **This suggests that there is no all-encompassing approach that covers the whole of Facebook, but rather a series of country-specific approaches for which Facebook can provide an important platform to spread messages confronting IS narratives and ideology.** In the UK, for example, most pressingly, Mosques and Muslim educational organisations are failing to reach out to young people via social media. These groups might consider using some of the more popular content types in order to reach a wider audience. Although anti-violence and satirical content is very popular in Indonesia, not much of it is produced, meaning its reach is limited. More explicit anti-IS and anti-violence posts from moderate Islam and Christianity Pages might reach a large, and different, audience.

Finally, this analysis can tell us something about how and why content is shared online, and the sort of reach that it has. However, it cannot say much about whether

and how that content is understood and acted upon in the real world. This remains an area for further research.

Demos – Licence to Publish

The work (as defined below) is provided under the terms of this licence ('licence'). The work is protected by copyright and/or other applicable law. Any use of the work other than as authorized under this licence is prohibited. By exercising any rights to the work provided here, you accept and agree to be bound by the terms of this licence. Demos grants you the rights contained here in consideration of your acceptance of such terms and conditions.

1 Definitions

- a 'Collective Work' means a work, such as a periodical issue, anthology or encyclopedia, in which the Work in its entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole. A work that constitutes a Collective Work will not be considered a Derivative Work (as defined below) for the purposes of this Licence.
- b 'Derivative Work' means a work based upon the Work or upon the Work and other pre-existing works, such as a musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Work may be recast, transformed, or adapted, except that a work that constitutes a Collective Work or a translation from English into another language will not be considered a Derivative Work for the purpose of this Licence.
- c 'Licensor' means the individual or entity that offers the Work under the terms of this Licence.
- d 'Original Author' means the individual or entity who created the Work.
- e 'Work' means the copyrightable work of authorship offered under the terms of this Licence.
- f 'You' means an individual or entity exercising rights under this Licence who has not previously violated the terms of this Licence with respect to the Work, or who has received express permission from Demos to exercise rights under this Licence despite a previous violation.

2 Fair Use Rights

Nothing in this licence is intended to reduce, limit, or restrict any rights arising from fair use, first sale or other limitations on the exclusive rights of the copyright owner under copyright law or other applicable laws.

3 Licence Grant

Subject to the terms and conditions of this Licence, Licensor hereby grants You a worldwide, royalty-free, non-exclusive, perpetual (for the duration of the applicable copyright) licence to exercise the rights in the Work as stated below:

- a to reproduce the Work, to incorporate the Work into one or more Collective Works, and to reproduce the Work as incorporated in the Collective Works;
- b to distribute copies or phonorecords of, display publicly, perform publicly, and perform publicly by means of a digital audio transmission the Work including as incorporated in Collective Works; The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. All rights not expressly granted by Licensor are hereby reserved.

4 Restrictions

The licence granted in Section 3 above is expressly made subject to and limited by the following restrictions:

- a You may distribute, publicly display, publicly perform, or publicly digitally perform the Work only under the terms of this Licence, and You must include a copy of, or the Uniform Resource Identifier for, this Licence with every copy or phonorecord of the Work You distribute, publicly display, publicly perform, or publicly digitally perform. You may not offer or impose any terms on the Work that alter or restrict the terms of this Licence or the recipients' exercise of the rights granted hereunder. You may not sublicense the Work. You must keep intact all notices that refer to this Licence and to the disclaimer of warranties. You may not distribute, publicly display, publicly perform, or publicly digitally perform the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this Licence Agreement. The above applies to the Work as incorporated in a Collective Work, but this does not require the Collective Work apart from the Work itself to be made subject to the terms of this Licence. If You create a Collective Work, upon notice from any Licensor You must, to the extent practicable, remove from the Collective Work any reference to such Licensor or the Original Author, as requested.
- b You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital filesharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works.

C If you distribute, publicly display, publicly perform, or publicly digitally perform the Work or any Collective Works, You must keep intact all copyright notices for the Work and give the Original Author credit reasonable to the medium or means You are utilizing by conveying the name (or pseudonym if applicable) of the Original Author if supplied; the title of the Work if supplied. Such credit may be implemented in any reasonable manner; provided, however, that in the case of a Collective Work, at a minimum such credit will appear where any other comparable authorship credit appears and in a manner at least as prominent as such other comparable authorship credit.

5 Representations, Warranties and Disclaimer

A By offering the Work for public release under this Licence, Licensor represents and warrants that, to the best of Licensor's knowledge after reasonable inquiry:

- i Licensor has secured all rights in the Work necessary to grant the licence rights hereunder and to permit the lawful exercise of the rights granted hereunder without You having any obligation to pay any royalties, compulsory licence fees, residuals or any other payments;
- ii The Work does not infringe the copyright, trademark, publicity rights, common law rights or any other right of any third party or constitute defamation, invasion of privacy or other tortious injury to any third party.

B except as expressly stated in this licence or otherwise agreed in writing or required by applicable law, the work is licenced on an 'as is' basis, without warranties of any kind, either express or implied including, without limitation, any warranties regarding the contents or accuracy of the work.

6 Limitation on Liability

Except to the extent required by applicable law, and except for damages arising from liability to a third party resulting from breach of the warranties in section 5, in no event will licensor be liable to you on any legal theory for any special, incidental, consequential, punitive or exemplary damages arising out of this licence or the use of the work, even if licensor has been advised of the possibility of such damages.

7 Termination

A This Licence and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this Licence. Individuals or entities who have received Collective Works from You under this Licence, however, will not have their licences terminated provided such individuals or entities remain in full compliance with those licences. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this Licence.

B Subject to the above terms and conditions, the licence granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different licence terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this Licence (or any other licence that has been, or is required to be, granted under the terms of this Licence), and this Licence will continue in full force and effect unless terminated as stated above.

8 Miscellaneous

A Each time You distribute or publicly digitally perform the Work or a Collective Work, Demos offers to the recipient a licence to the Work on the same terms and conditions as the licence granted to You under this Licence.

B If any provision of this Licence is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this Licence, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

C No term or provision of this Licence shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

D This Licence constitutes the entire agreement between the parties with respect to the Work licensed here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This Licence may not be modified without the mutual written agreement of Demos and You.

Counter-speech – which argues, disagrees or presents an opposing view – is a potentially important way to deal with extreme or offensive content online. It is fast, flexible and responsive, capable of dealing with extremism from anywhere, in any language, and retains the principle of free and open public spaces for debate. However, it is also likely that it is not always as effective as it could be; and some types of counter-speech could potentially even be counter-productive.

This report sets out the summary findings of phase II of this project, examining how speech which challenges extreme Islamist narratives is produced and shared. Countries included in this study are the UK, France, Morocco, Tunisia, Indonesia and India.

Jamie Bartlett is Director of the Centre for the Analysis of Social Media at Demos.

Alex Krasodonski-Jones is a Researcher with Centre for the Analysis of Social Media at Demos