



THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■

Journalism In The Age of Populism and Polarisation: Insights from the Migration Debate in Italy



in partnership with



Ca' Foscari
University
of Venice

CORRIERE DELLA SERA

About LSE Arena

Based at the Institute of Global Affairs (IGA) within the London School of Economics and Political Science (LSE), Arena is an innovative programme dedicated to overcoming the challenges of disinformation.

The aim of the Arena programme is to use high-quality research, analysis and evaluation to create effective best practices that can then be disseminated to journalists, public diplomacy teams and civic groups.

Arena seeks creative ways to counter the menace of unreality, stop the spread of hatred and division and foster a fact-based discourse that enhances security, enables democracy and entrenches trust. Its experimental research projects involve journalists, data analysts and activists who seek to both understand disinformation campaigns and reach audiences impacted by them.

About Ca' Foscari University of Venice

Established in 1868 as the first Business School in Italy, Ca' Foscari University of Venice is now one of Italy's best universities in the country. Located in the heart of Venice, Ca' Foscari has a national and international outstanding reputation for academic excellence in both teaching and research and has signed hundreds of international cooperation agreements with universities from all over the world, in order to promote mobility, research and training. The Laboratory of Data and Complexity coordinated by Walter Quattrociocchi is based in the Department of Environmental Sciences, Informatics and Statistics. Through an interdisciplinary and cross-methodological approach, its research team investigates mass social dynamics, (mis)information spreading, and the evolution of collective narratives in online social media.

Contributors

LSE

Peter Pomerantsev, Director, LSE Arena Program

Anne Applebaum, Professor in Practice

Sophia Gaston, Senior Visiting Fellow

Nicole Fusi, MSc International Migration and Public Policy student

Zachary Peterson, Independent Consultant

Ca' Foscari University of Venice

Walter Quattrociocchi, PhD, Head of Laboratory of Data and Complexity

Fabiana Zollo, PhD, Assistant Professor, Research Fellow in the Venice Center for Humanities and Social Change

Ana Lucía Schmidt, PhD, Postdoc

Antonio Peruzzi, Research collaborator

Corriere della Sera

Beppe Severgnini, Deputy Editor-in-chief

Andrea Federica de Cesco, Staff Writer

Davide Casati, Digital Editor

Supported in part by a grant from the Open Society Foundations.
Arena's contributions were supported by LSE Arena program funds.

I Introduction

Across the world – from the US to the UK, from Europe to South Asia and Latin America - politics and media are stuck in a spiral that incentivizes divisive rhetoric, hyper-partisanship and disinformation. The main beneficiaries of this spiral are a generation of politicians, often labeled ‘populist.’ What connects them is not their policies, but their ability to capture attention. They use intentionally inflammatory language and controversial ideas in order to focus attention on themselves and to divide electorates into crude wars of ‘us’ versus ‘them’. Sometimes they are supported by online squadrons of social media militias, as well as intensely biased publications with low editorial standards, that help push their messages. Traditional, ‘quality’ media that aspire to accuracy and balance find themselves caught in a catch-22: a failure to report on these politicians will result in accusations of censorship, but challenging them risks accusations of ‘fake news’ by the politicians themselves.

Invariably, ‘quality’ media end up reporting on these politicians, reinforcing their agendas and language. But the way they report is also shaped by the peculiar financial incentives of the current moment. As traditional media, especially print publications, struggle to survive, they find themselves operating in an online advertising market that favors ‘clickable’ stories – namely, the scandalous, personality-driven, polarising content that this new breed of politicians provide. The ad-tech market in which they operate is deeply un-transparent. It does not distinguish between publications that have a public service-spirited mission and editorial standards on the one hand, and partisan or deliberately misleading web-sites and social media pages on the other. The ad-tech market is in turn powered by the algorithmic architecture of the internet and social media, which is skewed towards highly emotive, hyper-partisan material that appeals to existing confirmation biases and feeds more ‘shares’ and ‘likes’. The very architecture of the internet rewards news organizations and individual users who take ever more extreme and polarising positions¹ - an algorithmic logic that in turn encourages the ‘populist’ politicians. They, in their turn, create content that mainstream media feels it is obliged to describe....and so the spiral spins on.

In this research project we asked whether it was possible to report on the polarising subject of migration in ways that didn’t play into this spiral. Over the course of a year Italy’s newspaper of record, Corriere della Sera, produced different types of content about migration, using different mediums (video and multimedia, text and infographics) and different techniques (fact-checking, human interest and constructive news among others). Corriere created over a hundred pieces of content, while the LSE and Venice analysed engagement with them on Corriere’s Facebook page. We wanted to go beyond the usual commercial metrics of ‘likes’ and ‘shares’, and looked for what we called ‘public service spirited’ metrics of engagement that could help us understand what sort of content fostered a more civil debate, smoothed polarisation and enhanced trust towards accurate content. Our research is not meant to be final, and our sample size was limited. Our aim was to encourage thinking and conversation about an editorial framework that avoids the polarising games of ‘populist’ politicians. Instead, we ask whether it is possible to create content that is both popular, accurate and encouraging of constructive, thoughtful engagement. This is hard enough at the best of times, but to do so in the context of the malign spiral we have described it is nearly impossible. There is little incentive to produce public-service-spirited content for a publication that has to compete in the current ad-tech market and internet architecture, especially, as in Italy, where subscription models have not caught on. In order to avoid playing into the media strategies of ‘populist’ politicians and a new public sphere to emerge the whole negative spiral has to be dismantled.

¹ The relationship between disinformation and social media has also been extensively explored by Walter Quattrociocchi of Ca’ Foscari University of Venice <https://www.pnas.org/content/113/3/554>, <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0181821>

How can this be done? Breaking the polarisation spiral will require, first and foremost, greater public oversight of the algorithms and social media models that currently encourage extremism. Such regulation is already well on its way in Europe, and public pressure is growing in the US. It is important any regulation is not focused on censorship and 'take-downs', but on encouraging accurate content, high editorial standards and providing people with a balanced diet of content rather than encasing them in 'echo chambers'. Breaking the polarisation spiral will also mean reforming the ad-tech system, to create incentives for content that is not just 'clickable', but also fosters more thoughtful engagement. As this report shows, it is possible to consider 'public service spirited' metrics of success, but the environment has to radically change if media are to focus on content that fosters a politics that revolves less around populist personalities, disinformation and polarisation.

Peter Pomerantsev, Director, LSE Arena Program

| Executive Summary

In 2018, LSE Arena, Ca' Foscari University of Venice and the Italian newspaper *Corriere della Sera* analysed the engagement of *Corriere* readers with content touching on the controversial and polarising topic of migration in Italy. The purpose was to address one of the most difficult problems in journalism today, which can be summed up in four related questions:

- Which types of journalism intensify polarisation, and which reduce it?
- How can one best communicate facts?
- How can we foster constructive engagement?
- Are there ways to avoid playing into the media strategies of “anti-establishment” politicians who make purposefully controversial statements in order to dominate the national debate, and then attack media who criticise them as “enemies of the people” or purveyors of “fake news”?

Studying *Corriere*’s public Facebook page, the project compared the quantity of content about migration with the numbers of real migrants actually entering the country; examined the sentiment of articles and comments on the *Corriere* Facebook page; looked at engagement metrics according to the quantity and proportion of Likes, Shares and Comments; and included an Annotation Analysis of over 20,000 comments to measure levels of

- Toxic speech (as a measure of civil discourse);²
- Criticism of media source (*Corriere della Sera*), i.e. the presence of comments directly criticising the publisher/journalist/newspaper as a source of information;
- Anti- and pro-migration comments.

Traditional media are more and more accused by populist politicians of being “fake news” and part of “the establishment”. In that context, we use “criticism of media source” as a metric to analyse whether and how it can enhance trust while providing fact-driven content

We analysed the content according to the medium (purely visual, text, multimedia, infographics), technique (data-driven/fact-checking, editorial, human interest, constructive news, pop culture, news with context and plain news) and theme (crime, refugee crisis, ethnic tensions, culture and society).

1 - The Political and Media Context

- Though the numbers of migrants arriving in Italy dropped drastically during 2018, the volume of content about the subject increased. This coincided with the arrival of a new government in June 2018, and the concerted efforts of Interior Minister Matteo Salvini to keep the issue high on the agenda through controversial policies and statements.
- Content about migration gets more engagement than other topics. The comments use more negative language than the average comments on *Corriere*’s Facebook page.
- Content that features Salvini achieves significantly more engagement than any other content on the subject of migration.

2 For a definition of terms, see Annex A (Methodology).

- On Twitter we can see how the debate about migration has moved from three disparate communities to two, namely “traditional media” followers versus “alternative media” supporters of the two “anti-establishment” parties Lega and Movimento 5 Stelle (the Five Star Movement). We appear to be seeing a polarisation of the media space not between “left” and “right”, but between “traditional” and “anti-establishment” media.

2 - On Styles of Journalism

- The technique that elicits the fewest number of comments critical of the media source (in this case *Corriere*) as a provider of information is straightforward, impartial and unemotional news reporting providing context. We take the metric “criticism of media source” as a way to examine levels of trust.
- Human interest stories, particularly about groups of people, get strong negative pushback, eliciting a high number of anti-migration comments, many comments using toxic language and a lot of criticism of the media source (*Corriere*). However, it should be noted that criticism is softer when a story is about the plight of a single immigrant than about groups.
- Constructive news – content that offers practical solutions – inspires less toxic language about migration than comment/editorial pieces. This offers an interesting comparison as these two techniques offer different ways to further the policy debate about migration; we used the degree of toxic language as a way to measure more or less civil debate.
- A data-driven approach elicits high levels of criticism of media source and high levels of anti-migration comments but not necessarily toxic language.
- Editorial pieces get the most engagement overall, while pieces with pop-culture references get the most likes.
- Fact-checking content inspires more comments that are critical of the media (in this case *Corriere*) as a reliable source of information, and draws out more anti-migration comments than regular news articles.

3 - Engagement by Format

- Forms of media that push statistics to the forefront, such as infographics, can provoke debate and discussion but receive significant pushback from anti-migration voices and elicit high levels of criticism of the media (in this case *Corriere*) as a reliable source of information.
- Video produces the lowest level of criticism of media source (*Corriere*) as a reliable source of information, potentially because “seeing is believing” (this assumption would need to be explored further in qualitative studies) or because anti-migration voices become quiet when faced with video evidence, or possibly because it is the most passive form of engagement (video generally receives a low level of comments).
- Multimedia pieces – video, text and photo combination – receive strong and supportive engagement. Multimedia pieces also perform relatively well in terms of criticism of source, suggesting relatively high trust for multimedia stories.

Conclusions and Recommendations

1 - Editorial Choices Need to Take into Account How to Improve Trust, Smooth Extreme Polarisation and Promote Civil Debate

In an environment where trust in traditional journalism is fragile, where populist politicians and “alternative” news sites are constantly pushing inaccurate narratives and provocative policies to deliberately polarise and toxify public discourse, newsrooms need to be more aware of the potential impact of their content on the quality of engagement. There is no single formula or magic bullet, and while our research should be seen in the context of other works in the field and future offline studies, our findings indicate that:

- Impartial, accurate reporting and providing context draws the least amount of criticism of the media as a reliable source of information (a proxy for institutional trust). Including photos and video together with text in a multimedia format appears especially effective, perhaps because it presents so much evidence that the report becomes harder to refute.
- Journalists’ instincts are often to reach for human interest stories to personalise and build empathy. However, human interest stories, and especially those about groups of people rather than individuals, can encourage strong negative comments from readers and high levels of criticism of media. We would need deeper research to understand why, but one reason could be that people feel they are being emotionally manipulated into taking a political position through such coverage.
- Infographics, fact-checking and a data-driven approach elicit strong pushback from audiences who do not agree with the data, and strong criticism of media. They are not necessarily a quick fix to convince of “the truth”.
- Putting forward strong opinions and policy prescriptions around polarising issues such as migration will inevitably elicit pushback and provoke an often toxic debate. However, a “constructive news” approach (namely news stories that include proposals for policy solutions) elicits a less toxic debate than regular editorial and comment content.

2 - Public Service-Spirited Media Metrics and Other Research Priorities

Our project explored online engagement metrics that try to reflect a civil, evidence-driven debate. For example, we looked at levels of toxic language (as a measure of more or less civil discourse) and criticism of journalists and the publication (*Corriere*) as a source of information (as a proxy for trust). Much more research is needed to define engagement metrics that can determine whether media content contributes to a more deliberative and evidence-based debate. Newsrooms need to look beyond “likes and

shares” to examine the quality of engagement they inspire and its impact on the public debate.

While this study only concerned itself with online reactions, a more holistic approach is necessary, one that partners long-term offline sociological surveys. It is important to measure engagement holistically, in terms of both the platform metrics and also public opinion. We need to be able to see the short-, medium- and longer-term impacts of engagement with media content on social views, trust and ideologies.

Our study focuses on how audiences engage with various journalistic techniques and mediums. Such research needs to be partnered with insights into how language and framing influence audience engagement. As the LSE Report *The European “Migration Crisis” and the Media: A Cross-European Content Analysis* shows, European media tend to describe migrants in terms of nationality and age. What would change, for example, if one were to describe them according to profession? Or focus on the political and military conflicts that produce migration rather than just focusing coverage on migrants entering Europe?

3 - Expose and Challenge an Internet Architecture that Skews Against Balanced, Accurate Content

Media that seek to uphold traditionally understood media standards of objectivity, accuracy and fairness are playing on an internet field that is easily gamed by actors using inorganic amplification in a non-transparent manner, and employing other forms of virally deceptive behaviour to push inaccurate, biased and hyper-partisan. Journalists and civil society must not only counter this with accurate and balanced content on contentious themes, but expose and challenge the internet architecture that enables it. Particular attention should be paid to an ad-tech infrastructure that currently encourages advertising to flow towards domains that promote inaccurate content, conspiracies and hate speech.³ Such domains can also pose as journalism while being closely allied to political campaigns.

3 Melford, Clare and Fagan, Craig, “Cutting the Funding of Disinformation: The Ad-Tech Solution”, Global Disinformation Index (2019). Available at: https://disinformationindex.org/wp-content/uploads/2019/05/GDI_Report_Screen_AW2.pdf.

| Background

1 - The Modern Reporter's Dilemma

A new generation of media savvy, “anti-establishment” and “populist”⁴ politicians have become skilled at making repeated, provocative and factually inaccurate statements on controversial themes such as migration or crime. They have learned to use social media to amplify these statements, but their strategy makes use of traditional media too. Broadcasters and reporters who are committed to the highest journalistic ethics also amplify their themes and narratives, even when they know them to be false or deliberately polarising, because they feel they have no choice but to report statements and policies from those who are running for office or exercising power, even while they know that efforts to debunk such statements will be less efficient than the statements themselves. Paradoxically, in seeking to tell the truth about this new breed of politician, traditional media risk further spreading the falsehoods and polarising narratives that they promote. One scholar goes so far as to argue that European media “have contributed to the legitimization of the issues, keywords and communication styles typical of populist leaders”.⁵

The dilemma is made more acute because this new generation of politicians has also learned that it can benefit by attacking and denigrating mainstream journalists, creating antagonism between themselves and what they call “fake” media. As Sophia Gaston and Peter Harrison- Evans argue in their study *Mediating Populism*:

“Populist politicians themselves have typically adopted a somewhat schizophrenic relationship to the press, simultaneously looking to attack and co-opt the traditional news media. On the one hand, major media organisations are prime targets within their wider anti-establishment critique, while on the other, they still represent the best route to public profile.”⁶

This forces mainstream journalists into an even deeper paradox: not only are they repeating and broadcasting falsehoods even when debunking them, they are repeating and broadcasting rhetoric and sentiments that actually serve, in the long run, to destroy their own credibility.

But ignoring this new breed of politician, or refusing to report their provocative or dishonest statements, is not a good option either. On the contrary, it can further reduce the credibility of traditional media – because if they do not report on these politicians, then others will. Non-traditional internet news outlets, which are often highly partisan in nature, as well as social media, entertainment channels and tabloids, will report information if “serious” journalists will not. During the 2016 presidential election in the US, for example, Donald Trump’s tweets pushed messages about the alleged dishonesty of

4 As Gaston and Harrison-Evans write in *Mediating Populism*, “there is some scholarly convergence around these two core aspects of populism – its (often exclusionary) construction of ‘the people’, and its antielitist critique of the ‘establishment’. [...] Perhaps most significantly from the perspective of this research, populism has also been defined as a style of political communication, and one that seeks to frame public and political discourse along the anti-establishment lines described above.” Gaston, Sophie and Harrison-Evans, Peter, *Mediating Populism* (Demos, 2018), pp. 7–8. <https://demosuk.wpengine.com/wp-content/uploads/2018/03/Full-Report-Mediating-Populism-March-2018-Demos-1.pdf>.

5 Mazzoleni, G., “Populism and the Media”, in D. Albertazzi and D. McDonnell (eds), *Twenty-First Century Populism: The Spectre of Western European Democracy* (London: Palgrave Macmillan, 2008), pp. 49–66.

6 Gaston and Harrison-Evans, *Mediating Populism*, p. 7.

his rival, Hillary Clinton, as well as about “mainstream media” bias that was allegedly covering up that dishonesty.⁷ Further amplification came through bots, trolls and search engine optimisation techniques, which help what were previously even the most fringe groups to push their agendas and frames.⁸ Traditional news media that did not at least report on this conversation were at risk of becoming irrelevant.

On top of this are the financial incentives for covering “populist” politicians and polarising issues. The recent rapid technological advances have pressured traditional journalistic business practices; the shift to online publication has led to a loss of advertising revenue and subscription income.⁹ Increasing financial pressure has restricted the news media’s ability to play the traditional role as “gatekeepers”. In France, for example, this has “weakened their ability to function as a public sphere informed by diverse and broad voices, which had supported reasoned and critical discourse”.¹⁰ Given that news outlets across the democratic world are almost universally suffering financially, the temptation to publish popular, controversial and emotional content is clear. The simplification of complex political ideas and the storytelling techniques favoured by “populist” politicians – personalisation, polarisation and confrontation – can resonate quite deeply, and have been shown to spread quickly. Online media outlets especially have a high affinity for the rhetorical influence of populists because both aim for the “quick kick/click” with a broad audience.¹¹

2 - Migration: A Challenging Topic

The above is particularly pertinent for coverage of migration and the refugee crisis. Stories about immigration, for example, can be linked to issues of national security, crime and loss of sovereignty. One study has shown how the Danish news media regularly use a conflict perspective (“foreigners in conflict with ‘x’ culture”) to tell stories about immigration, for example, in such a way as to make them more dramatic.¹² The LSE report *The European “Migration Crisis” and the Media: A Cross-European Content Analysis* concludes that “the European press systematically framed refugee and migrant arrivals to European shores as a ‘crisis for Europe’”.¹³

Journalists rely on framings to communicate complicated stories to their audiences. These framings can also construct dichotomies through which audiences understand an issue.¹⁴ Journalistic frames of irregular or illegal migration – most commonly associated with the arrival of migrants across the Mediterranean Sea into Italy, rather than, say, someone overstaying on a tourist visa – has an impact on citizens’ perceptions not only of a “distant other” but also of themselves as “belonging to and protected by, the nation state”.¹⁵ In other words, the issue is framed as an existential threat.

7 Lee, Jayeon and Xu, Weiai. (2017). “The more attacks, the more retweets: Trump’s and Clinton’s agenda setting on Twitter”. *Public Relations Review*, 44: 2 (2018), pp. 201–03.

8 Marwicz, Alice and Lewis, Rebecca, “Media Manipulation and Disinformation Online”, Data & Society Research Institute (2017), <https://apo.org.au/sites/default/files/resource-files/2017/05/apo-nid135936-1217806.pdf>.

9 Dell’Orto, Giovanna and Birchfield, Vicki L. (eds), *Reporting at Southern Borders: Journalism and Public Debates on Immigration in the U.S. and the E.U.* (London: Routledge, 2013), p. 6.

10 Benson, Rodney, “What Makes for a Critical Press? A Case Study of French and US Immigration News Coverage”, *International Journal of Press/Politics* 15: 1 (2010): pp. 3–24.

11 Esser, Frank, Stepinska, Agnieszka and Hopmann, David Nicolas, “Populism and the Media: Cross-national Findings and Perspectives”, in T. Aalberg, F. Esser, C. Heinemann, J. Strömbäck and C. d. Vreese (eds), *Populist Political Communication in Europe* (London: Routledge, 2016), p. 1.

12 Esser, F., Stepinska, A., & Hopmann, D. N. (2016). Populism and the Media. Cross-National Findings and Perspectives. In T. Aalberg, F. Esser, C. Reinemann, J. Strömbäck & C. d. Vreese (Eds.), *Populist Political Communication in Europe*. Routledge, 365–380.

13 Chouliaraki, Lilie; Georgiou, Myria and Zaborowski, Rafal, *The European ‘Migration Crisis’ and the Media: A cross-European press content analysis*, LSE Project Report (2017), p. 2.

14 Jacobsen, Christine M. (2015), “Communicating Irregular Migration”, *American Behavioral Scientist*, Vol. 59 (2015), p. 895.

15 Ibid.

Though one cannot draw a direct causal relationship between the media and politicians' framings and public perceptions, perhaps it goes some way to explain the disconnect between ordinary Italians' fears of immigrants locally and national-level anxiety. According to a 2018 report by More in Common, only 18% of Italians viewed the impact of immigration on Italy positively, while 57% felt its overall impact had been negative. The report's authors explained as follows:

"Negative sentiment about immigration has been heightened by concerns about security, the perceived loss of control of Italy's borders, and the failure of authorities to manage migration effectively. Many feel that there are too many people arriving for the country to handle the situation well, that Italy has been left alone in this crisis and that because of its economic circumstances Italy cannot afford to welcome more migrants or refugees."¹⁶

However, according to a survey conducted by Ipsos for Corriere della Sera in December 2018, what concerns Italians at a national level is different at a local level. Foreigners are – at a local level – associated with caretakers, with the children who attend school with their own children, or even with relatives and in-laws.

Given these framing problems, news media have often reached for human interest stories as a way to provide some insight into migrants' stories and reasons for travelling. However, this has had the unfortunate effect of reducing the entire migration situation in Europe to individual agency rather than a series of interrelated and systemic factors. Indeed, the type of long-form narrative favoured by some US media outlets allows readers to experience "in vivid detail what it would be like to live through" a migrant's ordeal, but it teaches the reader "very little about why this is happening" or "what might be done to help them".¹⁷ This has made the issue "ripe for partisan divides" rather than the cross-party and transnational policies that the issue demands.¹⁸ Such partisan divides are also reinforced by the way the debate about migration plays out in the media.

In a transnational study which examined framing patterns in the US, France and Norway, researchers claimed that major actors in the immigration debate such as "political parties and government bureaus" use their "high coordinated communications apparatuses" to ensure "their representatives stay 'on message' when discussing immigration in the news".¹⁹ Journalists often use competing viewpoints to try to provide balance to a story, and well-known and professionalised actors – political parties, NGOs, pressure groups – are easier to access than migrants themselves. This increasing professionalisation of immigration activism, "combined with the media's tendency to assign recurring narrative roles to actors" (which results in actors being "stuck" in their positions, unable to negotiate publicly for fear of retribution), means that the immigration debate across Europe and the US is "less a back-and-forth negotiation of competing ideas than a series of stand-alone statements repeated by fundamentally opposed actors". In the US, France and Norway, "immigration news ... seemed to lack a sustained dialogue between opposing advocate viewpoints", offering a "shouting match rather than a developing debate".²⁰

As we can see from this brief introduction, the modern reporter is surrounded by a barrage of Catch-22 situations.

16 Dixon, Tim, Hawkins, Stephen, Heijbroek, Laurence, Juan-Torres, Miriam and Demoures, François-Xavier, *Attitudes towards National Identity, Immigration and Refugees in Italy*, More in Common (2018), p. 9.

17 Benson, Rodney, "Why Narrative is Not Enough: Immigration and Genres of Journalism", in Dell'Orto and Birchfield (eds), *Reporting at Southern Borders*, p. 77.

18 Dell'Orto and Birchfield, *Reporting at Southern Borders*, p. 7.

19 Benson, Rodney and Wood, Tim, "Who Says What or Nothing at All? Speakers, Frames and Frameless Quotes in Unauthorized Immigration News in the United States, Norway, and France", *American Behavioral Scientist*, 59: 7 (2015), pp. 818–19.

20 Ibid.

... Provoked into relaying controversial positions by self-scandalising politicians – who in turn attack the credibility of “mainstream media” when one reports on them;

... Wanting to give depth and context – but having to compete with “alternative” news sites that exploit social anxieties and simplistic framings;

... Trying to balance differing viewpoints – but finding they are just descending into a partisan “shouting match”;

... Straining to narrate human interest stories – only to be attacked for choosing a side.

And all this while working in an environment where consumers can choose to follow a plethora of “alternative” news sites and stories, and can tailor their choice even further to suit their biases. Journalists hoping to present migration as a complicated situation with roots in multi-layered structural imbalances are fighting a difficult battle.

3 - *Corriere della Sera* and the Chronology of the Migration Debate in Italy

This research project looks at the case Italy’s newspaper of record, *Corriere della Sera*, and examines both its coverage of the migration topic in Italy as well as the public reaction to that coverage. The subject of migration was chosen by *Corriere* itself. As the digital editor of *Corriere* explained to the LSE and Ca’ Foscari University of Venice in early 2018, journalists at the publication unanimously felt that this was going to be one of the most contested political debates over the coming year. That insight proved to be correct.

The migration topic had its origins in a series of real-world events. Thanks to the war in Syria and the collapse of the Libyan state, the refugee crisis exploded between 2014 and 2017, when more than 100,000 people arrived in Italy annually. The issue was further pushed to prominence by “populist” political parties, leaders and hyper-partisan publications. Right from the start of the crisis, Matteo Salvini, leader of the nationalist Lega Nord party, railed against the perceived preferences granted to refugees over Italians. “Guess who comes last when it comes to housing, labour, and healthcare”, went a Lega Nord slogan in 2015,²¹ with posters showing migrants overtaking ethnic Italians in a queue.



Figure 1. Lega Nord poster in 2015.

21 Freschi, Federica, “La voce del leader: strategie comunicative per la creazione del panico morale e del nemico nella pagina Facebook di Matteo Salvini”, University of Milano-Bicocca (2015).

Lega Nord and the hyper-partisan publications that support the group focused their messaging on the supposed deterioration of areas where refugees lived, the threat to security due to an alleged increase in crime rates, the spread of diseases, and the fear of Islamic fundamentalism. They also sought to debunk the idea that refugees or asylum seekers are escaping genuine persecution.²²

Salvini has successfully pushed his agenda on social media, which he uses more actively than any other politician in Italy. An analysis²³ of his October 2017 Facebook feed shows him well in front of other politicians in terms of followers and engagement. He had a fan base of nearly 2 million in 2017, almost double that of his nearest rivals, then Prime Minister Matteo Renzi and Five Star leader Luigi Di Maio. In 2019 his Facebook fan base has nearly doubled. More than any other politician, he deployed anti-elite rhetoric in his messages, focused on themes of nationalism and ethnicity and the “ostracising of others”.²⁴

During 2017, this “alternative” and social media pressure was starting to get real political results. In February of that year, the European Border and Coast Guard Agency, usually known as Frontex, suggested that some NGOs were inadvertently helping refugee smugglers by rescuing the refugees they had launched in small boats into the Mediterranean. This story was picked up and amplified by the populist Five Star Movement as well as the Lega Nord, and it was heavily promoted on their social media channels. The then Minister of the Interior Marco Minniti felt pressured to act by this campaign.²⁵ Partisan, online-only news sites and Facebook pages were garnering so many “hits” that the government could not ignore the issue. In July 2017 Minniti decided to make NGOs sign a code of conduct. Many withdrew their boats as a result. Minniti’s actions allowed both the Lega and Five Star to claim victory: online viral posts which targeted emotional responses, and often distorted information, had achieved a solid policy outcome.²⁶

Salvini followed this with a decision, at the start of 2018, the year of an Italian general election, to rebrand his Lega Nord as “Lega” in its attempt to extend its appeal from beyond the north of the country and embrace the whole nation, with migration as his party’s main unifying message. Throughout this period the agenda on migration were being set by partisan groups and political actors.

22 Ibid.

23 Mazzoleni, Gianpietro and Bracciale, Roberta, “Socially mediated populism: the communicative strategies of political leaders on Facebook”, *Palgrave Communications*, 4: 50 (2018), pp. 1–10. DOI: 10.1057/s41599-018-0104-x

24 Ibid., p. 4.

25 Miles, Tom, “Dozens drown on Libya migrant boat; Italy impounds rescue ship,” Reuters (2019). Available at: <https://uk.reuters.com/article/uk-libya-security-migrants/dozens-drown-on-libya-migrant-boat-italy-impounds-rescue-ship-idUKKCN1SG20K>.

26 Binnie, Isla and Denti, Antonio, “Aid groups split over Italy’s new rules for migrant rescues,” Reuters (2017). Available at: <https://uk.reuters.com/article/uk-europe-migrants-italy-ngo/aid-groups-split-over-italys-new-rules-for-migrant-rescues-idUKKBN1AG2FR>.

Corriere della Sera, Data and Methodologies

Our study asked the following questions:

1. Did the quantity of stories about migration reflect actual peaks and troughs of migration flows into the country – or was the agenda being set by other factors (in other words, was the information cycle actually related to real-world events)?
2. Which mediums and techniques encouraged the most engagement and what was the nature of that engagement? For example: are there patterns in how audiences engage with human interest stories or data-driven stories? Video or text? What is the role of “constructive news” and fact-checking?
3. Which mediums and techniques helped to reduce polarisation, minimise distrust in media and generate constructive (or at least non-toxic) discussion?

The research does not aim to offer any final answers to the challenges journalists face in an environment where their credibility is under attack, and where a new breed of politician has learnt to skilfully amplify polarising narratives. The study IS limited to one topic (migration) in a given time frame (2017-18), on one country (Italy) and one language (Italian) that allows for a more isolated analysis (unlikely we get non-italian in our data given the language restriction, as compared to english or spanish). Thus the results can hardly be considered universal. Our analysis is meant to help identify further research, and advance the efforts of journalists to communicate accurate content and inspire an evidence-based, civil discourse. Can traditional news media remain faithful to core values, producing accurate content on difficult subjects, without increasing polarisation? Is there a way for them to report on important politicians without furthering the agendas and frames they promote? How can they debunk falsehoods without further amplifying them? How can they retain trust when constantly being accused of purveying “fake news”? Can we encourage more constructive engagement with factually accurate content?

In an era of attention scarcity, “post-truth” politics and disinformation, this is a difficult task. But we hope that the trends we observed will help the conversation among traditional news media and other communicators about how to cover not only the migration topic but other contentious issues also used by “alternative” news sources and “populist” politicians.

1- About the Newspaper

Corriere della Sera is Italy’s newspaper of record. Though perceived as tilting to the centre right because of its traditional readership, represented by the Milanese industrial bourgeoisie, it has a broad audience across the traditional political spectrum, one that is very divided in its views of migration.

In 2018, the NGO More in Common conducted a polling and segmentation analysis of attitudes to immigration and refugees in Italy and found, among other things, that *Corriere* is read prolifically by groups on different sides of the migration debate. More specifically, these include audiences which More in Common describes as Catholic Humanitarians, who are more “pro” migration; the “Security Conscious”, who are fearful of the influx of migrants; and “Disengaged Moderates”, who are neutral on the issue. While the More in Common study found that there are some groups in Italy who cannot be persuaded to think differently on the subject, both “Catholic Humanitarians” and the

“Security Conscious” are interested in reading fact-based material that challenges their preconceptions, even on this highly contentious topic.²⁷

2 - About the Data

The project uses data that can be divided into two main datasets: the Historical Dataset and the Experimental Dataset. The project collected anonymised data from Facebook users and public entities, such as Facebook pages, interacting with the *Corriere* Facebook page, as well as anonymised data of users interacting with the public Twitter accounts of the main national newspapers in Italy. The initiative has been duly notified by *Corriere* on its Facebook page and the collected data have been managed in agreement with GDPR EU 679/2016.²⁸

The Historical Dataset includes data from two social media platforms, Facebook and Twitter, each covering a different time period and scope. The Facebook data include all the posts on the *Corriere della Sera* Facebook page published from 1 January 2017 to 31 December 2018. It also includes all the reactions to these posts and all public comments. The Twitter data include all the tweets made between March 2018 and December 2018 on the official Twitter account of the main Italian national newspapers.²⁹

The Experimental Dataset, on the other hand, consists of a set of specific posts on Facebook, ranging from March 2018 to December 2018. During this period, the journalists from *Corriere della Sera* posted migration-related stories on Facebook, led by the news content at the time of publication, employing several different techniques and mediums. As soon as a piece was uploaded, the *Corriere* team alerted project partners at Ca' Foscari University of Venice, who in turn tracked audience reaction to the pieces. This process was broad by design in order to give us the greatest insight into audience reaction in the initial stages of the project.

The Experimental Dataset consists of a set of 165 Facebook posts, published over the course of 2018, referring to specific migration content on the *Corriere della Sera* website testing different **mediums**, **techniques** and **themes**. This dataset includes 25,179 anonymised comments that were labelled for a more in-depth analysis.

The Experimental posts were at the same time split according to their **target audience** – that is, any specific Facebook users targeted on the social media platform. The main experimental analysis works on non-targeted content (114 posts) – that is, standard Facebook posts; while some special case studies (51 posts) considered targeting users who fit an anti- or a pro-migration profile.

3 - About the Methodology of the Experimental Dataset

Each piece of content published by *Corriere* was classified by **medium**, **technique** and **theme**.

The **mediums** of communication were:

1. Purely Visual: video reports and photo galleries. Video was usually shot in an un-emotional, “news” style or was from amateur video. It regularly featured footage

27 Dixon, Tim, et al., *Attitudes towards National Identity, Immigration and Refugees in Italy*, p. 9.

28 For further information, please see Annex B.

29 For a comprehensive list please see Annex A.

from on board refugee ships. Popular video stories included an immigrant saving a young child from falling off a balcony.

2. Multimedia: usually longer news pieces with a mix of video, photo and text.
3. Article and Infographics: popular posts included polls about Salvini's decision to close Italian ports to NGO boats carrying migrants, and a study showing that Italy is the European country with the biggest difference between reality and perception regarding the number of migrants.
4. Plain Articles: text with a photo. This was the largest amount of content.
5. Quiz: there was only one quiz, including questions about some of the most common beliefs on immigration and public health.

The posts were also classified according to their migration topic into one of four possible main **themes**. These should not be seen as necessarily exclusive categories, as there was often overlap between them.

1. Crime: in which migrants can be both victims and perpetrators.
2. Refugee Crisis: articles about the journey of refugees trying to reach Europe, and political decisions and declarations related to it.
3. Ethnic Tensions: content related to racism and the often difficult coexistence of Italians and foreigners. One popular story, for example, was about a bus driver accused of deliberately refusing to stop a bus he was driving for migrants at a bus stop, and who was later defended by Lega when threatened with dismissal from his job.
4. Culture and Society: more analytical and background pieces about the social and cultural fabric of the places and the people that are dealing with migration issues. For example, in one article a journalist compares a politically correct approach to discourse on migration with Matteo Salvini's aggressive and divisive rhetoric; in another piece a sociologist explains why it is not true that immigrants are stealing Italians' jobs.

Finally, the posts can be classified according to the **technique** used when writing their corresponding article:

1. News Report: content reporting in an impartial and objective way what happened or what somebody said.
2. Article with Context: articles that go into more depth about a specific situation or contextualise it. For example, a story about an NGO worker describes the condition of minors on board a refugee ship, and another article carries a profile of the magistrate who investigated Salvini about the *Diciotti* saga (see below, p. 28).
3. Editorial: articles in which experts or journalists give their opinion or comment on a particular event/situation. Popular pieces included an editorial about the migrant shot to death in Calabria, and a journalist's comparison of a moral, politically correct approach in writing about migration with Matteo Salvini's tendency to attack everyone and be divisive.
4. Constructive News: content providing a possible solution to migration issues. These included, for example, a project in Milan that helps migrants and unemployed Italians get a job; or an interview with a sociologist who suggests humanitarian corridors and channels for legal migration are a solution to avoid the arrival of so many migrants to Italy.
5. Data-driven Pieces: these included fact-checking stories, as well as more analytical pieces on, for example, the data about migrant arrivals to Italy, or an audience poll about Salvini's decision to close Italian ports to NGO boats carrying migrants.
6. Human Interest – Individual: content that focuses on the adventures of individuals, often refugees. Examples included the story of Mamoudou Gassama, a migrant without a residence permit who risked his life to rescue a child who was going to

fall from a building, or the story of a Cameroonian female immigrant rescued by a Spanish NGO after being adrift for two days.

7. Human Interest – Groups: same as above but focused on groups. These included, for examples, the stories of different passengers on a migrant boat, or testimony from humanitarian workers regarding the psychological conditions of the minors at the Greek refugee camp.
8. Stories with Pop-culture References: articles related to popular characters such as football players, singers or actors. Popular content included a story about the four Italian athletes of African descent who won the Women's 4 x 100m relay in the Mediterranean Games, or an article about a basketball player who was aboard a boat that rescued a migrant who had been adrift for two days.

This classification system allowed the team to check that no one type of story was over-represented. However, by the very nature of newsrooms and the project, there was little the team could do to limit the number of “simple news stories”. These pieces of content were the type of standard journalistic article published by *Corriere* to announce news.

The most-engaged-with posts in the experiment were classified as “simple news stories”. Indeed, “simple news story” performance on Facebook so outpaced the median engagement data for every other type of content that it is necessary to control for its popularity. This is unsurprising as the “simple news story” variant on a particular piece of news was generally the first piece published by *Corriere* on that particular topic. Every subsequent piece suffered from no longer being “new”.

As tends to be the case on social media platforms, a few pieces accounted for a large share of the overall engagement on the posts measured. In this case, the top ten most-engaged posts in terms of Likes, Comments and Shares (taken separately) accounted for half (or more) of all engagement measured for all 114 posts.

- The top ten most-liked posts accounted for 52% of all likes for the total of 114 non-targeted articles.
- The top ten most-commented on posts on Facebook accounted for 45% of all comments for the total of 114 non-targeted articles.
- The top ten most-shared posts on Facebook accounted for 59% of total shares for the 114 non-targeted articles.

Migration-related stories rarely make the news for a positive reason, and stories with a more negative sentiment – reporting news that could be considered “bad” such as corruption or violence – were consistently the most engaged-with content on Facebook. However, when controlling for the weight of viral news stories, simple news stories – regardless of whether they were “good” or “bad” – tended not to over- or under-perform relative to any other technique.

Our research should be considered within an environment where journalists and editors have more opportunity to take audiences into account as they make difficult choices about how to present content. “Web analytics and social media have opened new channels to re-discover the audience,” write Raul Ferrer-Conill and Edson Tandoc Jr. “The audience has been making its way into the newsroom by contributing to journalistic content, by distributing and sharing news, or simply by interacting with news websites, leaving their digital footprints that are then considered in editorial decisions.”³⁰

30 Ferrer-Conill, Raul and Tandoc, Edson C., Jr., “The Audience-Oriented Editor: Making sense of the audience in the newsroom”. *Digital Journalism*, 6: 4 (February 2018), p. 436.
<https://www.tandfonline.com/doi/full/10.1080/21670811.2018.1440972>.

Historical Analysis of the Migration Topic between 2017 and 2018

The migration crisis of 2014-17 was really the culmination of a longer process. Migrants had been finding ways to cross the Mediterranean, and to enter Italy by other means, for many years. But several changes – specifically the collapse of Libya, which made it easier for migrants to take boats from the Libyan coast, and the war in Syria – quickly escalated the crisis. Still, to many readers, it may have seemed as if large numbers of migrants suddenly appeared from nowhere.³¹

Though it did not act immediately, in 2017 the centre-left Italian government did strike a deal with the Libyan government and the mayors of the most important Libyan cities to reduce migrant flows. Though some criticised the deal as inhumane,³² it substantially cut the numbers of people entering Italy. In August 2017, for example, the cumulative number of refugees arriving in Italy was 81.6% lower than it had been the previous year.³³ By the beginning of 2018, migrant arrivals by sea were also lower than the year before and they kept falling; the year ended with a drop of about 80% compared to 2017, and 87% compared to 2016.³⁴

Given this reduction, one might have expected the coverage of migration to decrease gradually as well. But our analysis of *Corriere*'s output shows the opposite. The number of articles mentioning migration on its Facebook page was already on average 7.85% of total content posted in 2017. But that figure rose to 11% in the second quarter of 2018, and then peaked at 14% in the third quarter of 2018, finishing the year at around 11% of all articles posted. So even as migration numbers were falling dramatically, the amount of coverage was rising.³⁵

³¹ Source: Italian Ministry of the Interior.

³² Ibid.

³³ Ibid.

³⁴ Ibid.

³⁵ This trend was not confined to *Corriere della Sera*, either. An analysis of coverage of migration in *La Stampa*, *La Repubblica*, *Corriere della Sera* and *Il Giornale*, *Avvenire* showed that the significant drop in migrant arrivals in the second half of 2017 and in 2018 did not coincide with a proportional decrease in the number of daily front-page headlines on immigration. Whereas in 2015 there was a headline for every eight migrants who landed on Italian shores, in 2018 the ratio was one headline for every two migrants. Again, Salvini himself appears to have been a driving force behind keeping migration at the top of the political agenda: the most common keyword used in relation to immigration in 2018, accounting for 8.2% of migration headlines, was "Salvini". The frames through which migration is discussed have shifted slightly, but remain focused on the frame preferred by the Lega and Five Star: threats to the Italian nation-state. In 2018 there was a reduction in the news related to migrant reception (17% of the total, compared with 54% in 2015) and an increase in those concerning migrant flows (47% vs 23% in 2015); news about society and culture remained stable (18%), while items on crime and security (11%), terrorism (5%) and economy and labour (2%) decreased. Over the course of 2018 migration news also grew on TV, which, according to the Osservatorio Demos-Coop, is the first information source for Italians.

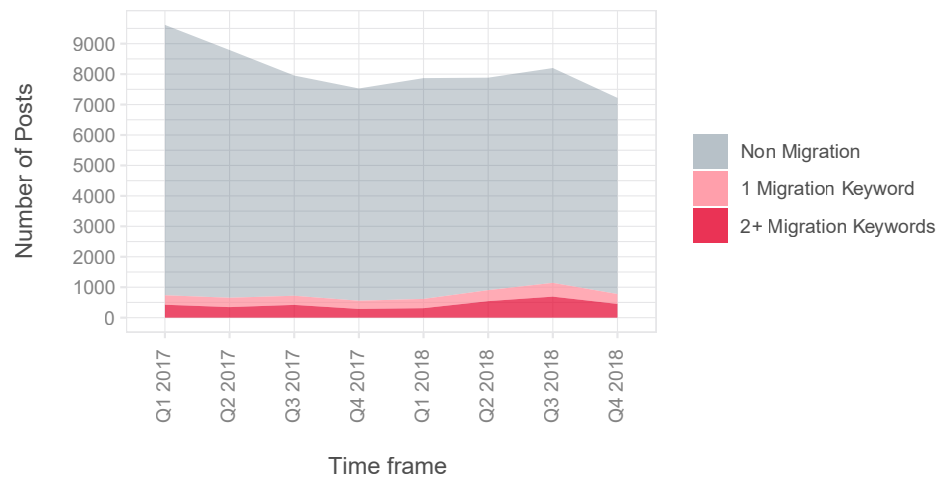


Figure 2. Number of posts, between 2017 and 2018, breaking down how many posts (and their corresponding article) contain words related to migration.

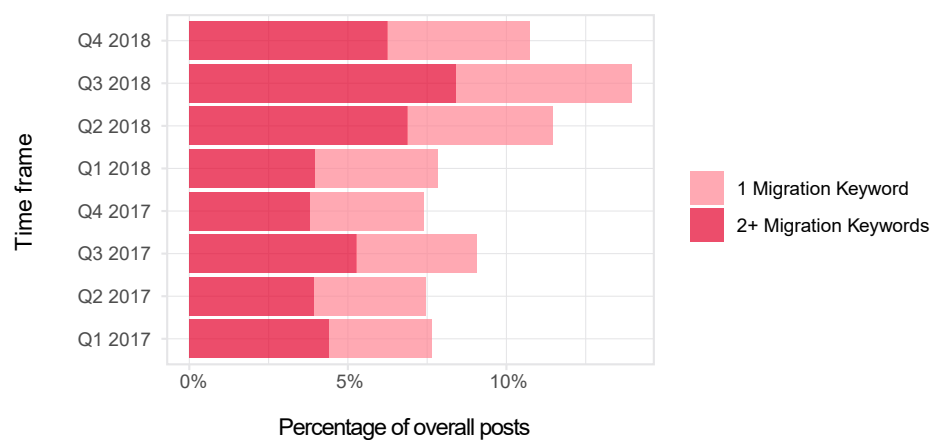


Figure 3. Percentage of content each quarter that contains words related to migration.

As the quantity of articles rose, so did the degree of reader engagement around them. In fact, articles about migration always have higher likes, shares and comments than other topics. Whereas the average number of shares for non-migration topics in 2017–18 was about 32, and the average number of comments about 27, for migration-related topics the averages were 57 shares and 94 comments. In the third quarter of 2018, there was a significant spike in this number, with articles about migration averaging 127 comments and almost 68 shares. The comments on posts about migration also present vocabulary of much more negative sentiment when compared to the non migration content.

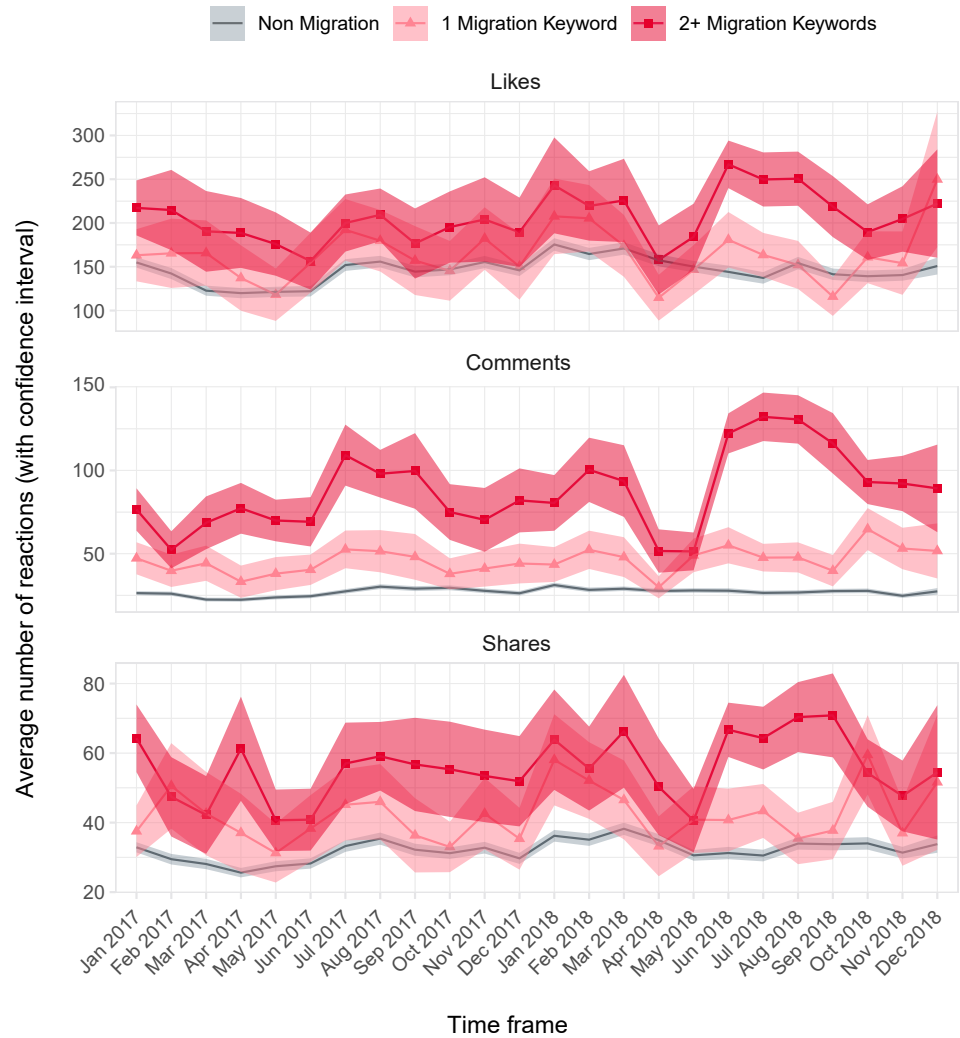


Figure 4. Average number of likes, comments and shares given to the posts during each month of 2017/18 with 95% CI. Breakdown is done by considering if the posts and its article contains 0, 1 or 2+ words related to migration.

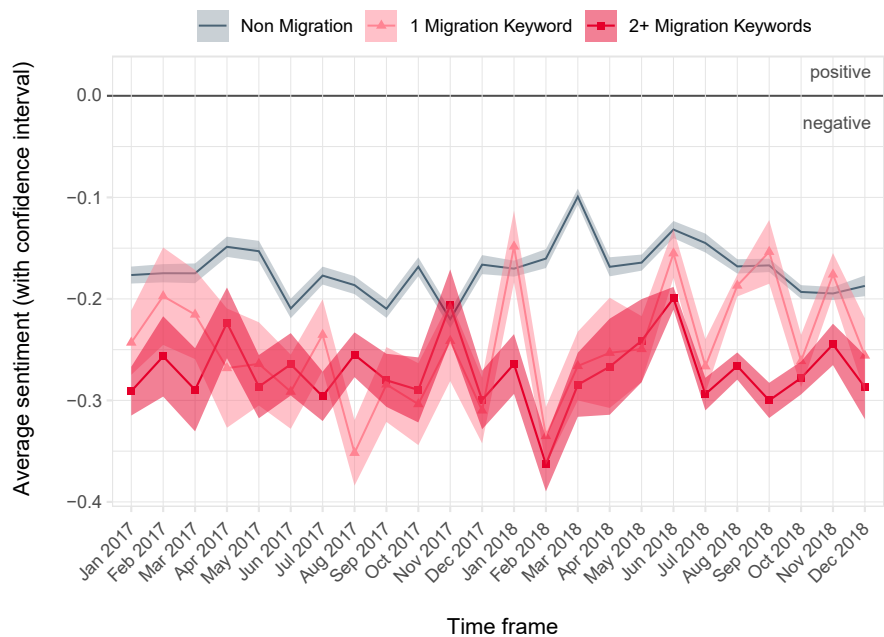


Figure 5. Average sentiment of the comments given to the posts during each month of 2017/18 with 95% CI. Breakdown is done by considering if the posts and its article contains 0, 1 or 2+ words related to migration. +1: positive, 0: neutral, -1: negative

What was the reason for this? June 2018 saw the arrival of a coalition government in Italy including Matteo Salvini, the leader of Lega, as Minister of Interior. And in that job, Salvini sought, at all costs, to keep the migration topic high on the agenda. He announced the shutdown of Italian ports to a migrant rescue ship for the first time, despite the fact that migrant arrivals were by then very few in number. He took the same measure against all NGO boats seeking to disembark migrants in Italy, claiming that Libya – against the evidence – was a safe port.³⁶ This decision sparked an international crisis and a fierce internal debate. In August 2018 Salvini prevented 190 refugees aboard the Italian coastguard ship *Diciotti* from disembarking for several days. By the end of 2018 his anti-migrant narrative had culminated in the so-called “security decree” (approved by the Italian parliament at the end of November), which clamped down on asylum rights and hardened security measures.

Salvini’s actions and rhetoric had an enormous impact on media content about migration in 2018. When we used the Watson Natural Language Understanding Service (an IBM tool that uses natural language processing techniques) to extract entities from the articles, Salvini’s was the only name that appeared repeatedly. At the same time, two of the most clicked-on pieces of content we tracked over the course of the project were directly focused on Salvini. The second most-clicked-on piece in our dataset is from June 2018, right after the formation of the coalition government, when the Italian government refused port of entry to the ship *Aquarius* and the 600 refugees on board. *Corriere* covered the *Aquarius* story in at least six different articles from different angles with different headlines, but the one which gained by far the greatest combination of shares, likes, comments and other engagement metrics had Salvini in the headline with his uncompromising assertion that “Our ports are closed to *Aquarius*”.

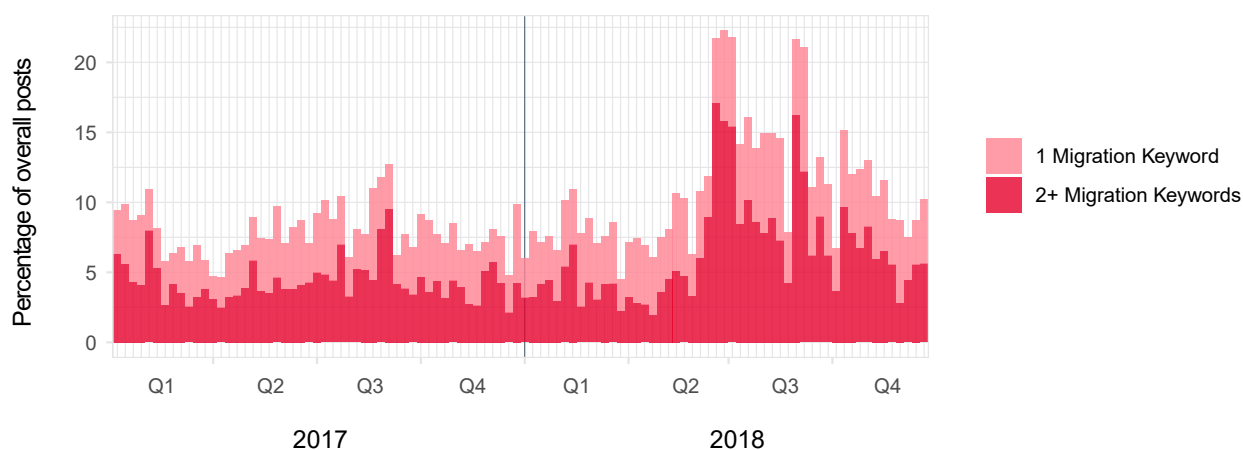


Figure 6. Percentage of content each week that contains words related to migration.

Salvini again drove engagement in August 2018, when it was made public that he was under investigation by a prosecutor in Catania for blocking entry for the *Diciotti*. Salvini took to his Facebook page, with its 3.7 million followers, and recorded a 20-minute video in which he railed against the EU and a host of other familiar targets. As of 19 December 2018, his original post had 2.8 million views, nearly 153k comments, and was approaching 70,000 shares. This is as “viral” as a politician can get in the modern social media environment, especially given that the content is in Italian and made solely for an Italian audience.

As the figures below show, since early 2018 Salvini features in a considerable percentage of the migration content on Facebook posts, and those that feature him get considerably more engagement.

³⁶ Source: Italian Ministry of the Interior.

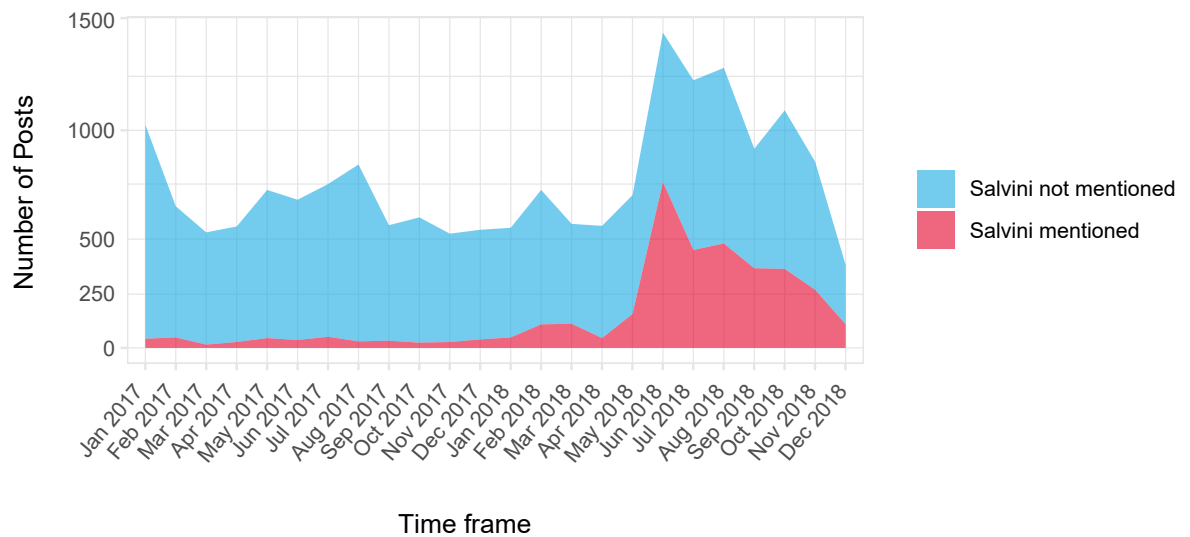


Figure 7. Number of posts that contain migration keywords between 2017 and 2018 and the breakdown on how many mention Salvini in the post or its corresponding article.

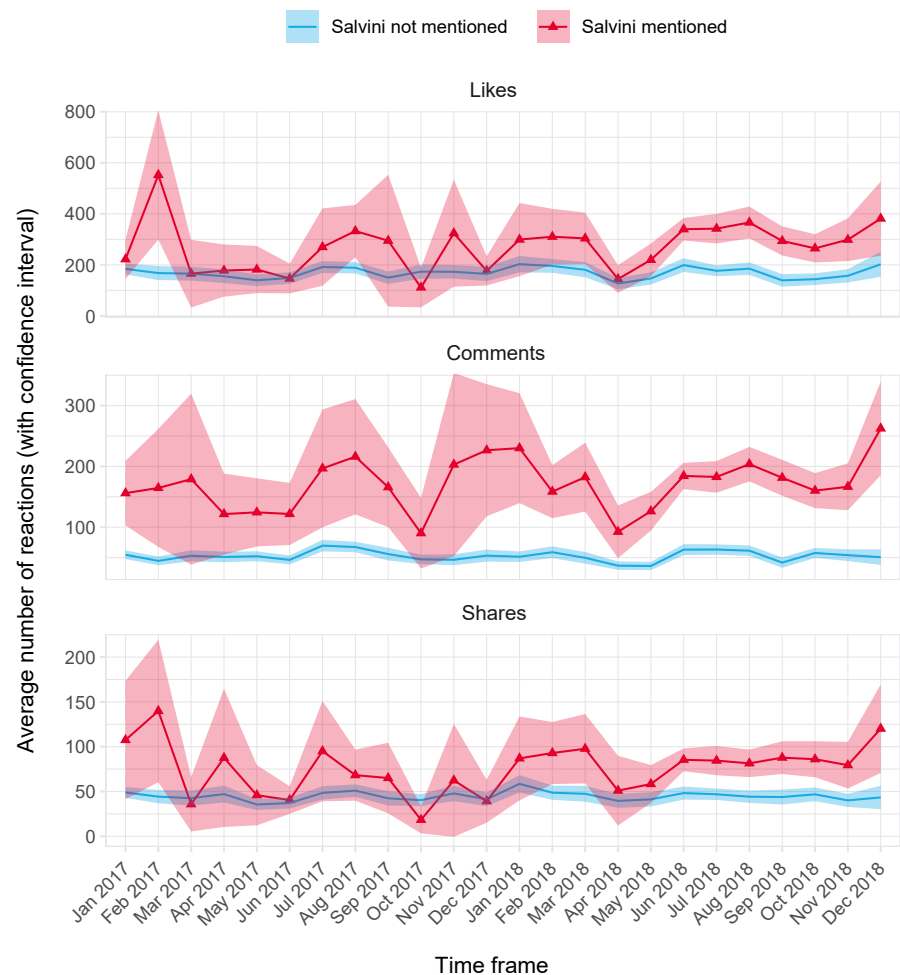


Figure 8. Average number of likes, comments and shares given to the posts during each month of 2017/18 with 95% CI. Breakdown is done by considering if each migration post and its corresponding article mentions Salvini or not.

Corriere covered the *Diciotti* story in ten Facebook articles, but the one covering his own Facebook video hugely outstripped all other articles in terms of engagement, and ended up being the single most “engaged-with” Facebook post in the period we were measuring.

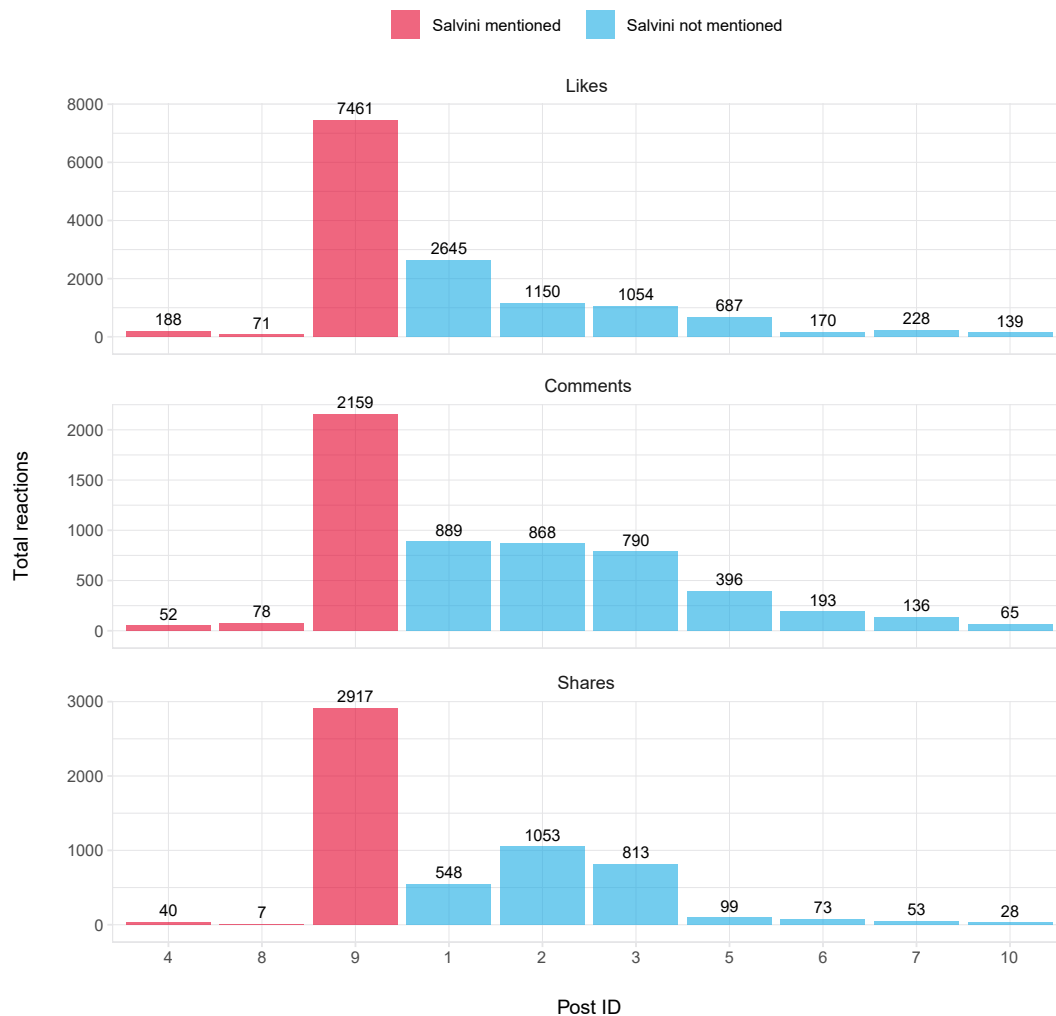


Figure 9. Total of likes, comments and shares received by the different posts that talked about the Diciotti story. The posts are numbered in order of appearance and color coded to distinguish between those that mention Salvini and those that don't.

While Salvini and the new government managed to keep the migration topic high on the agenda, partisan online news sites have also been changing the overall dynamics of the conversation. Our network mapping of the migration topic on Twitter over 2018 shows how it was dominated by users retweeting traditional news media. By June, when the coalition government was formed, there was a shift from users retweeting from media aligned with either Lega or M5S to users retweeting from both, thus merging the two communities. During the same period the level of engagement with the issue on Twitter also increased.

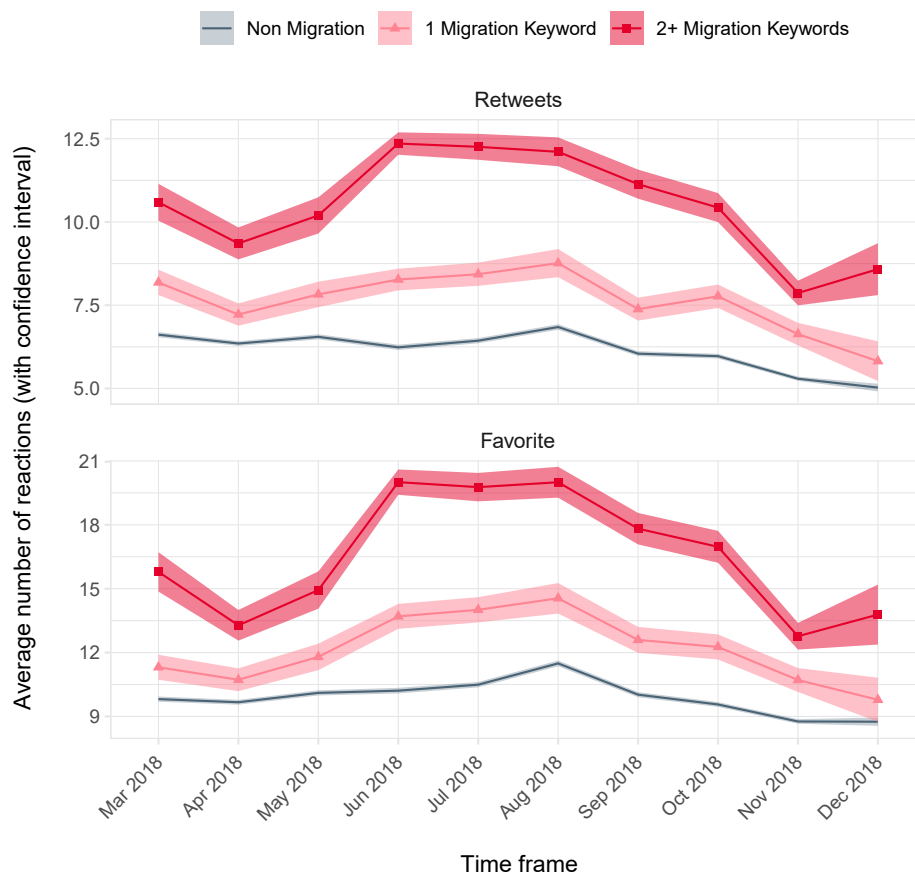


Figure 10. Average number of retweets and favourites given to the different tweets published by the national newspapers between March 2018 and December 2018 with 95% CI. Breakdown is done by considering if the tweet and its article contain 0, 1 or 2+ words related to migration.

Our network mapping of the migration topic on Twitter over 2018 also shows that the communities of sentiment about the topic changed over time. Before the formation of the coalition government in June, the discussion around migration on Twitter consisted of three echo chambers (or communities): those allied with mainstream national news, those aligned with the far right, and those aligned with Five Star. After the coalition's formation the Five Star community merged with the right wing, suggesting a country divided into two polarised news communities. This change persisted until November 2018 and can easily be traced in the retweeting patterns of users in the Twitter communities that emerged.

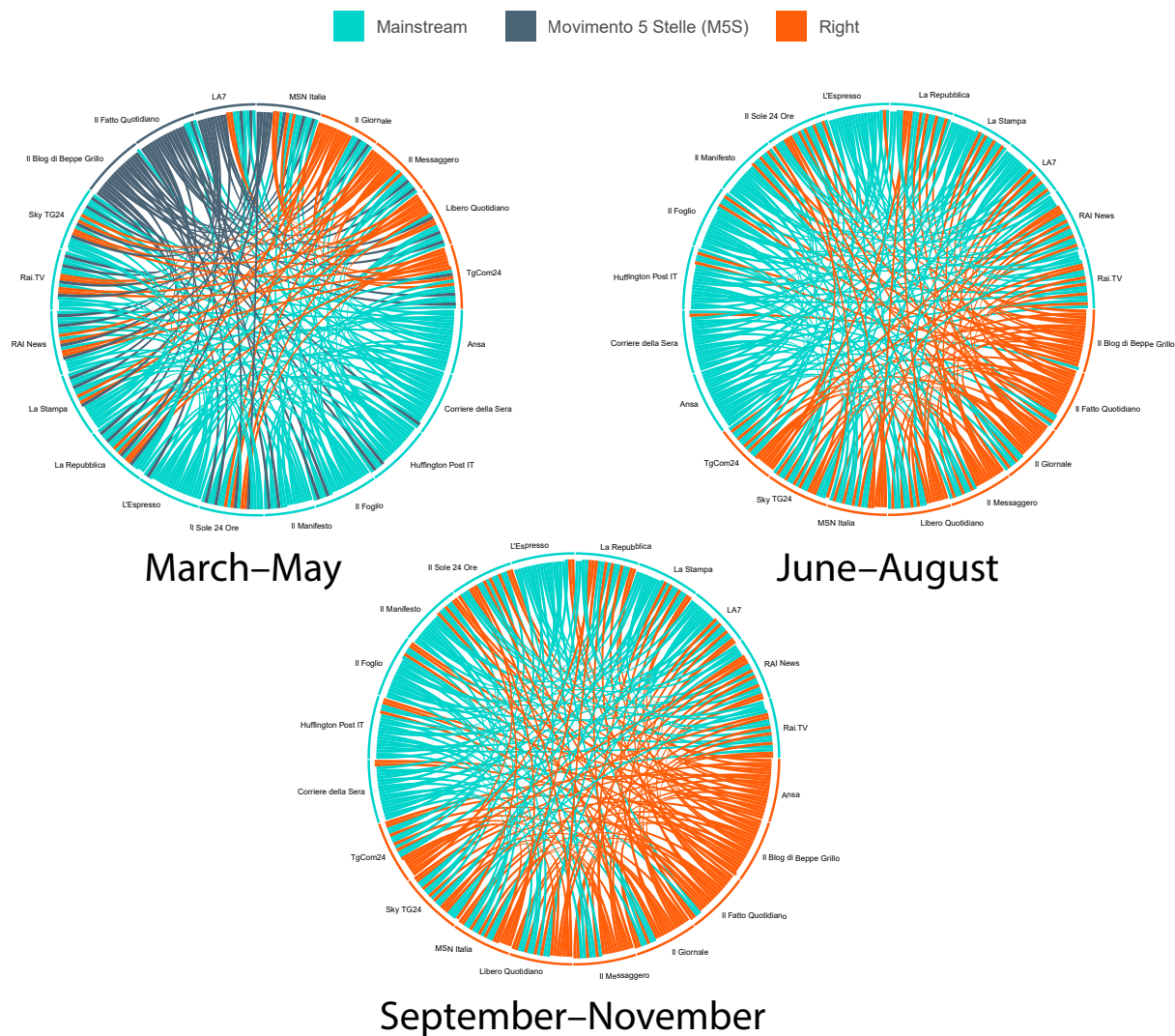


Figure 11. Communities that emerge from the retweeting behaviour of the users around migration tweets on national newspapers Twitter accounts. The color denotes the communities that appear while applying the FastGreedy community detection algorithm.

There is no straightforward causal effect between negative views of irregular migration in the general population and the media coverage of it.³⁷ However, it is noteworthy that according to the survey “Gli italiani e il resto del mondo”,³⁸ immigration was the second most feared issue for Italians in 2018 – just as it was in 2014, when the Mediterranean migrant crisis began in earnest. If anxiety levels remained high even when the numbers were already going down, they could have been stoked by Salvini’s rhetoric and the amount of coverage it received in the media.

³⁷ Jacobsen, “Communicating Irregular Migration”, p. 894.

³⁸ The survey has been conducted by Ipsos for Ispi (and Rai News 24) every year since 2014.

Key Takeaways:

- Though the numbers of migrants arriving in Italy dropped drastically during 2018, the amount of content about the subject increased. This coincided with the arrival of a new government in June 2018, and the determined efforts of Interior Minister Matteo Salvini to keep the issue high on the agenda through controversial policies and statements.
- Content that features Salvini far out-performs any other content on the subject of migration. His name is the only one to appear regularly in content about the subject.
- Content about migration achieves more engagement than other content, and the comments use more negative language than the average comments on *Corriere's* Facebook page.
- On Twitter we can see how debate about migration has moved from three disparate communities to two: “traditional media” followers versus “alternative media” supporters of the two “anti-establishment” parties, Lega and the Five Star Movement.

Experimental Analysis of Engagement and Reaction on Non-Targeted Content

Migration was a central theme of the Italian debate throughout 2017-2018, and remains so. This gives rise to our central question: are there ways of covering the issue that sharpen or weaken hyper-partisan polarisation, that enhance more constructive conversation, and that communicate facts better than others?

Our aim was to explore how readers engage with *Corriere's* content on migration. We classified content according to:

- Mediums (video, text, infographics, etc.)
- Themes (crime, culture and society, etc.)
- Journalistic techniques (human interest story, constructive news, fact-checking, etc.)

We then analysed engagement with *Corriere's* content through online metrics, comment annotation and sentiment of language in articles and comments. We focused on three engagement metrics (likes, comments and shares) characterised by different meanings (respectively appreciation, discussion, willingness to distribute the content) and different levels of effort (a like requires minimal effort, sharing and commenting more).

We extracted the sentiment of all migration posts and comments appearing on the *Corriere della Sera* Facebook page. To do so, we relied on Watson Natural Language Understanding. We evaluated each comment and post in a scale ranging from -1, a fully negative sentiment, to +1, a fully positive sentiment. When the score is 0 the sentiment is neutral – neither positive nor negative. The sentiment of a comment refers to the language used within the comment, not the intention of the user. Thus a comment on a negative topic like crime could be negative even when the framing of the comment is positive.

As for comment annotation, we hired three Italian mother-tongue graduate students to manually annotate a set of 26,524 users' comments³⁹ labelled according to:

- Toxic Speech, i.e. the presence of vulgar, aggressive, rude, disrespectful or unreasonable language which could lead a user to abandon a conversation. During the project we could not find any way to define "constructive" speech, so we took levels of toxic speech as an indication of whether content inspired more or less civil discourse.
- Criticism of Media Source (*Corriere della Sera*), i.e. the presence of comments directly criticising the publisher/journalist/newspaper as a source of information. We used this metric to consider whether and how traditional media - which are more and more accused by populist politicians of being "fake news", and part of "the establishment" - can enhance trust while providing fact-driven content.

³⁹ The total number of comments was 26,886, but a certain amount were spam or impossible to decipher.

- Pro- and Anti-Migration Stance, i.e. whether a comment was in favour, against, or neutral about the migration phenomenon in Italy. Comments that were explicitly in favour or against a politician were not considered a stance on migration, unless the comment was *explicit* with respect to the topic.

To perform the engagement metrics and reactions analysis and ensure comparability, we relied on the set of non-targeted posts.

1 - Engagement Metrics and Reactions by Medium

The mediums considered were: Articles, Article+Infographics, Multimedia, Purely Visual and Quizzes. The figure below shows the number of the articles that use each of the mediums. As we can see, the standard Articles are disproportionately represented and we only have one quiz.

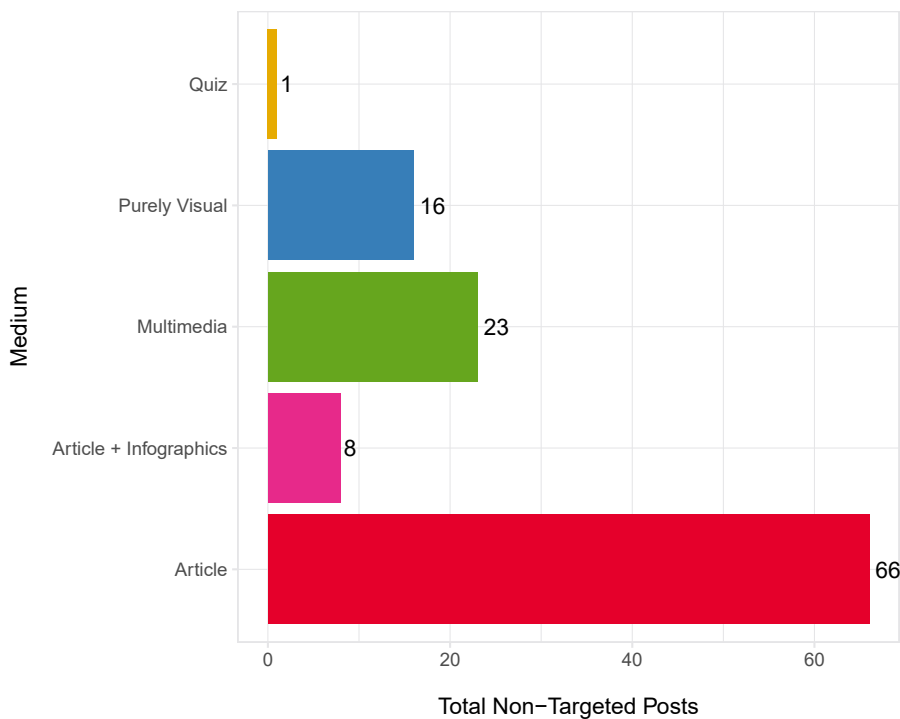


Figure 12. Number of non-targeted posts grouped by their medium.

1.1 - Engagement (Likes, Shares, Comments)

When it comes to engagement based on “likes” and “comments”, multimedia pieces scored highest (214 likes, 132 comments in median). Articles with infographics get the highest number of shares in median (77.5), which might suggest that they are used to prove points to other readers.

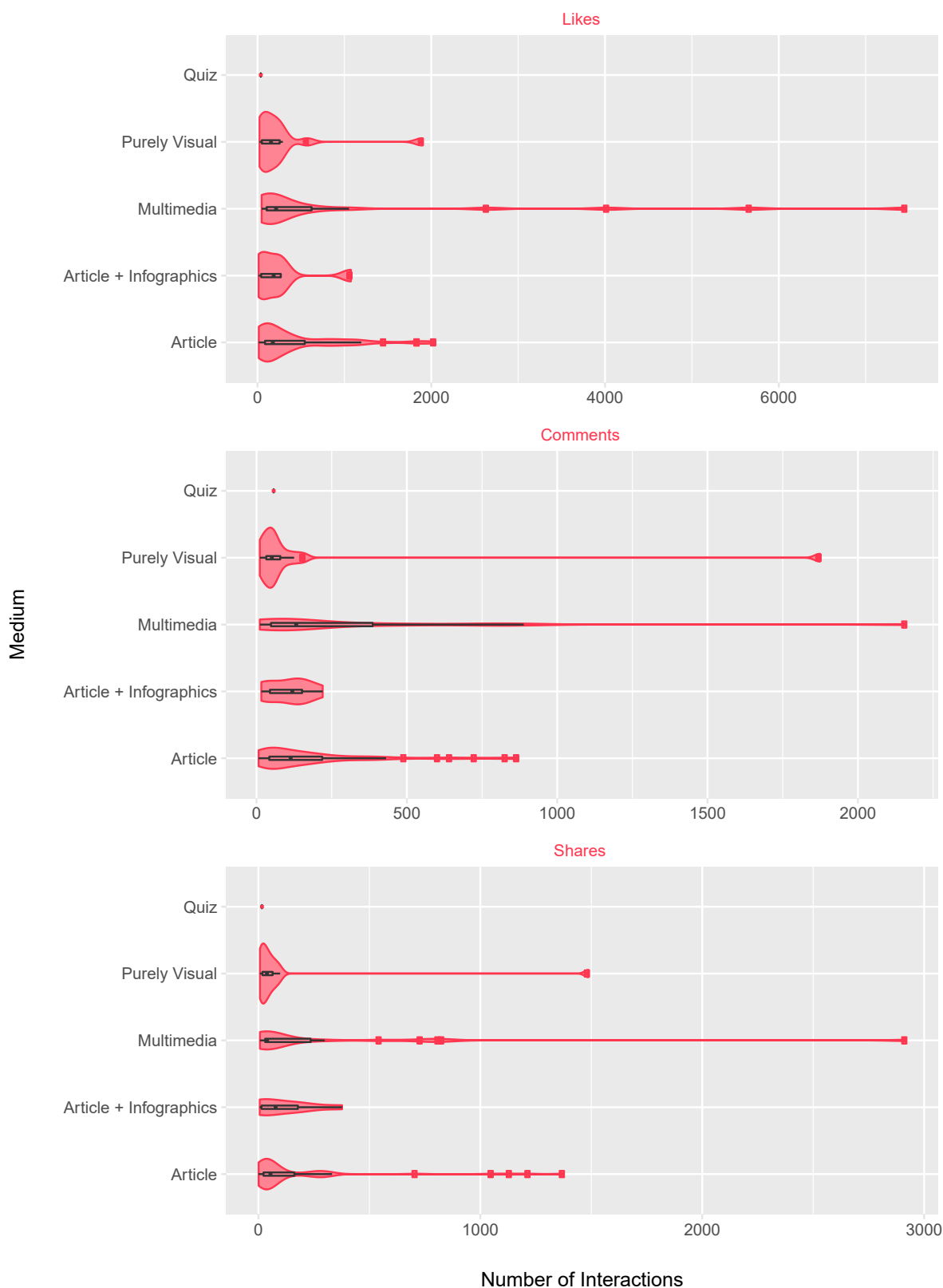


Figure 13. Distribution of the likes, comments and shares of the posts across the different mediums.

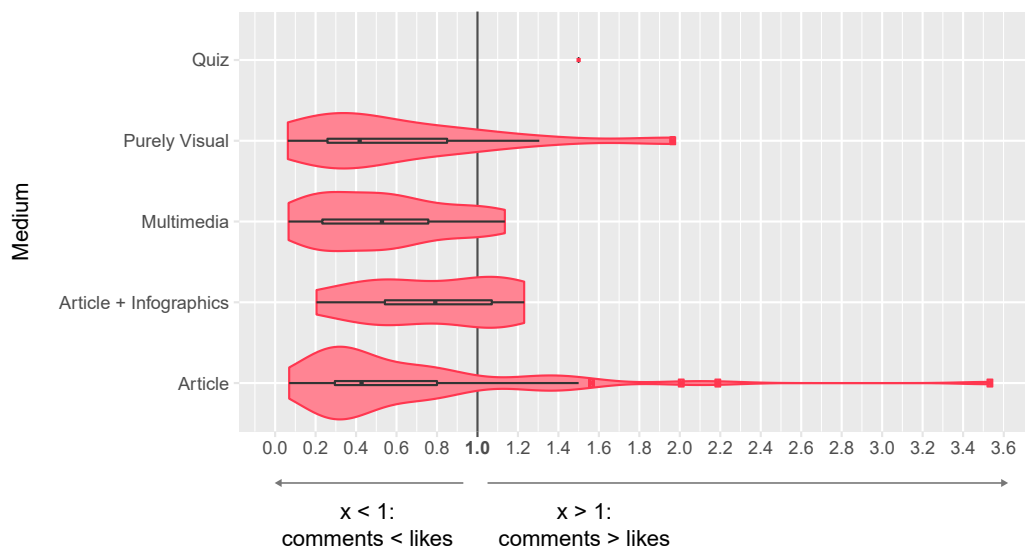


Figure 14. Distribution of the ratio comments over likes of each post across the different mediums. If the value is above 1 then there are more comments than likes. If it's 1 then likes and comments are equal. If the value is below 1 then there are more likes than comments.

To explore this more deeply, let us examine the nature of the comments.

1.2 - Comment Annotation

The meaning of these engagement metrics comes into clearer focus when we match them with comment annotation.

The highest number of comments labelled as critical of source was for (our only) quiz (22%), followed by infographics (ca. 10% in median), suggesting that a “numbers” approach elicits strong pushback from readers. The medium that elicits the least criticism of media source is multimedia content (ca. 1% in median), suggesting that “seeing is believing” and inspires more trust – or perhaps that when faced with video evidence, anti-migration groups feel discouraged from commenting.

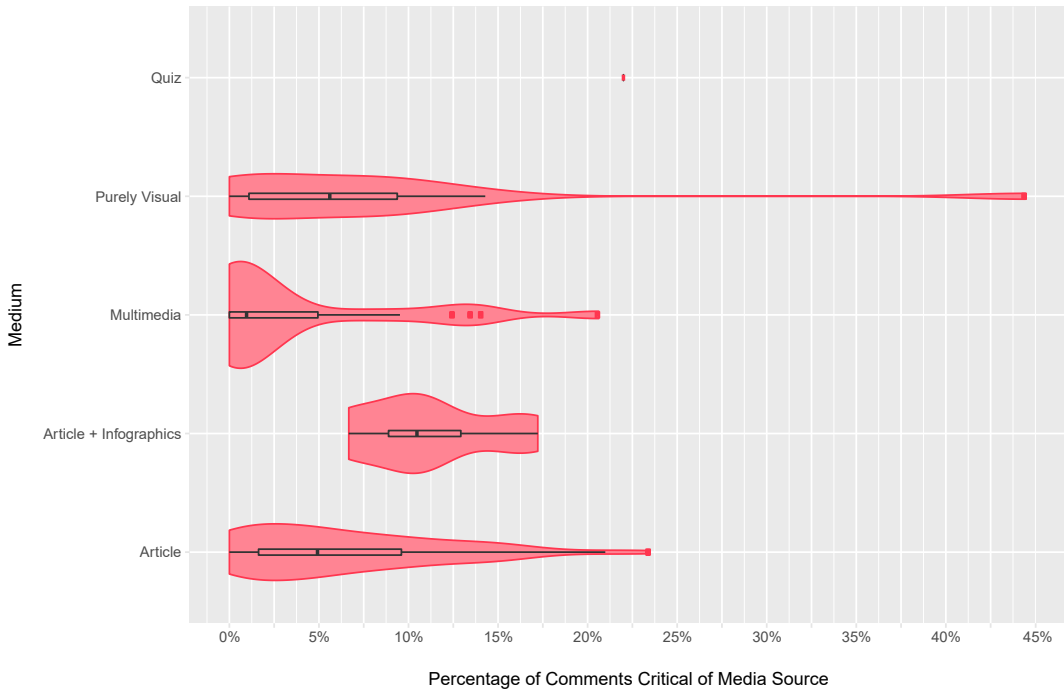


Figure 15. Distribution of the percentage of comments critical of the source of each post across the different mediums.

This pattern is reinforced when we look at anti- and pro-migration comments. Purely visual content gets the lowest number of anti-migration comments (-0.12 position score in median) and the second lowest percentage of toxic comments (7.9% in median), suggesting again that anti-migration audiences become shy when faced with footage from refugee boats or showing positive actions by migrants (such as the video of a migrant saving a young child).

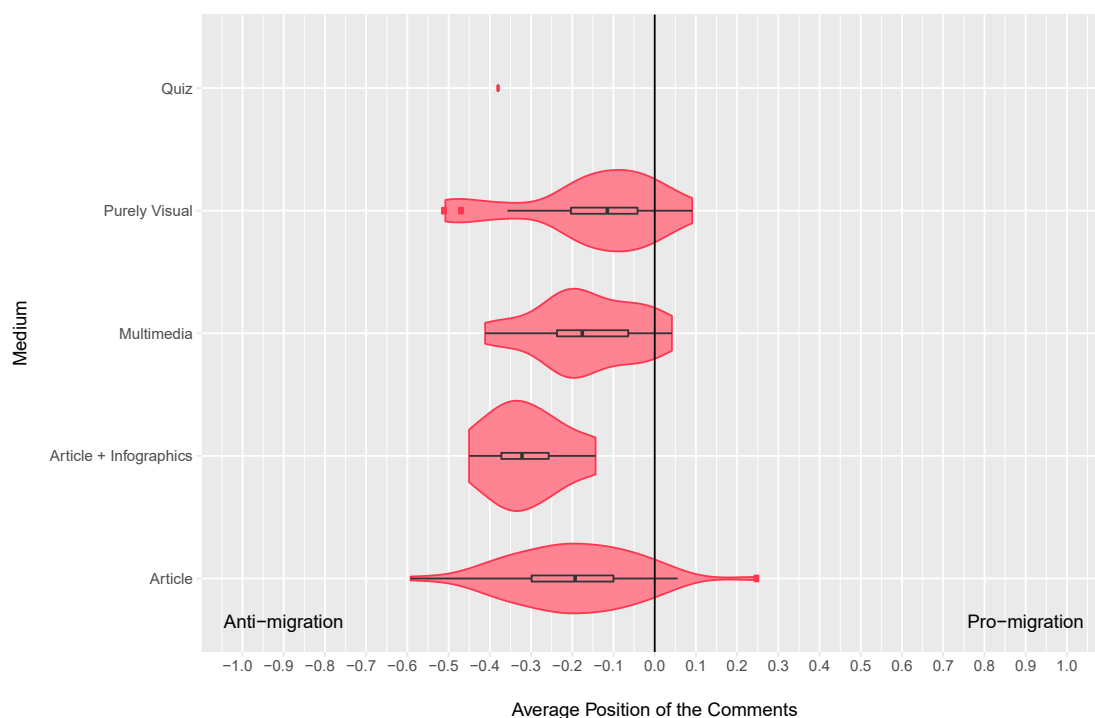


Figure 16. Distribution of the average migration position of the comments of each post across the different mediums, where -1 is fully anti-migration, 0 is neutral or undefined and 1 is fully pro-migration.

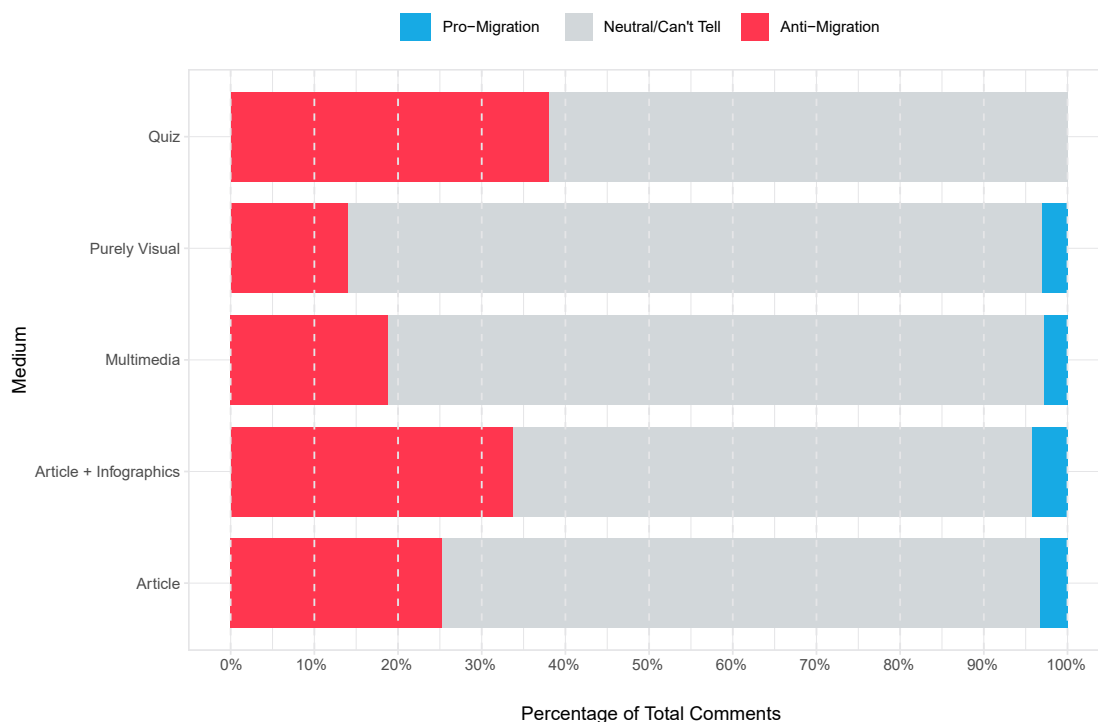


Figure 17. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by medium.

The highest proportion of anti-migration comments was elicited by the quiz (38%), followed by infographics (34%), suggesting that pushing dry data to the forefront of content can get pushback. However, it should be noted that infographics also elicited the highest proportion of pro-migration comments (4.3%), suggesting they are used to reinforce arguments for both sides of the debate.

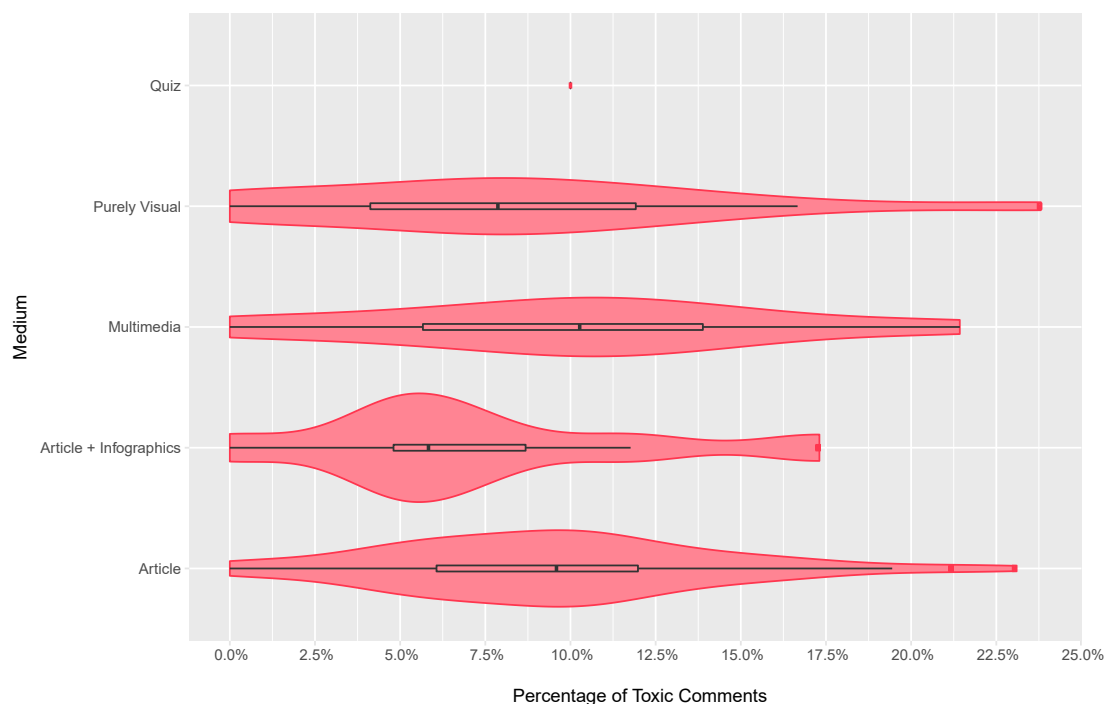


Figure 18. Distribution of the percentage of toxic comments of each post across the different mediums.

Infographics got the lowest percentage of toxic comments (5.8% in median), suggesting that while they do provoke pushback, the temperature of the debate is more tepid.

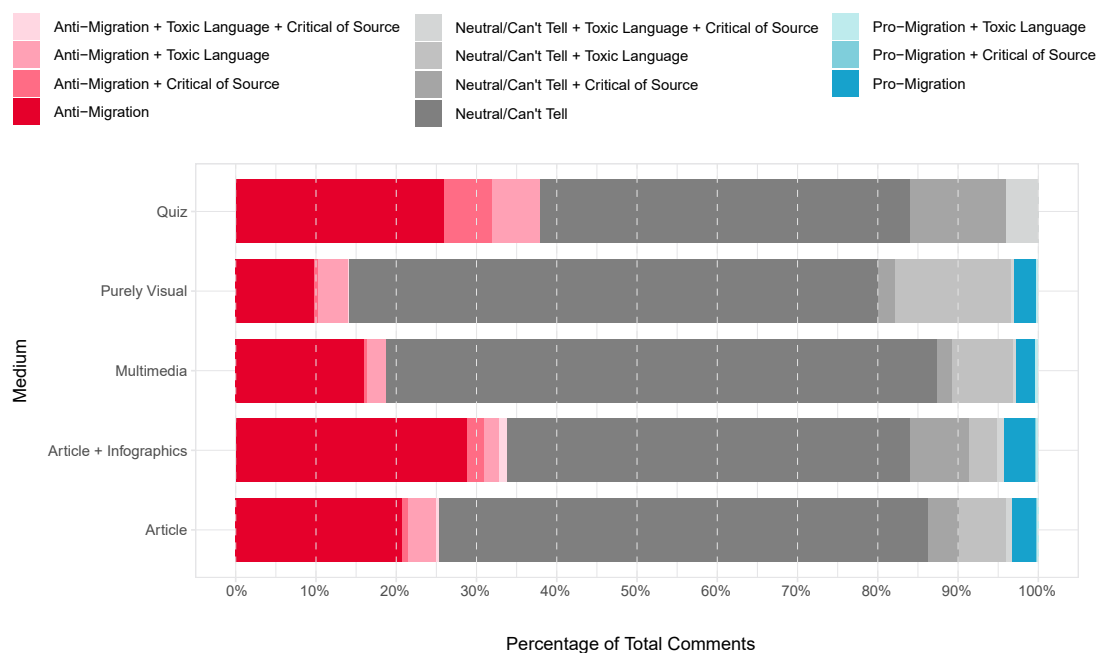


Figure 19. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by medium while also considering the toxic and critical of source labels.

Key Takeaways:

- Mediums that push facts to the forefront, such as infographics, can provoke debate and discussion but receive significant pushback from anti-migration voices and elicit high criticism of media. However it is also a fairly tepid debate, with low levels of toxic speech.
- Video produces the lowest level of criticism of media, potentially because “seeing is believing” (this assumption would need to be explored further in qualitative studies) or because anti-migration voices become quiet when faced with video evidence.
- Multimedia pieces receive strong and supportive engagement, with high numbers of likes and shares, while the relatively low number of comments suggests the pieces are less polarising. Multimedia pieces also perform relatively well in terms of criticism of media, suggesting relatively high trust for such stories, though not as high as for purely visual stories.
- Creating friction between the reader and the content could provoke negative reactions (though we would need many more quizzes to test this hypothesis).

2 - Engagement Metrics and Comment Annotation by Theme

Of the 114 non-targeted migration-related articles we analysed during 2018, 12.3% were about crime, 41.2% about the refugee crisis, 14% about ethnic tension and 32.5% about culture and society.

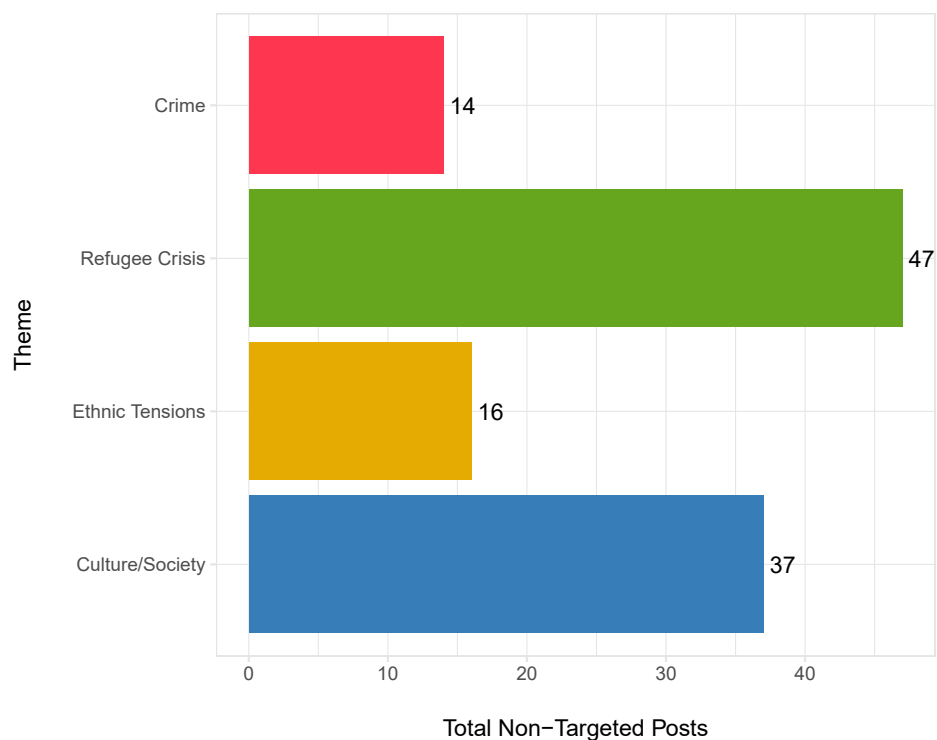


Figure 20. Number of non-targeted posts grouped by their theme.

2.1 - Engagement (Likes, Shares, Comments)

Culture and Society content seemed to get the lowest overall engagement (139 likes, 57 comments, 35 shares in median), while Ethnic Tensions received the highest amount of likes (307.5 in median) and comments (192.5 in median). Crime got the highest number of “comments-over-likes”, suggesting that when people engage with this content they want to make their opinion widely known (7 comments for every 10 likes in median).

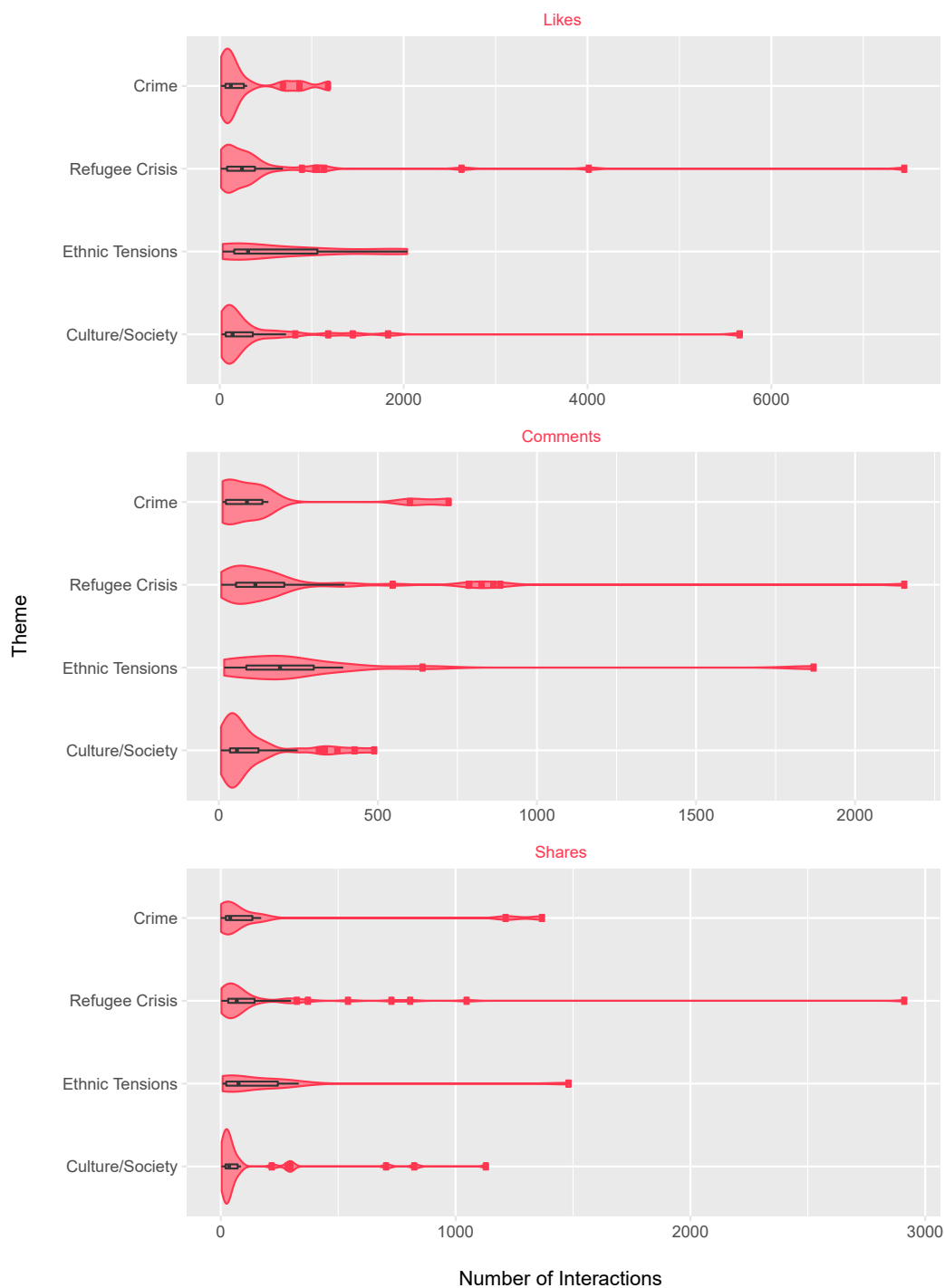


Figure 21. Distribution of the likes, comments and shares of the posts across the different themes.

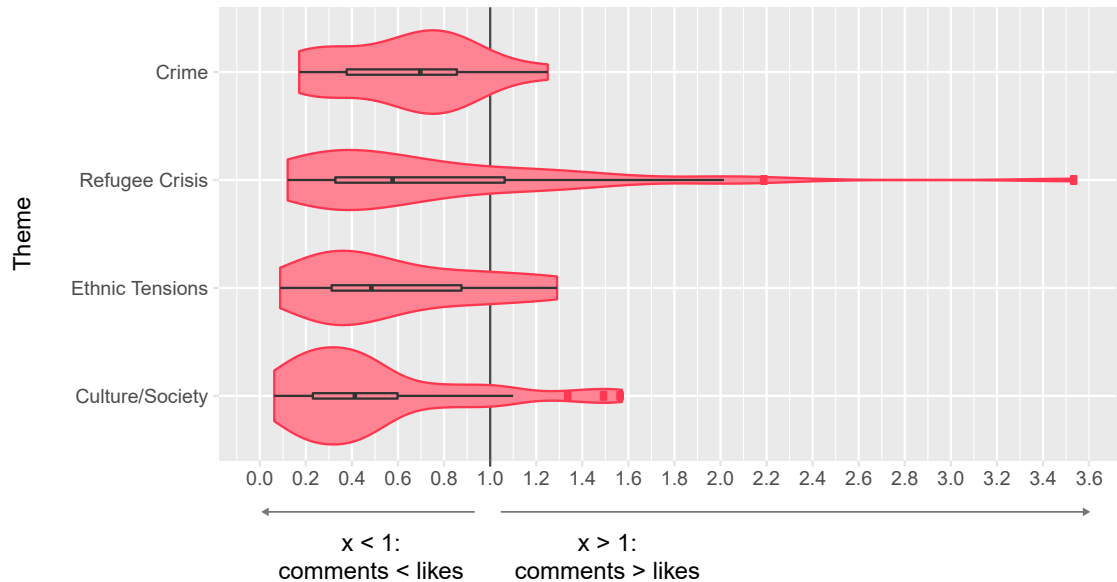


Figure 22. Distribution of the ratio comments over likes of each post across the different themes. If the value is above 1 then there are more comments than likes. If it's 1 then likes and comments are equal. If the value is below 1 then there are more likes than comments.

2.2 - Comment Annotation and Sentiment Analysis

The posts with themes related to Culture and Society proved to elicit the most comments critical of the source (ca. 11%), followed by the posts about Crime (ca. 4%). The Ethnic Tensions theme and the Refugee Crisis theme generated the lowest level of critical comments from the readers (respectively, 2.8% and 2.7%).

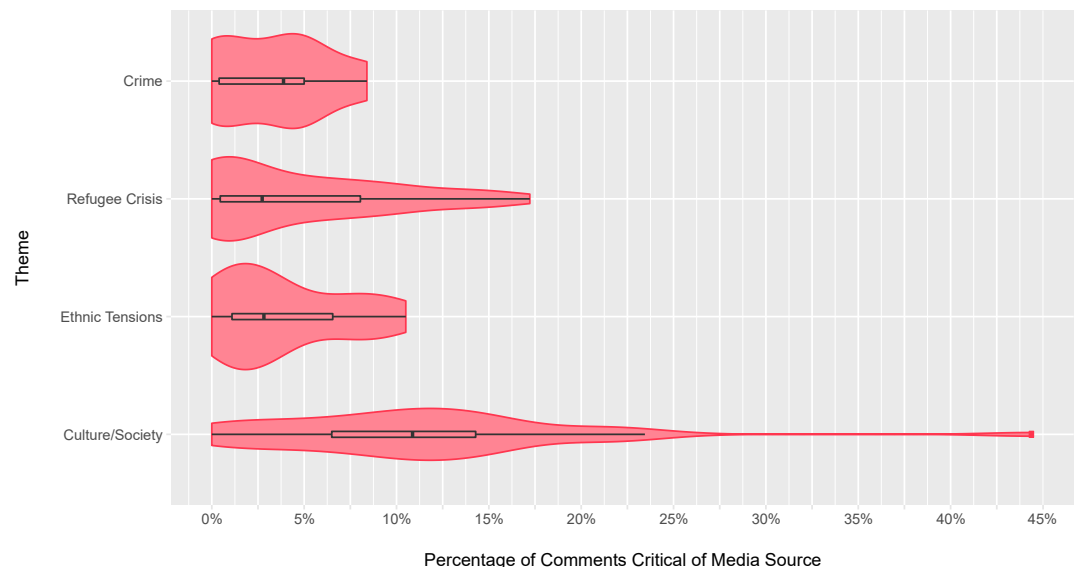


Figure 23. Distribution of the percentage of comments critical of the source of each post across the different themes.

Content about the Refugee Crisis provoked the most anti-migration comments (-0.24 score in median): keeping this theme high on the agenda seemed to bring out the anti-migration base.

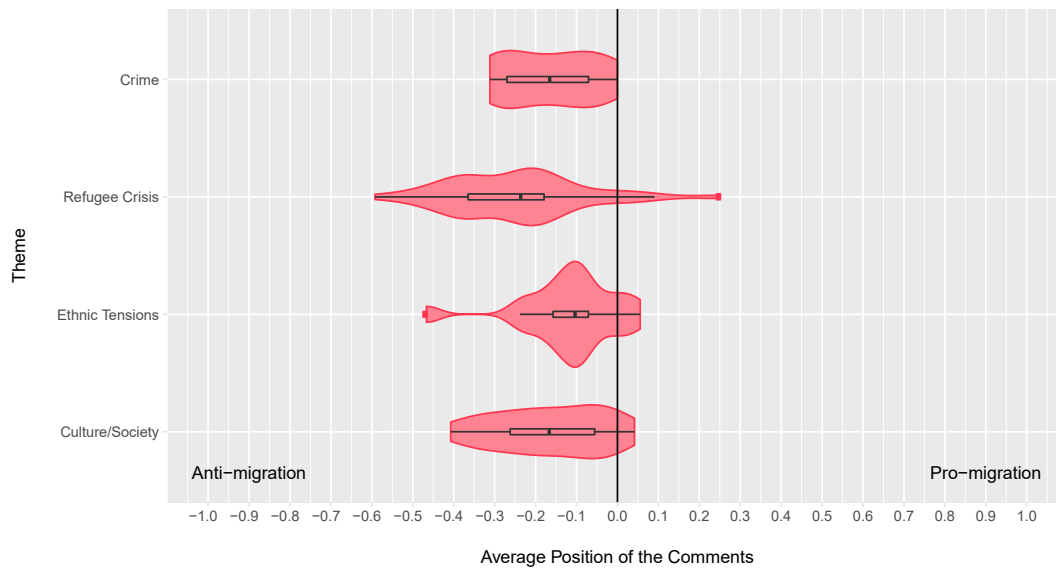


Figure 24. Distribution of the average migration position of the comments of each post across the different themes, where -1 is fully anti-migration, 0 is neutral or undefined and 1 is fully pro-migration.

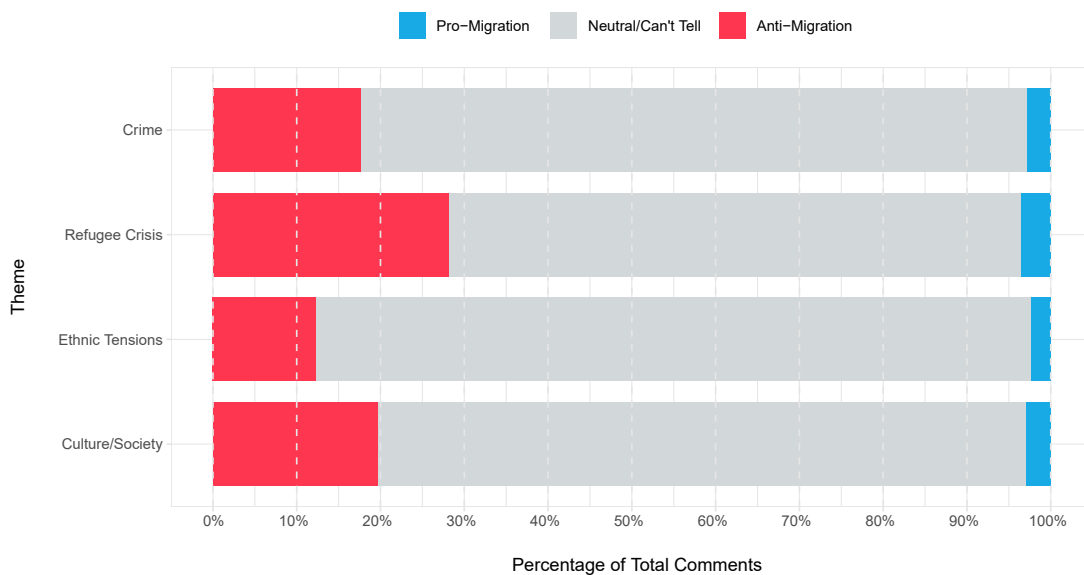


Figure 25. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by Theme.

Ethnic Tensions content produced the highest number of toxic comments (10% in median), suggesting that the most hotly contested issues dealt with the impact of migration on domestic social and ethnic tensions.

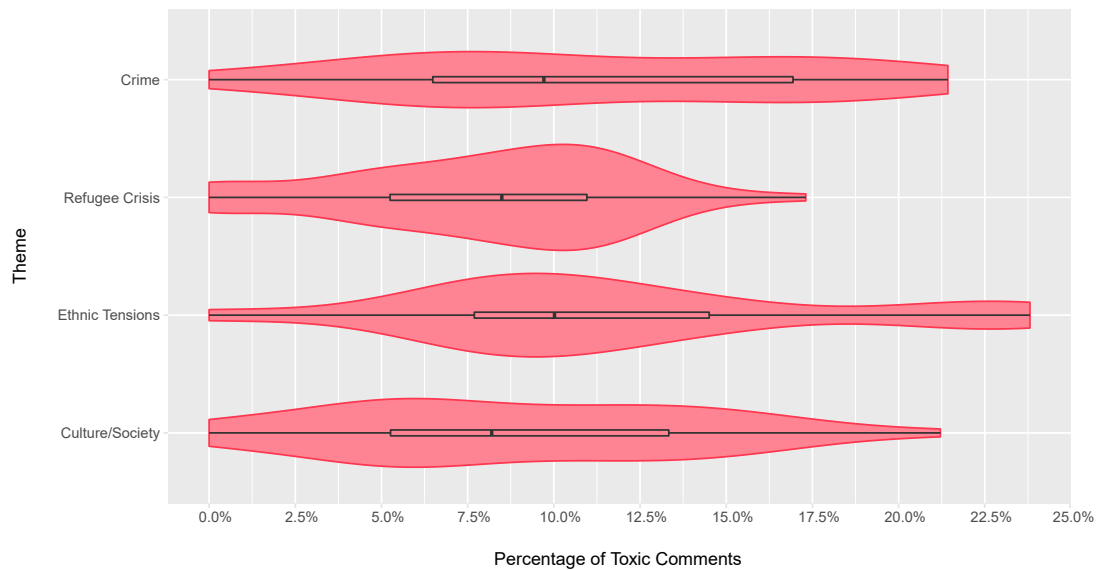


Figure 26. Distribution of the percentage of toxic comments of each post across the different themes.

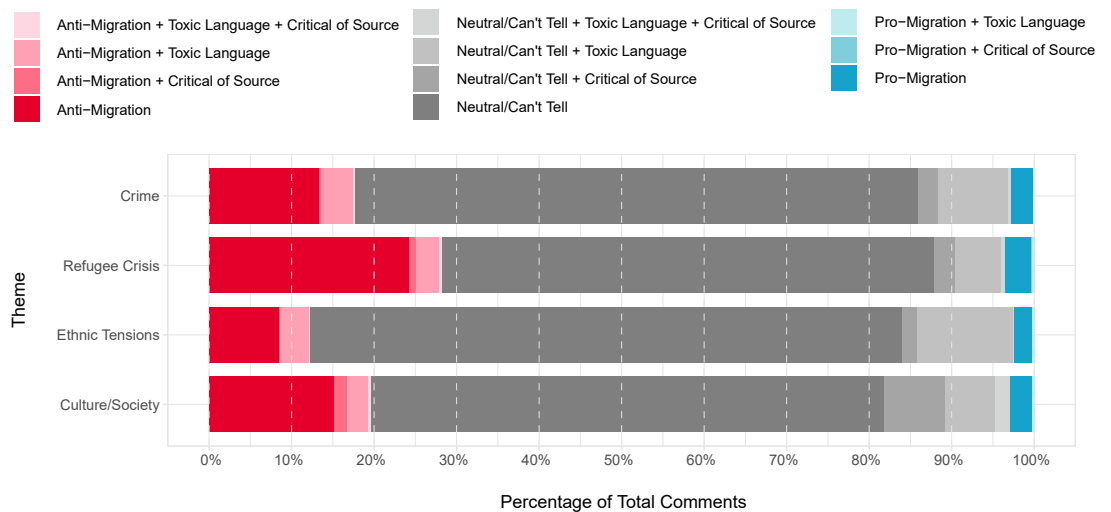


Figure 27. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by theme while also considering the toxic and critical of source labels.

Culture and Society content attracted the lowest percentage of toxic comments (toxicity: 8.2% in median) compared with other themes, and the highest percentage of positive sentiments (-0.20 sentiment score), but this did not lead to more constructive engagement. Quite the opposite: Culture and Society stories stand out as the ones that elicited the most criticism of media source (ca. 11%) and a high number of anti-migration comments (20% in proportion; only Ethnic Tension got more: 28% in proportion).

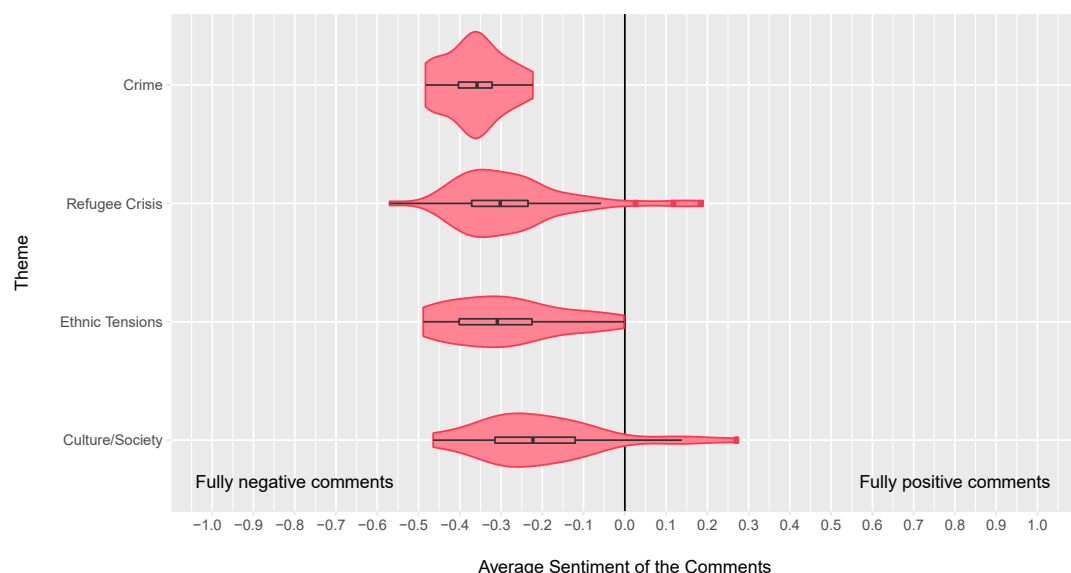


Figure 28. Distribution of the average sentiment of the comments of each post across the different themes, where -1 is negative, 0 is neutral and 1 is positive.

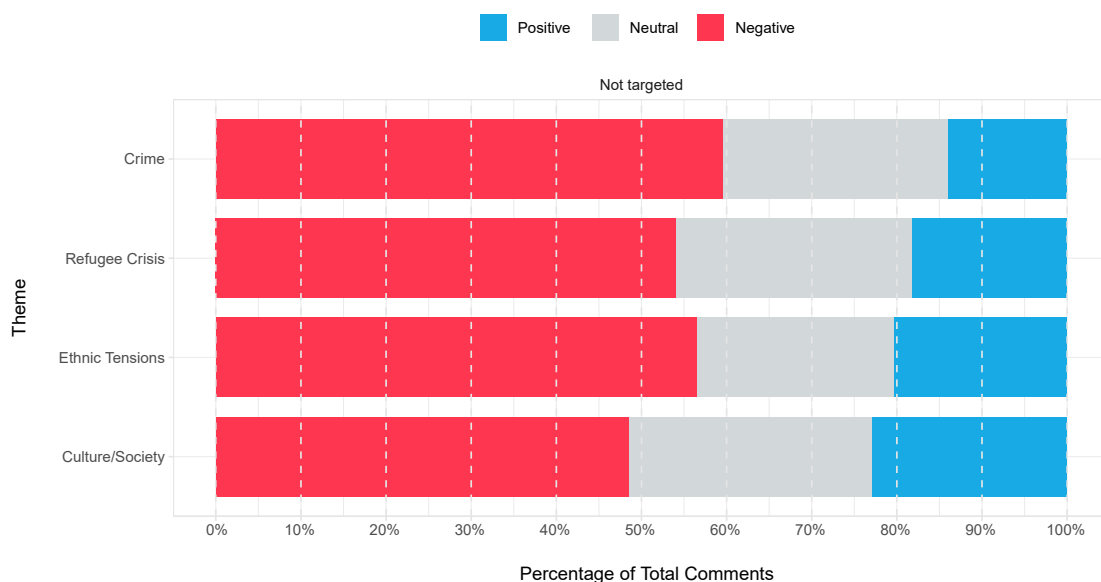


Figure 29. Proportion of all comments that are negative, neutral or positive in their vocabulary broken down by theme.

Key Takeaways by Theme

- **Ethnic Tensions** content about the everyday impact of migration on society produces the highest number of comments and also the most toxic comments.
- Stories about the refugee crisis get the most anti-migration comments, perhaps indicating why certain political forces want to keep them high on the agenda to “bring out their base”.

3 - Engagement Metrics and Engagement by Technique

As explained above, we broke down our content into eight techniques:

- **News report:** content reporting in a straightforward and objective way what happened or what somebody said.
- **Article with Context:** articles that go into or contextualise a specific situation.
- **Editorial:** articles in which experts or journalists give their opinion or comment on a particular event/situation.
- **Constructive News:** content providing a possible solution to migration issues.
- **Data-driven pieces:** fact-checking stories, as well as more analytical pieces.
- **Human Interest – Individual:** content that focuses on the adventures of individuals, often refugees.
- **Human Interest – Groups:** same as above but focused on groups.
- **Stories with Pop-culture References:** articles related to popular characters such as football players, singers, actors.

The breakdown of the non-targeted Facebook posts across the different techniques can be seen in the following figure, which shows that the standard news reports constitute a large percentage of posts.

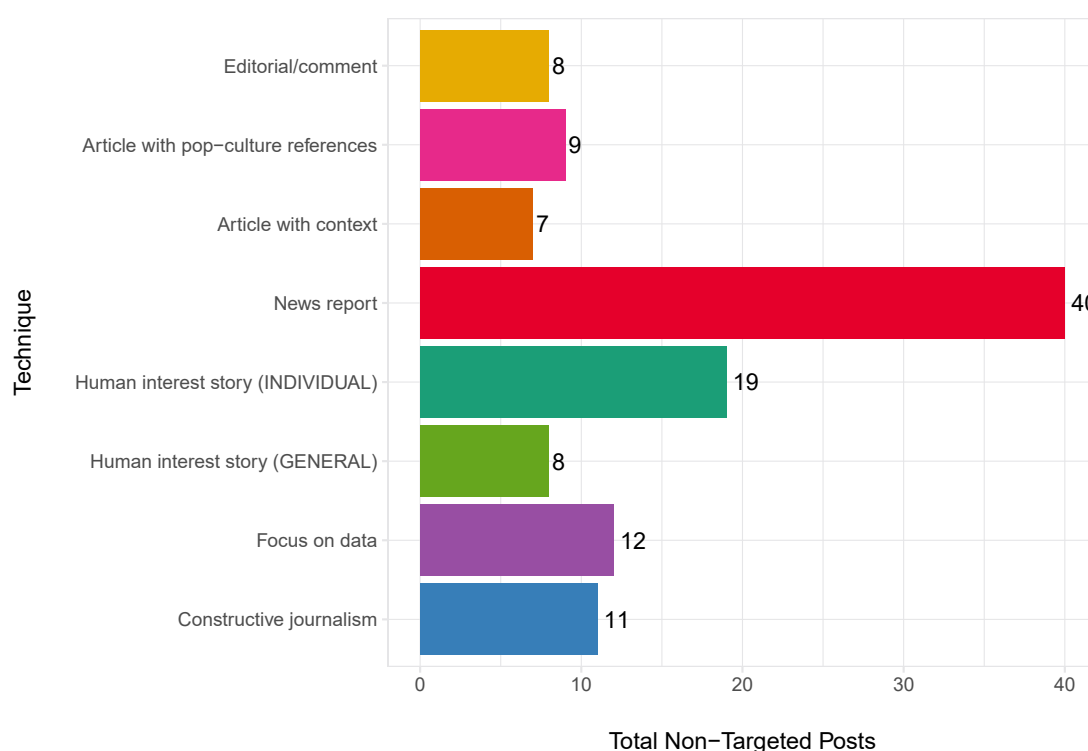


Figure 30. Number of non-targeted posts grouped by their technique.

3.1 - Engagement (Likes, Shares, Comments)

Editorial pieces got the highest level of engagement across likes (389 in median) and ranked second for both comments and shares (with 55.5 comments and 59.5 shares in median, respectively). Articles with pop-culture references got a higher number of com-

ments (163) and news reports a slightly higher share (74.5 in median), but this could be due to their topicality. Human Interest (General) pieces got the lowest response across all metrics (respectively, 78.5 likes, 32.5 comments, 21 shares in median).

Although content with pop-culture references got high likes and comments (respectively, 312 and 163 in median), there were low numbers of shares (35 in median). This suggests a softer level of approval (people superficially agreed with the story, and might chat about it, but they did not feel the need to push it further).

Though Human Interest (General) and Focus on Data got low numbers of likes (respectively, 78.5 and 124 in median), they got a high proportion of comments-over-likes (respectively, 0.75 and 0.80 in median). Possible reasons for this become clearer in the comment annotation.



Figure 31. Distribution of the likes, comments and shares of the posts across the different techniques.

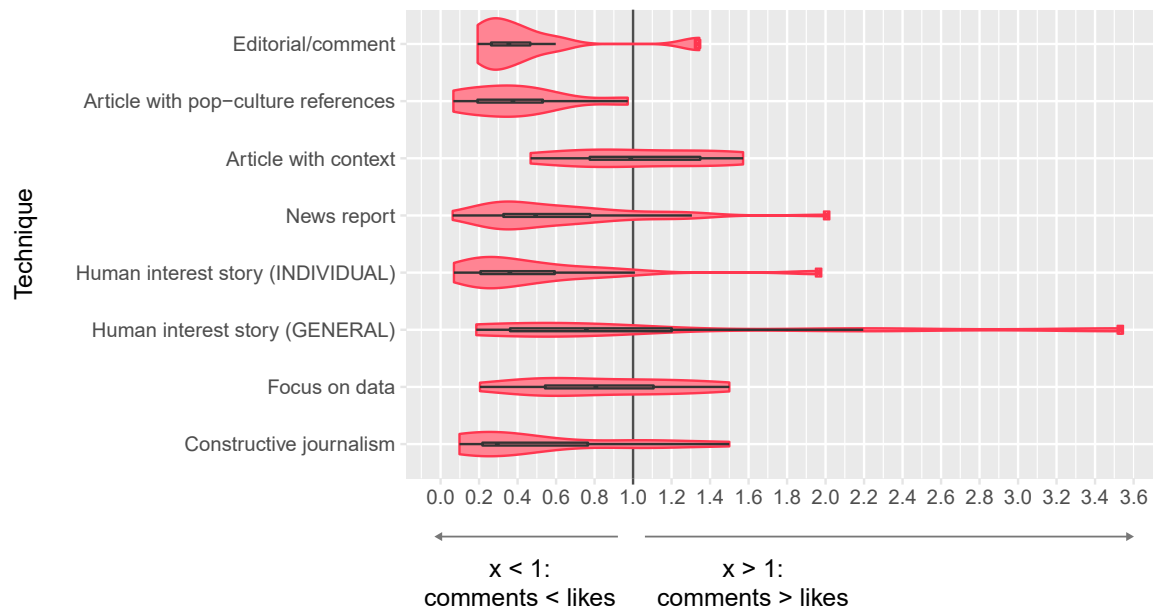


Figure 32. Distribution of the ratio comments over likes of each post across the different techniques. If the value is above 1 then there are more comments than likes. If it's 1 then likes and comments are equal. If the value is below 1 then there are more likes than comments.

3.2 - Comment Annotation

Articles with Pop-culture References and Constructive Journalism got the lowest level of negative sentiment in comments (respectively, -0.15 and -0.24 in median). Looking across the comment annotation, Constructive News also got the lowest number of toxic comments (6% in median), suggesting it inspires a more civil discourse.

It is perhaps more instructive to compare Constructive Journalism with Editorial pieces, as both are about putting forward innovative ideas on the subject of migration. While both elicited high levels of criticism of media source (respectively, 9% and 8.5%), there was a (slightly) lower level of negative language and toxic speech in the comments for Constructive Journalism (-0.24 sentiment score and 6% of toxic language for Constructive Journalism, as opposed to -0.27 and 11.3% for Editorial/Comments), again suggesting the former produces a more civil debate about polarising issues. Editorials, however, got more overall engagement (389 likes, and 155.5 comments, in median, for Editorials, as opposed to 152 and 56, respectively, for Constructive Journalism).

Straight, classical news pieces (News reports) received the lowest level of criticism of media source (1.5% of critical comments in median), hinting that being "impartial" and "objective" could be the best way to win trust. Articles with Context produced the second lowest criticism of media source (5% in median). Both were relatively well shared, liked and commented on. (News reports: 260 likes, 133.5 comments, 74.5 shares in median; Articles with Context: 107 likes, 65 comments, 32 shares in median)

By far the strongest criticism of media source came in response to Human Interest (General) pieces (12.3%). They also got sizeable anti-migration (-0.31 sentiment score) and toxic comments (9.7%). One of the stories attracting the most criticism of media source, for example, was a seemingly very warm story about a soccer team made up purely of migrants (21% toxicity; this article ranked sixth highest in the overall ranking by toxicity). It should be noted that human interest stories about individuals attracted less criticism of media source (6% against 12.3% in median), suggesting that when faced by the plight of individuals, readers are less likely to attack the journalists. However, human interest stories about individuals still provoked noticeably more criticism of media

source than straight news stories and articles with context (respectively, 6% against 1.5% and 5%). It is worth asking whether some readers felt they were being somehow emotionally manipulated by such stories and pushed back strongly.

Focus on Data got the highest number of anti-migration comments in median (-0.32), high levels of criticism of media source (9.3% in median, ranking second) but relatively low toxicity (7%, ranking second-to-last). This suggests that a data-driven approach does little to win over anti-migration audiences, though it generates relatively mild debate.

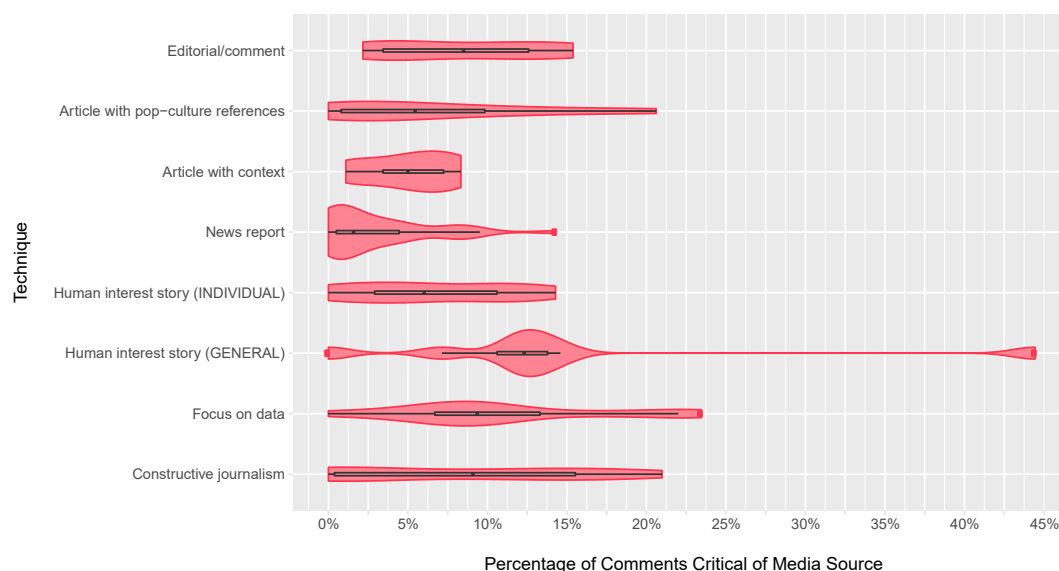


Figure 33. Distribution of the percentage of comments critical of the source of each post across the different techniques.

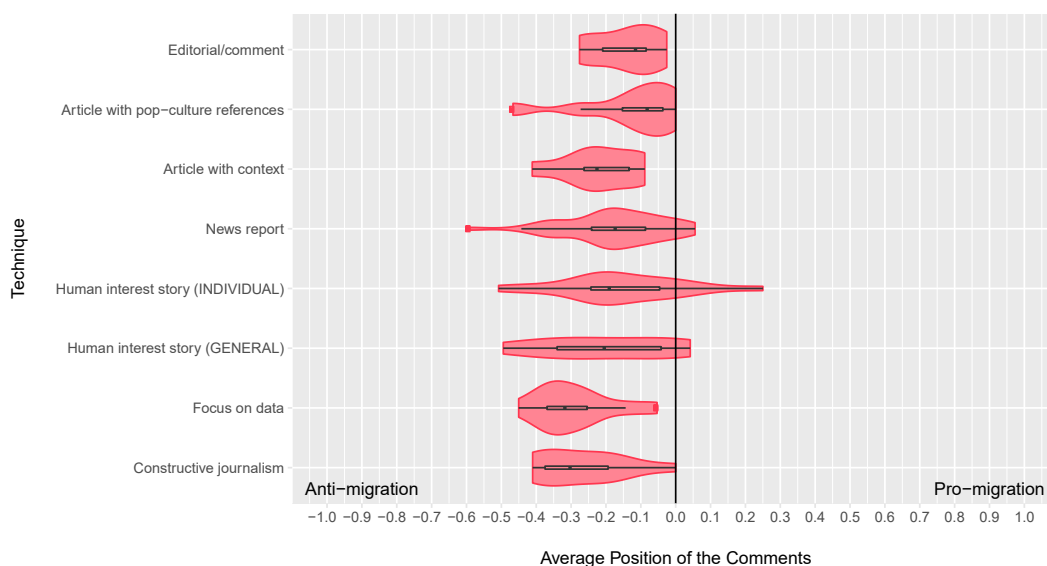


Figure 34. Distribution of the average migration position of the comments of each post across the different techniques, where -1 is fully anti-migration, 0 is neutral or undefined and 1 is fully pro-migration.

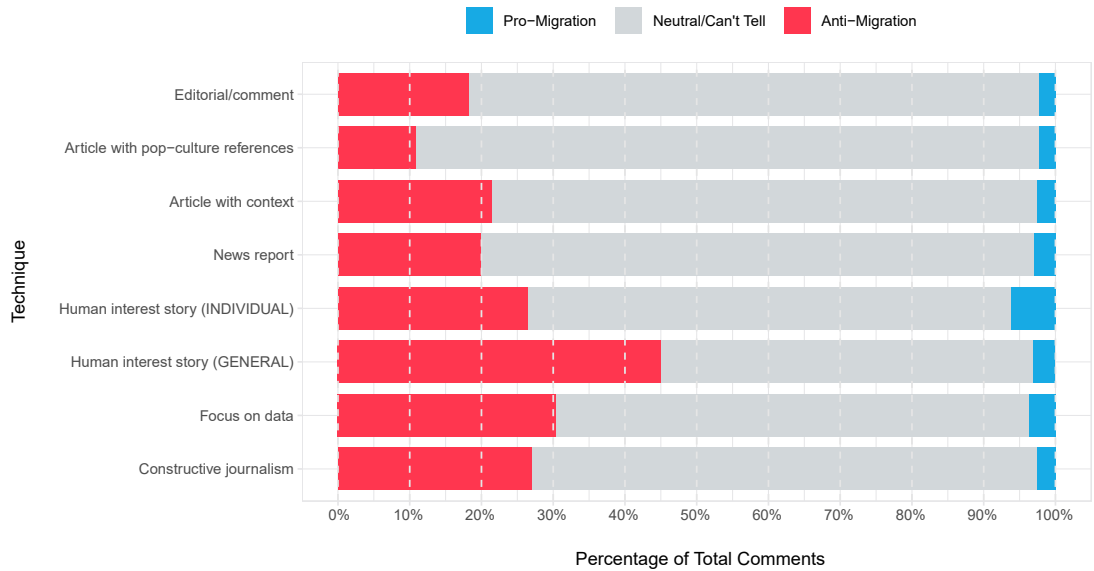


Figure 35. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by technique.

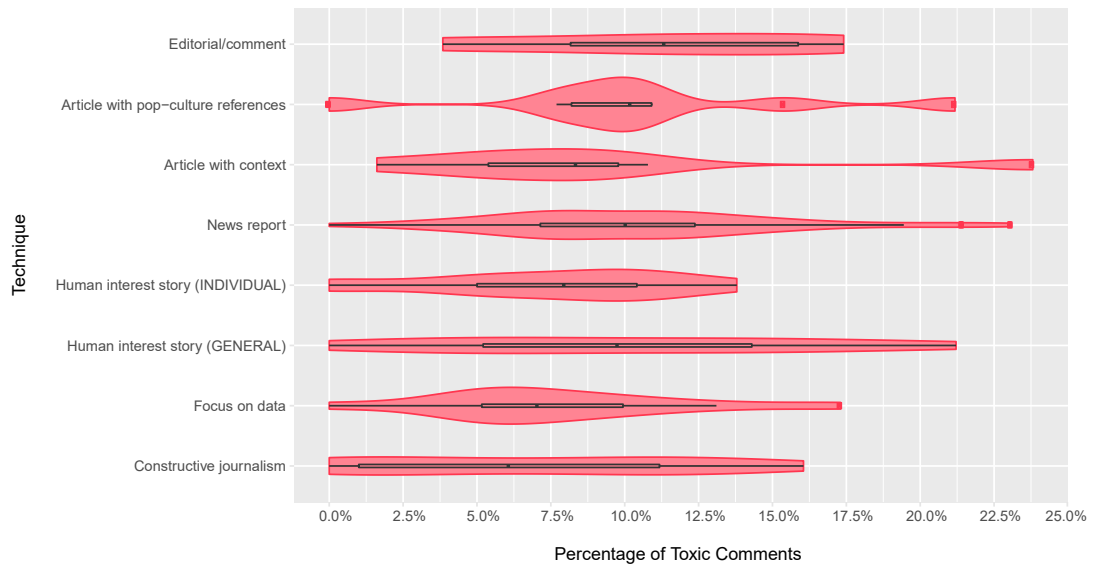


Figure 36. Distribution of the percentage of toxic comments of each post across the different techniques.

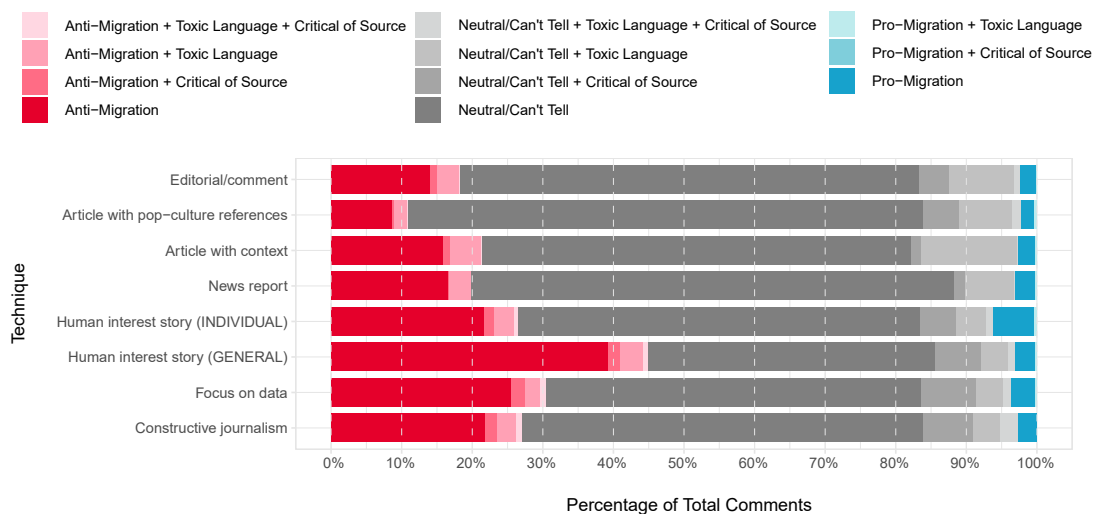


Figure 37. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by technique while also considering the toxic and critical of source labels.

Key Takeaways by Technique:

- The articles that elicit the most trust are straightforward, impartial and unemotional news reports and articles with context.
- Human interest stories, particularly about groups of people, get strong push-back, eliciting a high number of anti-migration comments, many toxic comments and a lot of media source. Perhaps some audiences push back when they feel they are being emotionally manipulated into taking a certain side on this topic.
- Constructive news inspires a less toxic debate about migration than editorial pieces. However, any type of strong opinion or policy proposal on the subject will elicit a lot of criticism of media source and bring out the anti-migration commenters.
- A data-driven approach elicits high levels of criticism of the media source and of anti-migration comments.
- Editorial pieces get the most engagement overall, while pieces with pop-culture references get the most likes.

| Special Case Studies

We undertook a number of special case studies to explore and test various hypotheses. Though the sample size of these tests was sometimes small, they pointed to avenues for future research.

1 - Comparing the General Audience with a More Polarized Audience

As explained in the previous sections, a small percentage of the content in the Experimental Dataset was simultaneously targeted by *Corriere* at partisan audiences who identified themselves as either very pro- or anti-migration, on their Facebook page. The pro-migration groups, for example, were followers of humanitarian NGOs, while the anti-migration ones followed far-right parties. This analysis was done on 42 posts, 21 targeted and 21 non-targeted.

Would directing content at strongly pro and anti partisan groups make the engagement more polarised – or were all audiences already so polarised that it would make no difference? Results seem to indicate that the latter could well be the case, as most content showed similar metrics to non-targeted posts.⁴⁰

Among the techniques tested in this case study, the target audience did not significantly affect the percentage of toxic comments.

Comments on Constructive Journalism remained low in terms of toxicity even in this more polarised environment (11% as opposed to 9%), indicating that this medium smooths emotional reactions; it also got relatively low criticism of media source in targeted posts.

Focus on Data pieces got 10% of toxic comments regardless of the audience type. Human Interest stories focused around individuals also got 10% of toxic comments when non-targeted, and 8% when targeted. Human Interest stories focused around groups got 15% of toxic comments when non-targeted, and 16% when targeted. The standard News report attracted the most toxic comments from all audience types, with 23% of toxic comments when non-targeted and 20% in the targeted posts.

⁴⁰ Levels of engagement on targeted content were usually lower, because targeted audiences are always smaller than the non-targeted and perhaps because these audiences had already been exposed to these stories.



Figure 38. Proportion of total comments that are toxic, broken down by technique and target.

When looking at the percentage of comments critical of the source, the scenario is different.

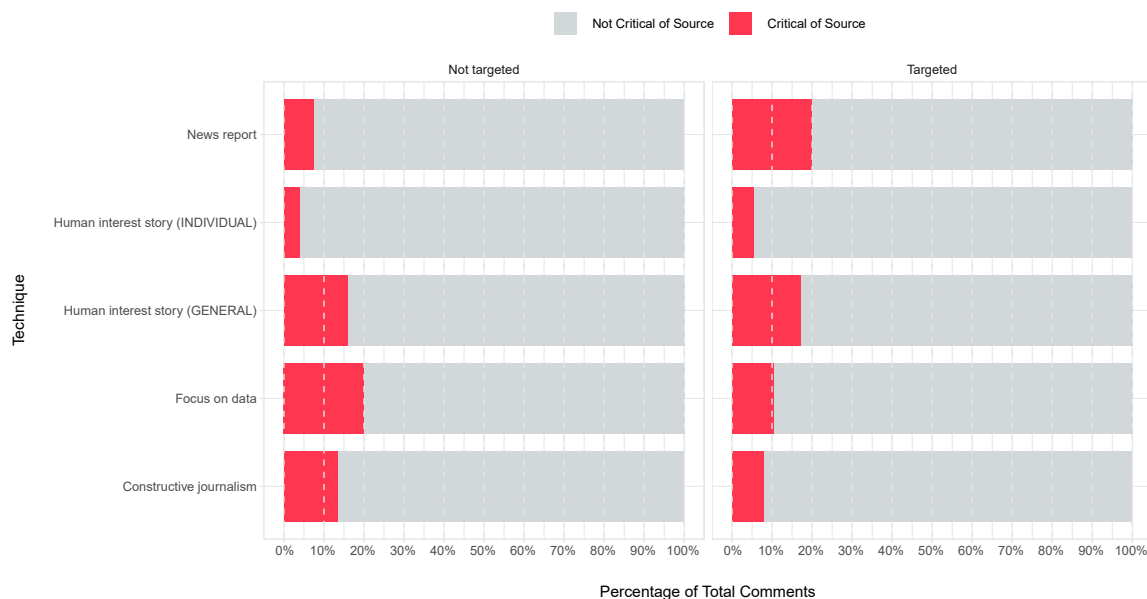


Figure 39. Proportion of total comments that are critical of the source, broken down by technique and target.

Regular news reports received more criticism of source among targeted audiences, as one would perhaps expect in an a priori hyper-polarised environment. They also received more anti-migration comments. The opposite is true, however, when considering the Focus on Data and Constructive Journalism pieces. The Focus on Data piece got double the level of comments critical of source (20%) when the content was non-targeted than when it was targeted (10%). Equally, with Constructive Journalism non-targeted pieces got 13.6% of critical comments versus 8% for non-targeted content. It is interesting to note that both Human Interest story techniques (General and Individual) did not show significant differences between audiences in terms of the percentage of critical comments.

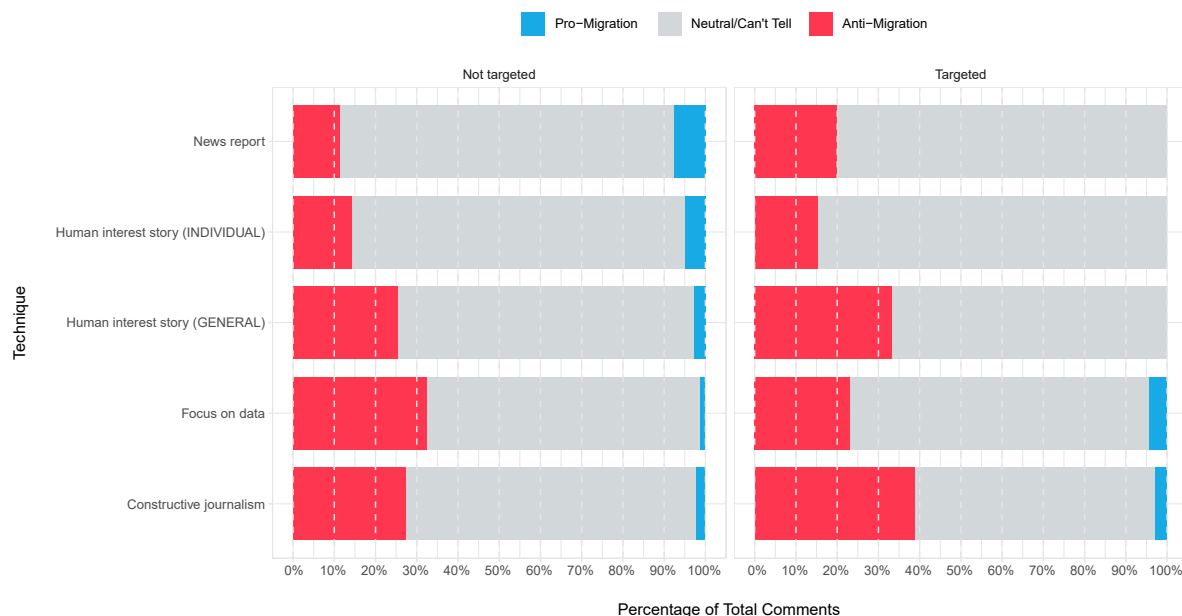


Figure 40. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by technique and target.

When considering the position on migration taken in the comments we can see some differences in the techniques, depending on the targeted audience. With the Constructive Journalism technique, 27.5% of the comments were anti-migration and 2% pro-migration on the non-targeted content, versus 38.9% and 2.8% respectively on the targeted content. The Focus on Data technique, on the other hand, got 32% anti-migration and 1% pro-migration comments on the non-targeted content as opposed to 23% and 4%, respectively, on the targeted content.

Human Interest stories presented an interesting case: there were no pro-migration comments at all on the targeted content (while the non-targeted content showed 2.8% in Group stories and 4% in Individual stories). The Human Interest content based on groups of migrants also showed significant differences between the targeted (33%) and non-targeted content (25%). The Human Interest stories about individuals, by contrast, presented very similar results for anti-migration comments (14% of total comments when non-targeted and 15% when targeted).

When we look at the various mediums employed to cover the stories, we can note other differences in the way the targeted and non-targeted audiences reacted.

Purely Visual mediums and Multimedia received noticeably more anti-migration comments in the targeted posts (47% and 25% respectively) than in their non-targeted posts (20% and 13% respectively). In the targeted content, Purely Visual and Multimedia content only displayed pro-migration comments (4.7% and 5.7% of total comments respectively).

For content making use of infographics there were no significant differences between the targeted (43% anti-migration and 2.8% pro-migration) and non-targeted content (40% anti-migration and 2.5% pro-migration). There was a similar result for the standard article, with 26% of the comments being anti-migration and 1.9% pro-migration in the non-targeted content, and 27% anti-migration and 3% pro-migration in the targeted content.

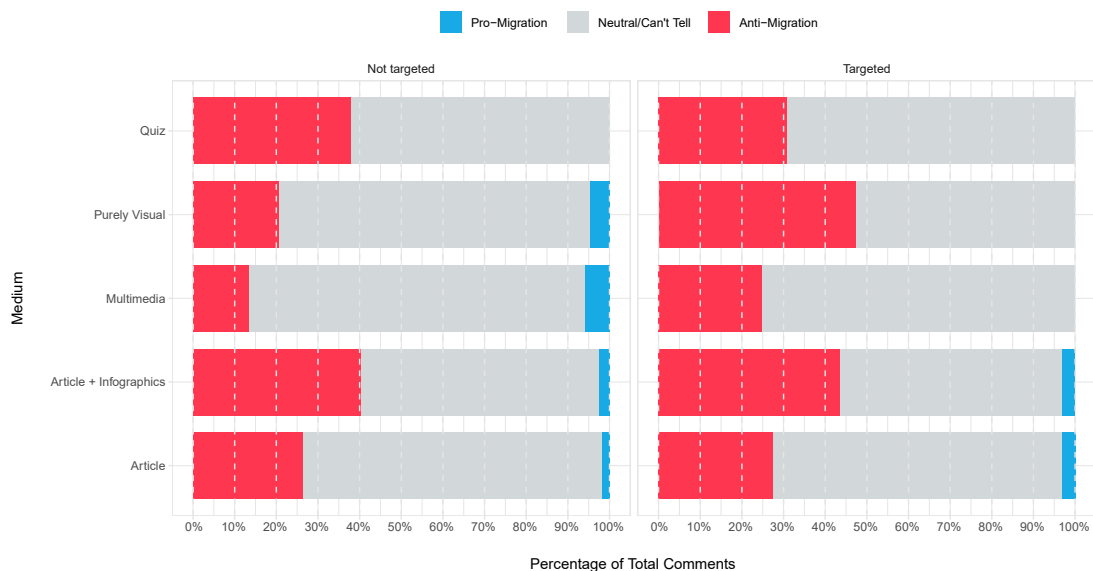


Figure 41. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined by medium and target.

Toxicity is also a factor to consider, and it seems that the articles with infographics got slightly more toxic comments in targeted posts (14% vs 10%), suggesting that they made polarisation more acute in this context. The number of toxic comments on content that focused on data (which often featured infographics) rose in the more targeted environment.

The rest of the mediums, however, did not display significant differences between the audiences, with the exception of the standard article, which attracted 12% of toxic comments when the audience was non-targeted and 8% when the audience was intentionally polarised.

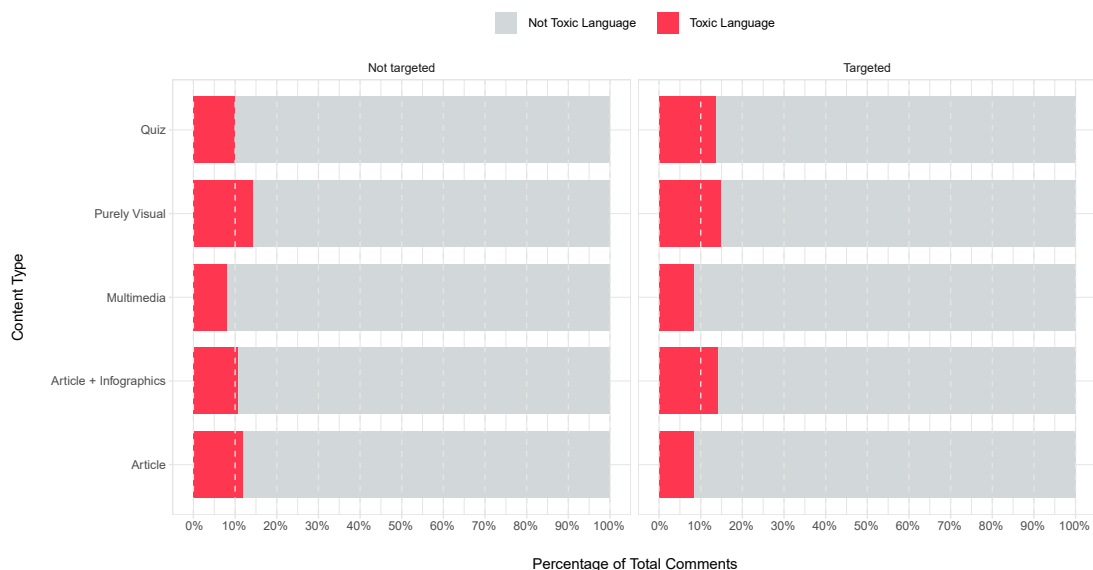


Figure 42. Proportion of total comments that are toxic, broken down by medium and target.

2 - Fact-checking Pieces

From the pieces that focused on data we separated six specially created fact-checking-style pieces and compared them to a selected set of techniques. While many *Corriere* pieces can include elements of fact-checking, these were specially designed to focus on debunking commonly held myths: is it true that migrants receive €35 per day, for example, or that they spread diseases? This analysis consisted of 90 non-targeted posts of which six were fact-checking pieces.

News reports had much higher cumulative average engagement than all the other techniques, including fact-checking, with 292 median likes, 135 median comments and 76 median shares per post – though we should bear in mind the sample size of news stories was much larger and levels of engagement were influenced by the timeliness of news articles. Fact-checking, however, remained a close second, with 154 median likes, 97 median comments and 74 median shares. Owing to the distribution of the data this was hardly a solid result, as it was probably down to the timely nature of their appearance. Other results were much more significant.

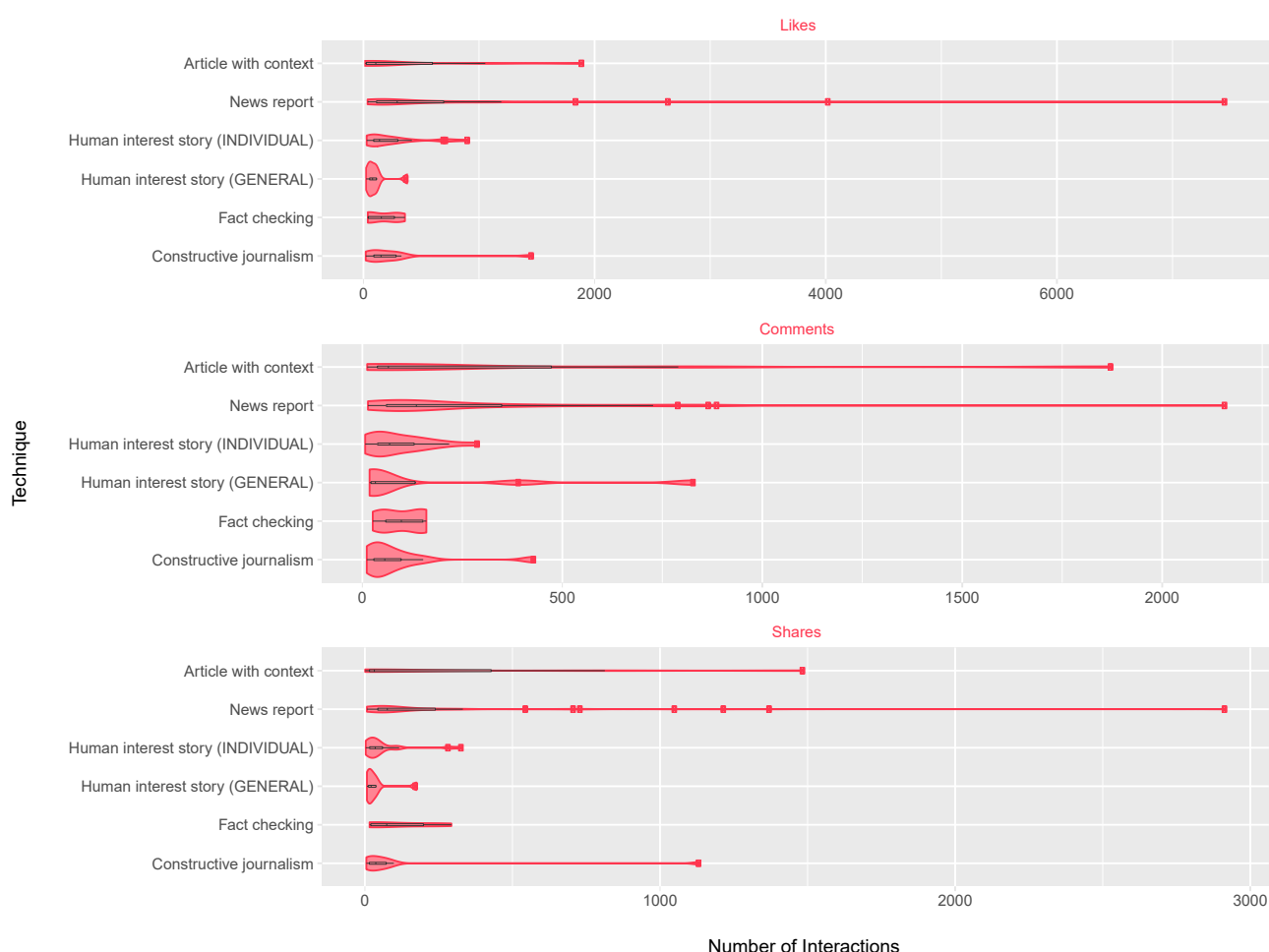


Figure 43. Distribution of the likes, comments and shares of the posts across the different techniques, considering fact-checking.

Fact-checking got slightly fewer toxic comments (8.2%) than the standard news report (9.76%), human interest stories based on groups (9.73%) or articles with context (8.33%), but considerably more than constructive journalism (6%).

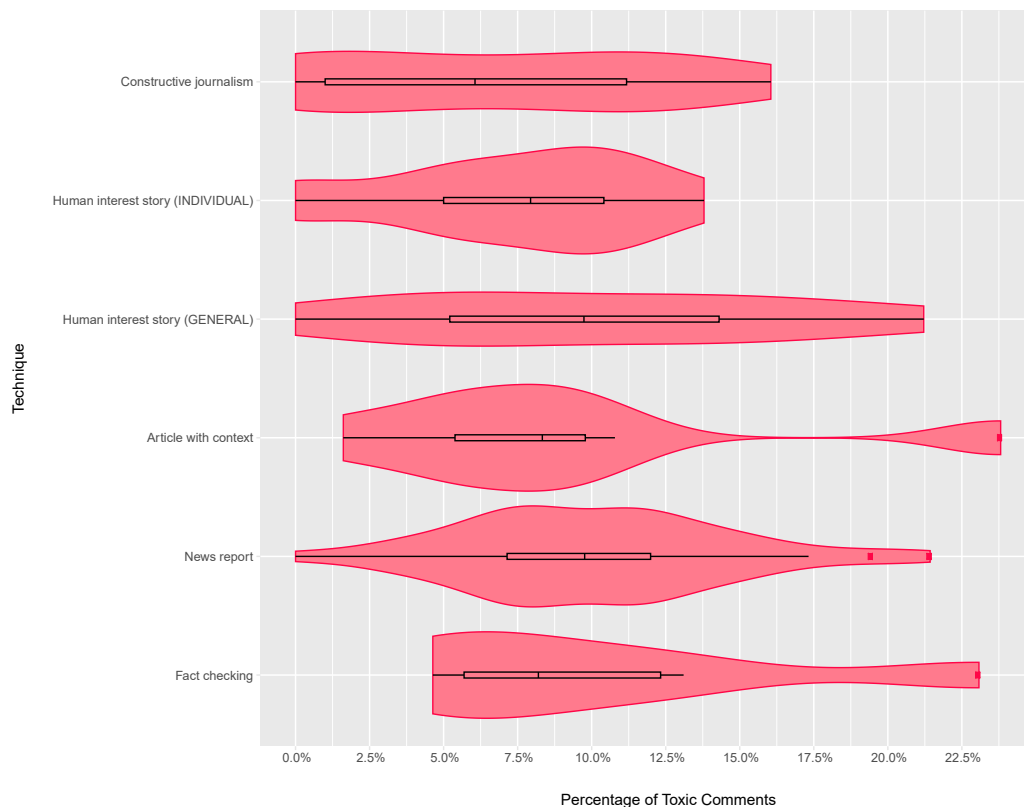


Figure 44. Distribution of the percentage of toxic comments of each post across the different techniques, considering fact-checking.

With regard to the percentage of comments critical of the source, we see that fact-checking got the median highest percentage of critical comments (14.6%), followed by human interest stories about groups of migrants (12.3%). The rest of the techniques get considerably less pushback, with the standard news report getting the least at 1.58%.

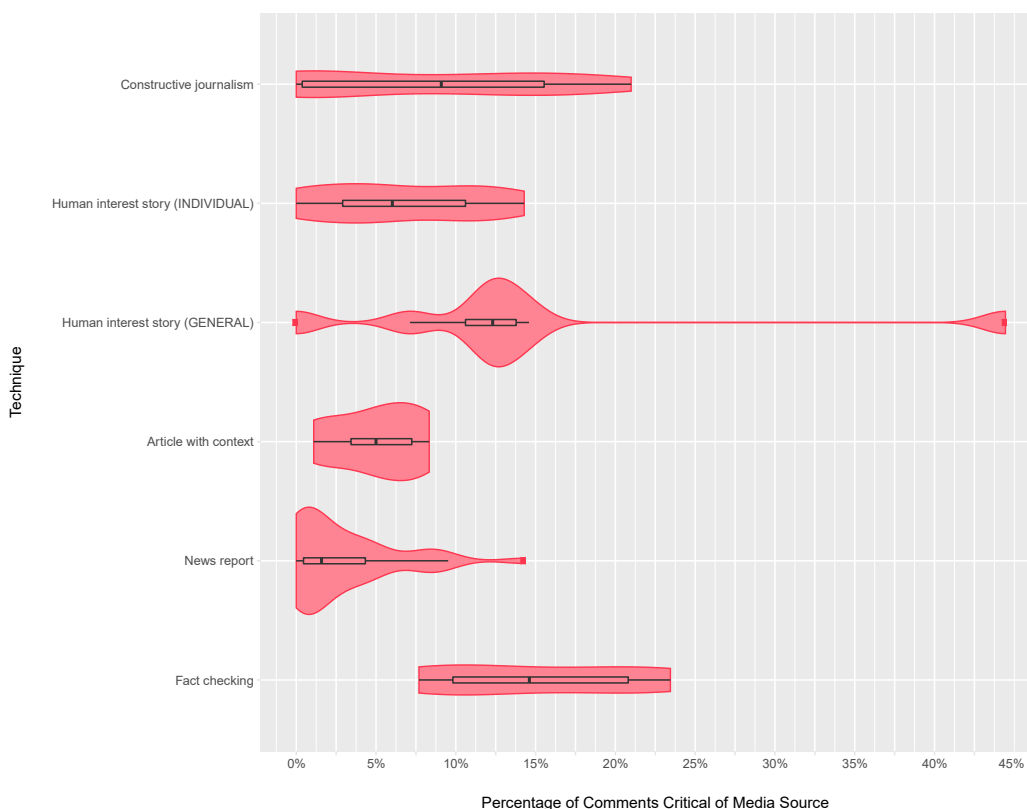


Figure 45. Distribution of the percentage of comments critical of the source of each post across the different techniques, considering fact-checking.

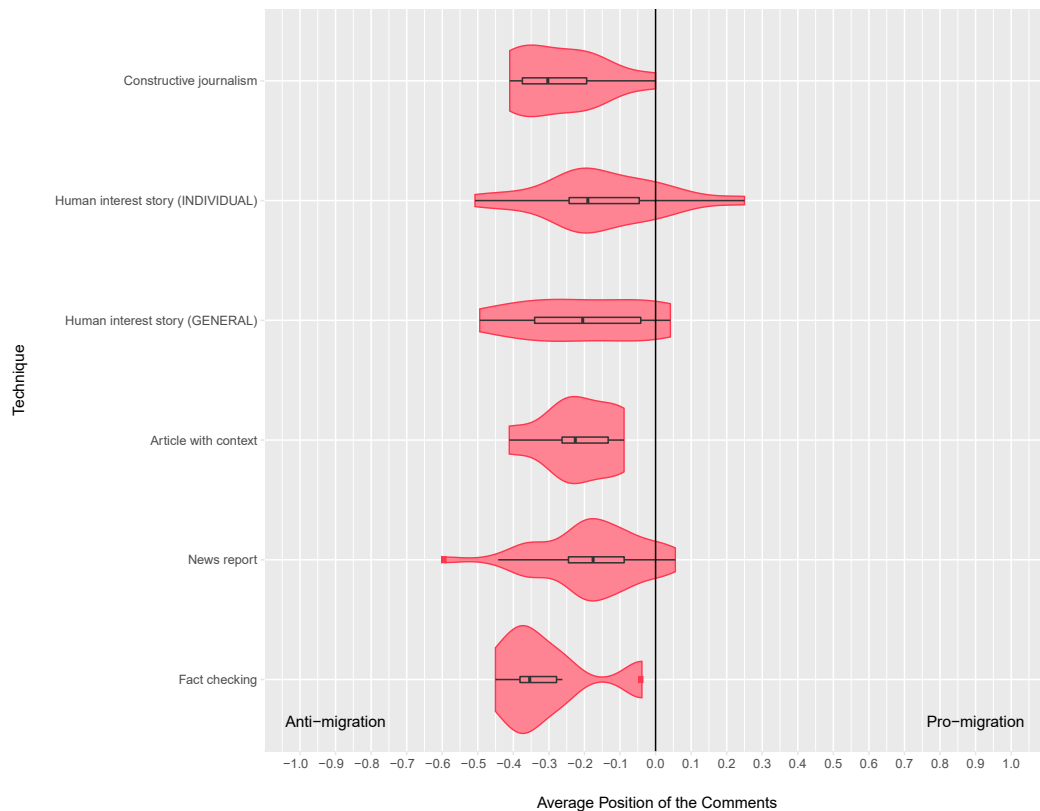


Figure 46. Distribution of the average migration position of the comments of each post across the different techniques considering fact-checking, where -1 is fully anti-migration, 0 is neutral or undefined and 1 is fully pro-migration.

Fact-checking also got the most anti-migration comments, with a median value of -0.35 (where -1 indicates all the comments are anti-migration and 1 that they are all pro-migration). Constructive Journalism also got a large number of anti-migration comments (-0.3) and the standard news report gets the least (-0.17).

All this seems to indicate that fact-checking generates more pushback from readers but the conversation is not highly charged with toxic language. Data do not generate much support – or much emotion.

3 - The Effect of a Thumbnail: Data vs Photo

For one article *Corriere* experimented with two different types of thumbnails – the headline picture that you see on Facebook – to advertise the story. The message of the repost reflects that of the article linked on Facebook.

The story was about how to use the money saved from the arrival of fewer refugee ships to integrate the refugees already in the country. One version used an infographic in the thumbnail, while the other showed a photo of smiling migrants picking up leaves.

Both posts were, at the same time, directed to two different audiences: the untargeted general audience of the *Corriere* Facebook page; and clearly polarised users, who were either pro- or anti-migration.



Corriere della Sera

5 October 2018 · 🌐

I richiedenti asilo in Italia sono ancora migliaia: integrarli conviene. E per farlo bisogna renderli economicamente autosufficienti



CORRIERE.IT

Migranti, ecco come e perché investire i soldi risparmiati grazie al calo degli sbarchi nell'integrazione

Figure 47.

Headline: Italy is spending less time on the arrivals of migrants – integrating those who are already inside the country is the best way to spend the savings: here's why

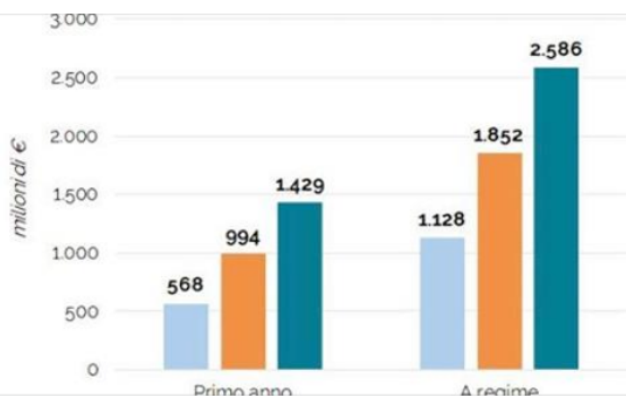
Copy: The number of asylum seekers is still high in Italy – integrating them is a great way to spend Italy's money. The best solution is to make them economically independent)



Corriere della Sera

28 September 2018 · 🌐

Politiche per l'integrazione ben finanziate possono evitare che rifugiati e richiedenti asilo pesino a lungo sulle casse dello Stato e persino trasformarli in un beneficio sia per le entrate fiscali sia per il Paese nel complesso



CORRIERE.IT

Migranti, dal calo degli sbarchi 2 miliardi di euro all'anno: ecco perché conviene investirli nell'integrazione

Figure 48.

Headline: Migrants, 2 billion euros each year coming from the reduction in arrivals: That's why it's better to invest in integration.

Copy: Well-financed integration policies may help asylum seekers and refugees avoid becoming a longterm burden on Italy's national finances. Such policies can even become a benefit in the form of tax revenues and for Italy as a whole.)

As regards overall engagement, the post with the photo showing migrants out-performed the one with the infographic in all the metrics – that is, likes, comments and shares. This was true for both the non-targeted and the targeted versions.

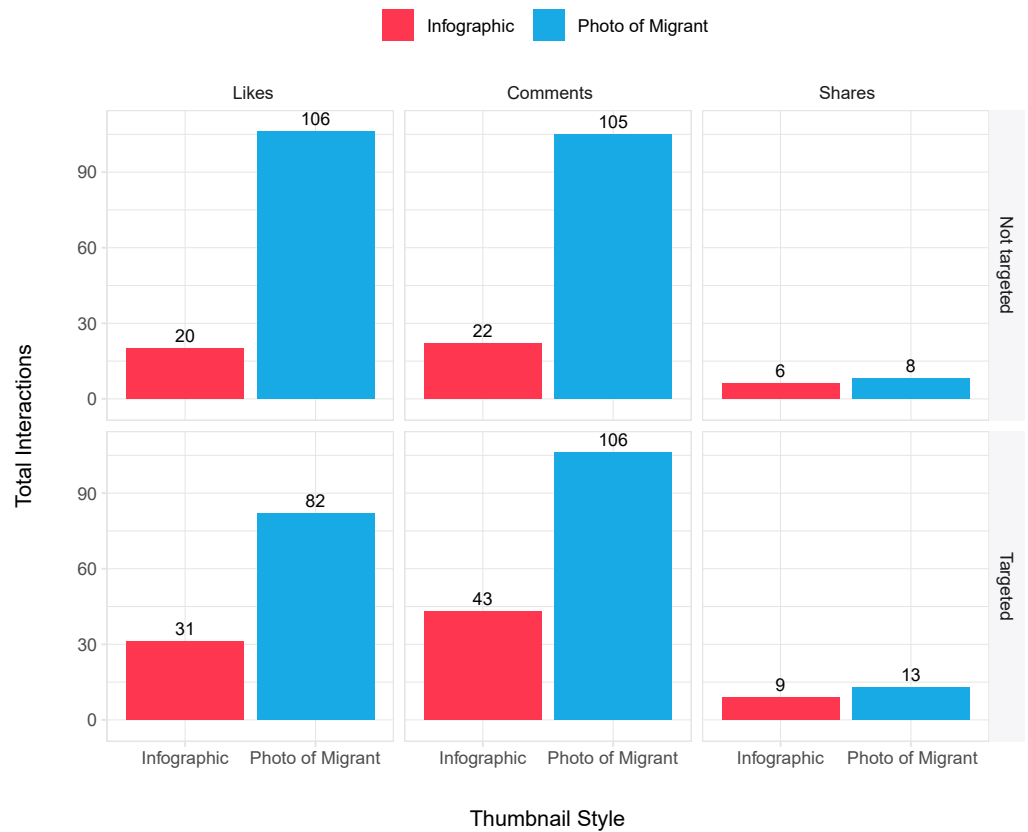


Figure 49. Total likes, comments and shares received by the 2 posts with different thumbnails.

In terms of the comments, the non-targeted post with the photo of the smiling migrants received a higher percentage of toxic comments (11.76%) and criticism of the source (15.69%) than the non-targeted data piece (5.25% and 10.53% respectively). In the targeted posts, however, the article with the photo of the migrants received a higher percentage (13.83%) of critical comments (9.76%) but a lower percentage of toxic comments (12.77%) than the data piece (17.07%).

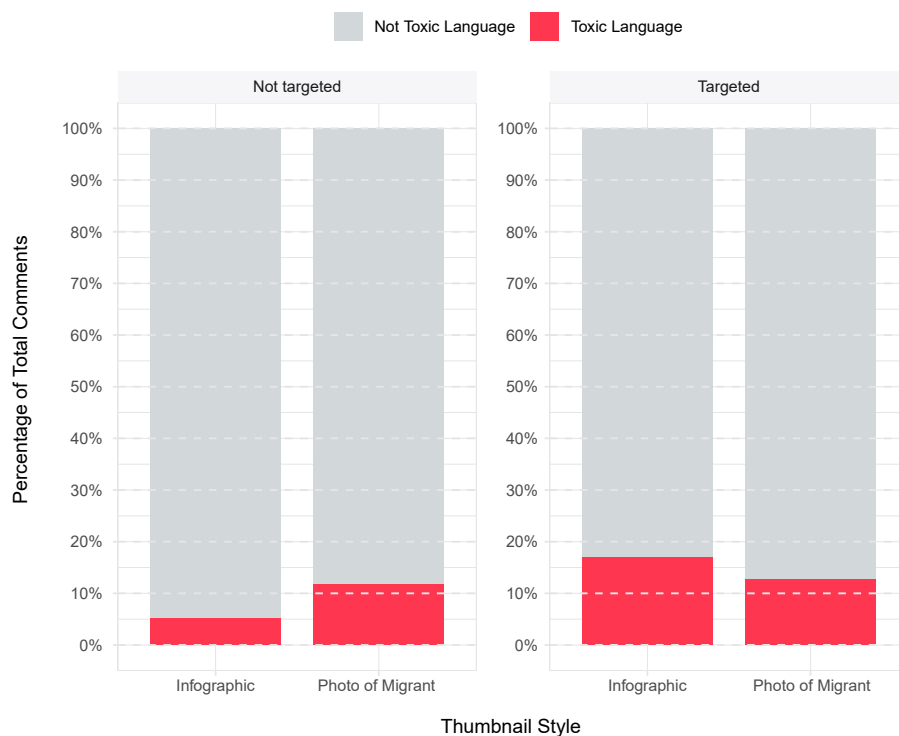


Figure 50. Proportion of toxic comments given to each post with different thumbnails.

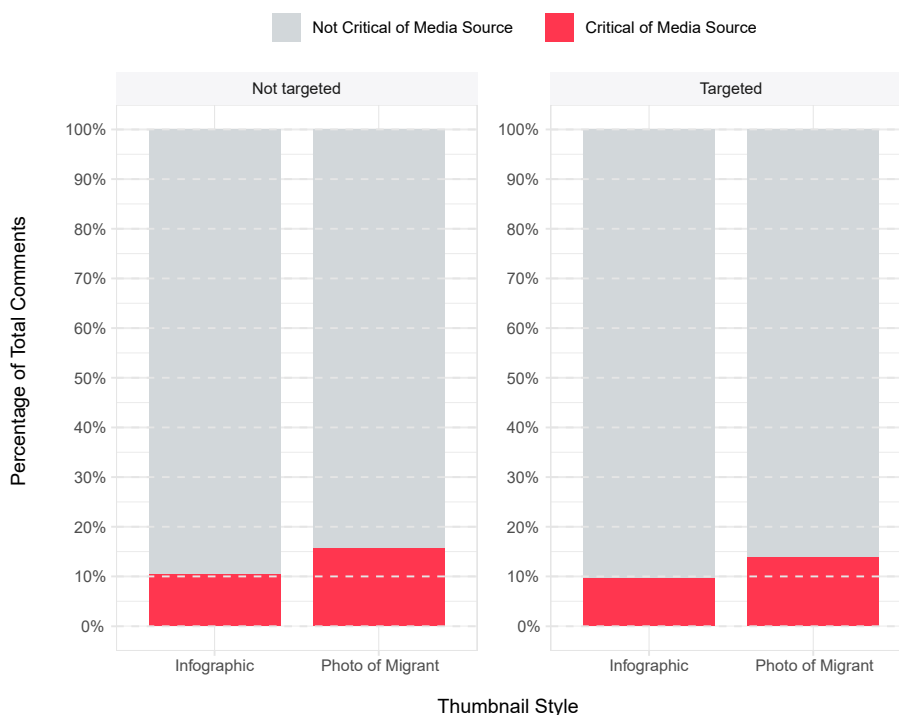


Figure 51. Proportion of comments critical of the source given to each post with different thumbnails.

Looking at the position with regard to comments on migration, we noticed that the percentage of anti-migration comments did not differ substantially between the two non-targeted thumbnails, though we observed a higher percentage of pro-migration comments in the non-targeted data piece at the expense of comments that were neutral about migration. In contrast, we observed that the percentage of anti-migration comments was higher for the targeted article with the photo of the migrant than for the piece displaying data.

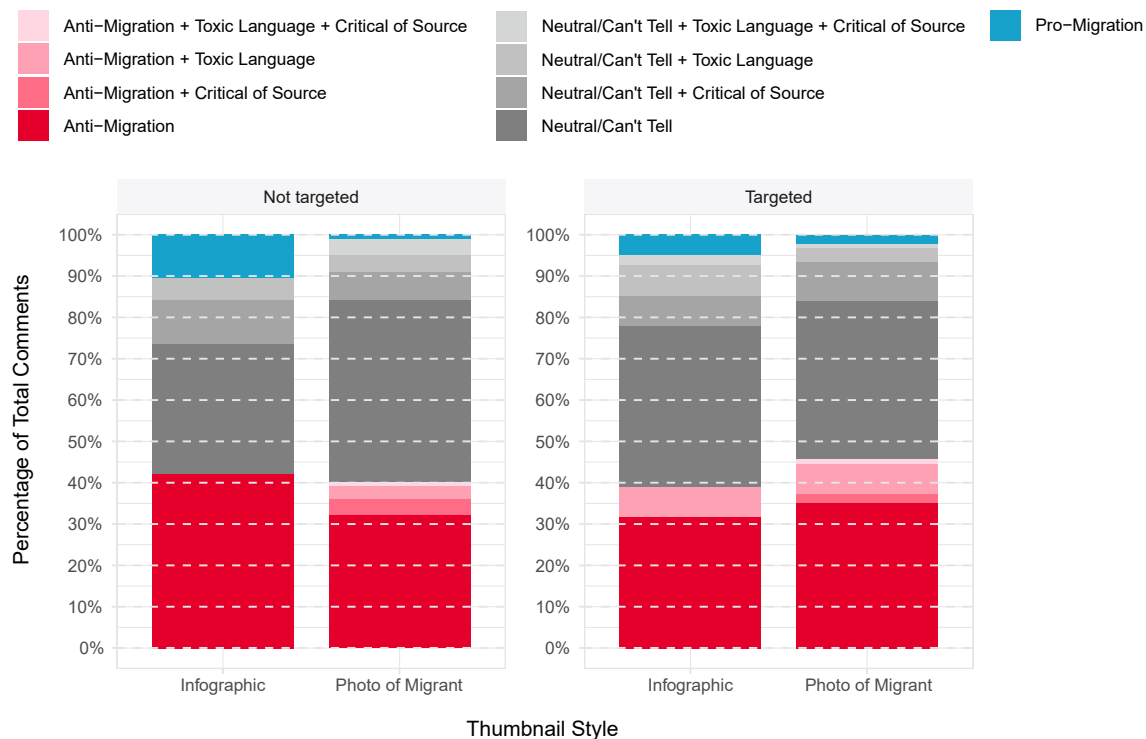


Figure 52. Proportion of all comments that are anti-migration, pro-migration and neutral/undefined while also considering the toxic and critical of source labels of each of the posts with different thumbnails.

To summarise, the article with the photo of the migrants got higher overall engagement, suggesting that an emotional picture may capture readers' attention. In terms of the general audience, that is non-targeted posts, it elicited more comments critical of source and a higher proportion of toxic comments, suggesting it inflamed push back.

| Annex A: Methodology

1 - Introduction: Structure of the Data Analysis

This project aimed to analyse data from social media platforms to monitor readers' reaction to fact-based journalistic outputs concerning the migration issue in Italy. The project developed in two main stages based on the datasets.

In the first stage, we monitored the very recent historical trends related to the online impact of news concerning the migration issue in Italy. This first step was implemented relying on all the Facebook posts published by *Corriere della Sera* between January 2017 and December 2018, as well as on the Twitter public messages published by the official accounts of the major national Italian news outlets between March 2018 and December 2018.

In the second stage, we focused our attention on a set of 165 Facebook posts (and corresponding articles) published by *Corriere della Sera* between March and December 2018 concerning the recent migration issue in Italy. In these articles, several different techniques, targets and content types were tested.

2 - Data

2.1 - Data description

The Historical Dataset gathered data from two platforms: Facebook and Twitter.

From Facebook, only content from the *Corriere della Sera* Facebook page was downloaded – that is, all the posts published from 1 January 2017 to 31 December 2018. We also collected all the reactions given by the users on those posts, as well as all the public comments. The entire Facebook dataset comprises 65,600 posts, 5,181,577 comments and 35,037,478 reactions. A breakdown of the dataset can be seen in Table 1.

	Historical Dataset First Stage	Experimental Dataset Second Stage
Posts	65,060	165
Likes	26,023,596	65,517
Reactions	35,037,478	96,943
Comments	5,181,577	26,886
Users	2,769,763	70,591

Table 1. Number of Facebook posts, likes, reactions, comments and users in each of the stages.

From Twitter, we downloaded the timelines of 52 Italian news outlets covering the time range between March and December 2018. The dataset includes tweets, i.e. Twitter public messages, and a sample of retweets, analogous to share actions on Facebook.

After a preliminary analysis of the data, we decided to focus exclusively on the 21 national outlets. The final dataset comprises 228,921 tweets and 555,044 retweets from 21 accounts.

	Name	Website	Handle
1	Ansa	https://www.ansa.it/	Agenzia_Ansa
2	Avvenire	https://www.avvenire.it/	Avvenire_NEI
3	Corriere della Sera	https://www.corriere.it/	Corriere
4	Huffington Post IT	https://www.huffingtonpost.it/	HuffPostItalia
5	Il Blog di Beppe Grillo	https://beppegrillo.it/	beppe_grillo
6	Il Fatto Quotidiano	https://www.ilfattoquotidiano.it/	fattoquotidiano
7	Il Foglio	https://www.ilfoglio.it/	ilfoglio_it
8	Il Giornale	https://www.ilgiornale.it/	ilgiornale
9	Il Manifesto	https://ilmanifesto.info/	ilmanifesto
10	Il Messaggero	https://www.ilmessaggero.it/	ilmessaggeroit
11	Il Sole 24 Ore	https://www.ilsole24ore.com/	sole24ore
12	L'Espresso	https://espresso.repubblica.it/	espressonline
13	La Repubblica	https://www.repubblica.it/	repubblica
14	La Stampa	https://www.lastampa.it/	LaStampa
15	LA7	https://www.la7.it	La7tv
16	Libero Quotidiano	https://www.liberoquotidiano.it/	Libero_official
17	MSN Italia	https://www.msn.com/it-it	msn_italia
18	RAI News	https://www.rainews.it/	RaiNews
19	Rai.TV	https://www.raiplay.it/	RaiPlay
20	Sky TG24	https://tg24.sky.it/	SkyTG24
21	TgCom24	https://www.tgcom24.mediaset.it/	MediasetTgcom24

Table 2. List of the national Italian newspapers and their respective Twitter accounts.

In the second stage, the Experimental Dataset consisted of 165 posts on migration published by *Corriere della Sera* between May and December 2018, with all their respective user interactions.

3 - Analyses

3.1 - Detection of Migration-related Content in the Historical Dataset

In order to differentiate the posts and tweets related to migration we established a filter that tried to match a selected list of words to the content of the posts or tweet, and the linked article. To do this we extracted the article from the website to avoid false positive matches with the context of the article – that is, related articles, relevant news, and

ads; and we tokenised the text to compare the words individually. We kept count of the number of matches between each post or tweet with our set of keywords and split the results into three categories:

- No migration-related keywords: no matches were found in the post/tweet and article; thus the content is not related to migration.
- 1 migration-related keyword: only one match was found in the post/tweet and article.
- 2+ migration-related keyword: two or more matches were found in the post/tweet and article; thus the content is migration-related.

The list of words is as follows: *asilo politico, diritto d'asilo, rifugiato, rifugiata, rifugiati, rifugiate, richiedente asilo, richiedenti asilo, richiesta di asilo, richieste di asilo, richiesta d'asilo, richieste d'asilo, clandestino, clandestina, clandestini, clandestine, immigratorio, immigratoria, immigratorii, immigratorie, immigrante, immigranti, immigrare, immigrato, immigrata, immigrati, immigrate, immigrazione, immigrazioni, migrante, migranti, migrare, migratore, migratrice, migratori, migratrici, migratorio, migratoria, migratorii, migratorie, migrazione, migrazioni, emigrante, emigranti, emigrare, emigrato, emigrata, emigrati, emigrate, migratorio, migratorii, migratoria, migratorie, emigrazione, emigrazioni, musulmano, musulmana, musulmani, musulmane, islam, baobab, burqa, hijab, niqab, ramadan, sharia, moschea, moschee, guardia costiera, ong, sprar, unhcr, unicef, aquarius, ius soli, ius sanguinis, extracomunitario, extracomunitaria, extracomunitari, extracomunitarie, razziale, razziali, razzismo, razzista, razzisti, razziste, razza bianca, negro, negra, xenofobia, xenofobie, xenofobo, xenofoba, xenofobi, xenofobe, barcone, barconi, sbarco, sbarchi, scafista, scafisti, scafiste, naufragio, naufragi, porto sicuro, protezione umanitaria, decreto sicurezza, corridoio umanitario, corridoi umanitari, sprar, bossi-fini, regolamento di Dublino, rimpatrio, rimpatri, espatrio, espatri, rimpatriare, espatriare, profugo, profuga, profughi, profughe, eunavformed, frontex, oseghale, sacko.*

The breakdown of the posts and tweets by matches with the relevant keywords can be seen in Tables 3 and 4.

	All Tweets	2+ Migration-related Keyword	1 Migration-related Keyword	No Migration-related Keywords
Tweets	228,921	19,167	13,392	196,362

Table 3. Number of tweets considering the presence of migration related keywords in the posts and/or article.

	All Posts	2+ Migration-related Keyword	1 Migration-related Keyword	No Migration-related Keywords
Posts	65,060	3,469	2,638	58,953
Likes	26,023,596	1,549,471	1,044,045	23,430,080
Reactions	35,037,478	2,140,672	1,462,944	31,433,862
Comments	5,181,577	537,687	316,442	4,327,448
Users	2,769,763	560,246	527,229	2,649,753

Table 4. Number of Facebook posts, likes, reactions, comments and users considering the presence of migration related keywords in the posts and/or article.

3.2 - Analysis of Engagement

The analysis of engagement aimed to assess the level of users' interaction with on-line-published contents. Different online platforms allow for different possible interactions. For example, Facebook allows users to interact with posts by leaving a reaction (like, love, haha, wow, sad, angry), commenting or sharing the post. Analogously, Twitter allows users to interact with a tweet by leaving a favourite, commenting and retweeting.

We exploited basic engagement analysis to investigate how users interacted with contents published online concerning the migration issue in Italy. In this respect, engagement analysis may offer useful insights concerning the themes, techniques and content types preferred or rejected by users.

3.3 - Sentiment Analysis

Sentiment analysis in its basic form tries to extract sentiment polarity from text. The simplest form of this analysis consists of giving each word in the text either a polarity label (positive, neutral or negative) or a score between -1 and 1, where positive values mean the word is positive, zero means the word is neutral and negative values mean the word is negative, and the higher the score, the greater the strength of the sentiment in the respective polarity. In order to establish the sentiment of the text, the labels are counted or the scores are averaged in order to establish the overall sentiment. There are many improvements to simple sentiment analysis, achieved by considering the various complexities of the text, such as the type of word (adjective, noun, verb, etc.), negation, punctuation or emojis.

In order to extract sentiment from posts and comments appearing on the *Corriere della Sera* Facebook page, we relied on Watson Natural Language Understanding, an IBM service which exploits natural language processing techniques to derive sentiment from text.

3.4 - Stance Analysis and Comment Coding Activity

In order to assess the stance of *Corriere's* readers with regard to the migration issue, we hired three Italian mother-tongue students at LSE to annotate the set of 26,886 users' comments in the Experimental Dataset with the perceivable stance conveyed through each comment. Some of the comments were replicated in order to compare the agreement between the coders and the self-agreement of each coder on a subset of the comments.

The comment coding activity was carried out on Goldfinch,⁴¹ an online annotation platform brought by the **Department of Knowledge Technologies** of the Jožef Stefan Institute of Ljubljana. The annotators were allowed to display a sample of the comments on *Corriere's* posts devoid of any information about either the commenter or the post. In more detail, the annotators were asked to annotate three characteristics of the comments: the conveyed position on migration, the presence of rude [offensive?] language and the criticism directed towards the publisher/journalist/newspaper.

With regard to the position on migration, the annotators were asked to determine whether a comment was in favour of, against or neutral about the migration phenomenon in Italy. The annotators were also allowed to express uncertainty about their judgement.

41 http://annotate.ijs.si/goldfinch/kt/_home/Index.

With regard to the presence of toxic language, we asked the annotators to signal whether a comment explicitly contained vulgar, aggressive, rude, disrespectful or unreasonable language which could lead a user to abandon a conversation. We explicitly asked them to disregard comments containing irony or sarcasm.

Lastly, the annotators were asked to signal any comment directly criticising the publisher/journalist/newspaper as a source of information.

The annotators also had the option of removing from the comments set those comments considered not relevant (empty, not in Italian or spamming).

These features were decided as the result of an initial experiment in which the researchers at Ca' Foscari University and the journalists at *Corriere della Sera* tried to code multiple features in a set of 500 comments. Based on the results of this experiment, we tightened the definitions of the features and settled on the simplest definitions that yielded the greatest consensus: position on migration, toxic speech and criticism expressed towards the journalist/publisher/newspaper. One of the possible features we considered was "constructive comment" but the resulting low agreement on this indicated that the definition was too subjective to yield useful results.

The exact instructions provided to the annotators for their comment coding activity on Goldfinch were as follows:

For every comment choose:

- *The position on the migration phenomenon: choose thumb up if the comment is explicitly pro immigration, the circle if the comment does not express any explicit position regarding immigration or thumb down if the comment is explicitly anti immigration. If you are not sure of your choice, check the Not Sure box. Please note that comments in favor of or against some politicians do not necessarily express a pro/anti immigration position.*
- *The presence of rude language: Check the box Toxic if the comment explicitly contains vulgar, aggressive, rude, disrespectful, unreasonable language which could lead the user to abandon a conversation. The word Toxic should not include irony or sarcasm - the language used must be explicitly rude.*
- *Criticism of the editor/journalist/newspaper: Check the box Critical of the source if the comment directly criticizes the publisher or journalist or newspaper as a source of information.*

In addition, you [the annotator] can decide to exclude the comment by clicking on the trash icon if it does not seem relevant (it's spam, it's not in Italian, it's empty, it only tags a person, etc.). If you want to annotate the comment at a later time, you can skip the comment by clicking on the double arrow.

4 - Data Visualisation: Violin Plots

A violin plot allows for synthetic representation of the distribution of observed data and comparison of different distributions. It is very similar to box plots but they also show the probability distribution of the data at different values, that is, how likely a given value is.

On the one hand, a box plot is a simple representation of data starting from five summary values of a distribution: the median, the upper and lower quartiles, and the upper and lower extremes. The median is represented by a thick vertical line within a rectangle or box, whose width is determined by the distance between the upper and lower quartiles. Two horizontal lines, called whiskers, are then traced respectively between the lower extreme and the lower quartile and between the higher quartile and the higher extreme. Values beyond the lower and upper extremes are called outliers and are depicted as dots.

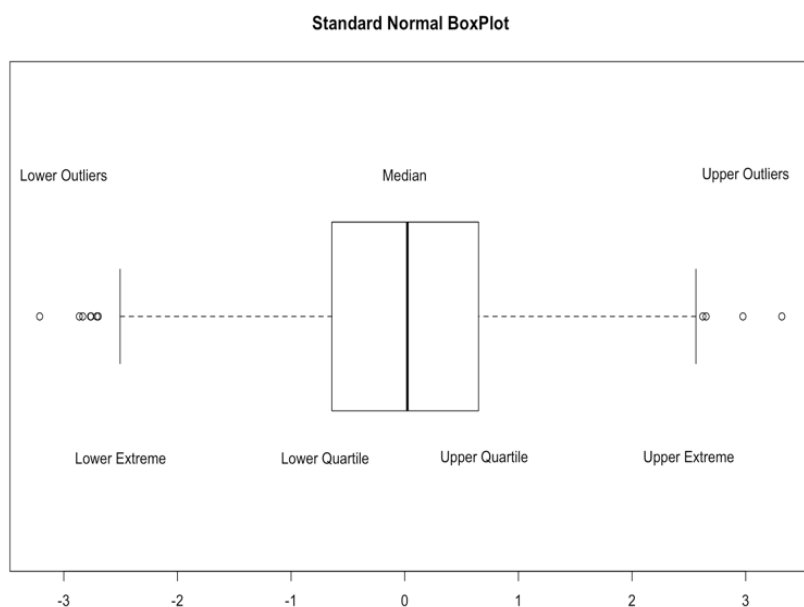


Figure 53. Example of a boxplot.

A violin plot, on the other hand, will include all the data that is in a box plot (a marker for the median of the data; a box or marker indicating the interquartile range) and the probability density of the data rotate on the sides where the bigger the density the higher the frequency in the data.

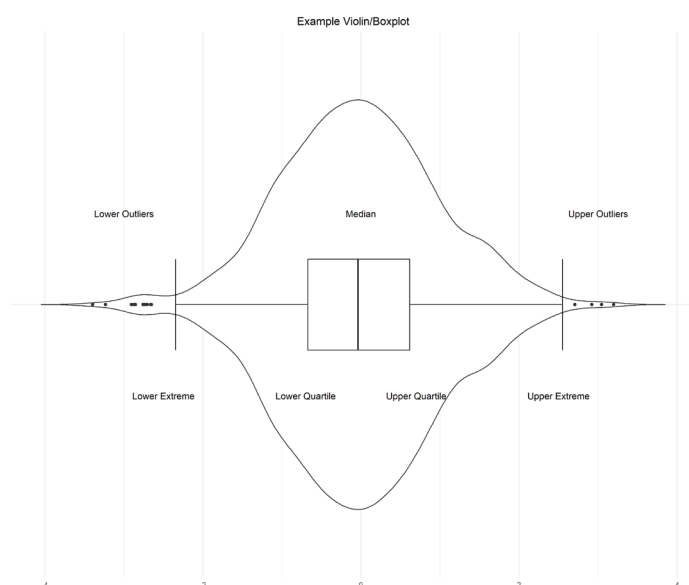


Figure 54. Example of a violin plot.

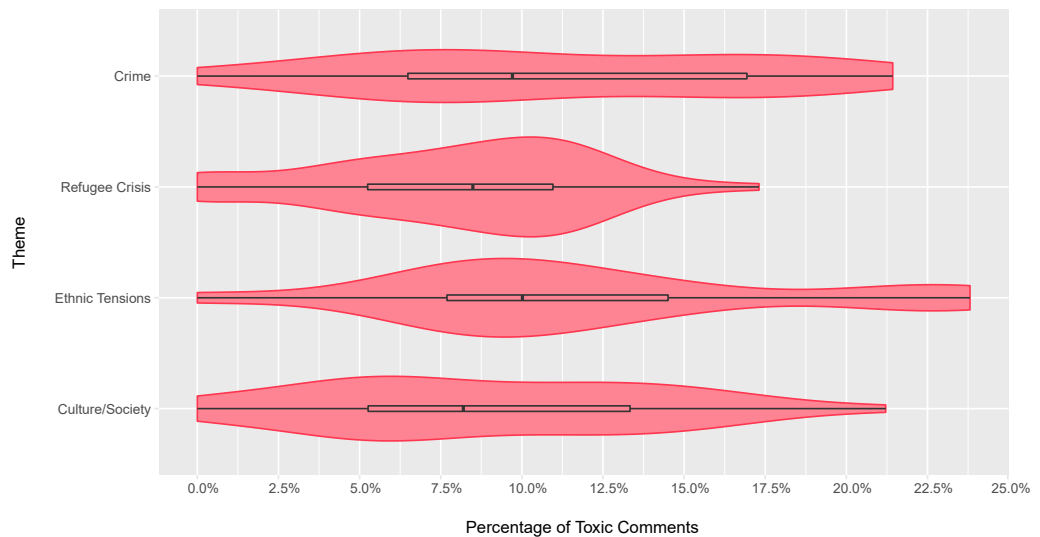


Figure 55. Example of a violin plot as shown in our report.

In our work, we relied on violin plots to compare engagement, sentiment and stance for different groups of posts. For example, Figure 55 displays the violin plots for the level of language toxicity present in the comments of several posts grouped according to the theme of the story.

5 - Community Detection

Community detection allows identification of the presence of sub-communities within the Twitter network. In particular, in our research, we exploited community detection to identify communities of news outlets having in common interaction with the same Twitter users.

We started by tracing the interactions between anonymised users and the Twitter accounts of the major Italian news outlets. The result was a bipartite network formed by a set of users and a set of pages where links only existed between a user and a page if the user retweeted anything on that page.

We can represent the bipartite network with a matrix where each column is a user and each row is a page, and the content of each cell equals 1 if the user retweets content of that page, and 0 otherwise. If we multiply the matrix by its transpose, we get the projection of the bipartite network. This new matrix will have a row and column for each page, and the content of each cell will represent the number of common users between the two pages that define that cell – that is, the number of users who retweet content on both pages.

Once we have the network of pages linked by their common users, we can apply different community detection algorithms to detect communities – that is, groups of pages that are strongly connected.

Figure 56 offers a graphical representation of the process. First, the bipartite network linking users to the Twitter accounts of news outlets is created (panel (a)). From that, we derive the weighted network pairwise, linking news outlets according to the number of common users (panel (b)). Finally, the community detection algorithm enables the presence of communities within the network (panel (c)) to be discovered. In the example below, the algorithm detects the presence of two communities (community 1: blue nodes and community 2: red nodes).

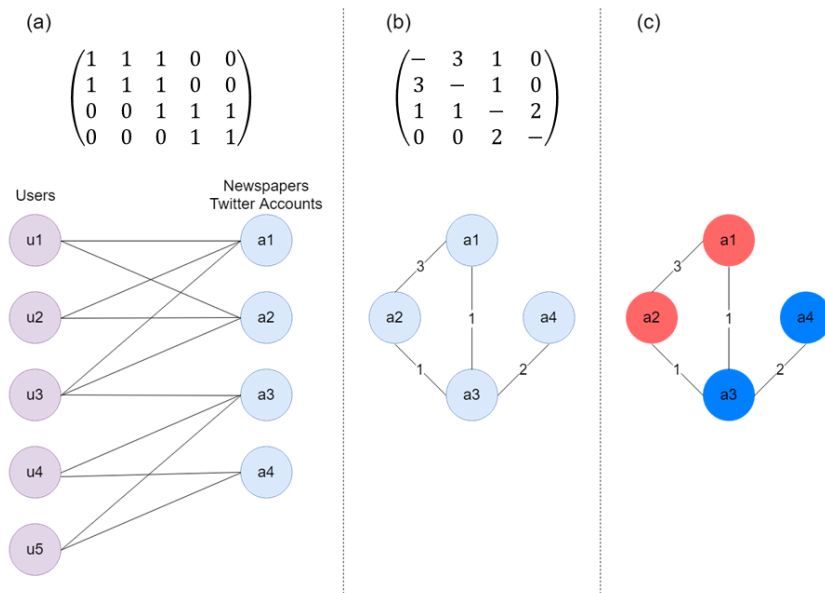


Figure 56. Diagram shows the different steps taken to obtain the Twitter communities.

(a) shows a small example of a bi-partite network of Twitter users and Twitter Newspaper accounts with its corresponding matrix. (b) shows the projection of that network, with its corresponding matrix. (c) shows how the different communities the nodes belong to (red or blue) when applying the FastGreedy algorithm to the network.

Annex B: Data Management and GDPR Compliance

1 - About the Data Collection Process and Management

The project collected data from the Facebook account of *Corriere della Sera*. Specifically, it collected personal data from Facebook users and public entities. The project also collected data from Twitter, in particular from public Twitter accounts of Italian news outlets. For this reason, the collected data have been managed in agreement with GDPR EU 679/2016.

- NO geographical data were collected, directly or indirectly, for the research.
- NO genetic data were collected for the research.
- NO biometric data were collected for the research.
- NO data concerning health were collected for the research.
- NO children were involved in the research.

Data were stored in the form of tables (CSV, RData).

1.1 - Facebook Data

The entire Facebook data collection process was performed exclusively by means of the Facebook Graph API,⁴² which is publicly available through one's personal Facebook user account. We used only publicly available data. Users with privacy restrictions were not included in our dataset. Data were downloaded from Facebook pages that were public entities. When allowed by users' privacy specifications, we accessed public personal information, which was limited to the users' interactions with the Facebook page of *Corriere della Sera*. Moreover, in this project we used fully anonymised and aggregated data.

We abide by the terms, conditions, and privacy policies of Facebook.⁴³ Owing to an update to Facebook API policy on 5 February 2018, we are not allowed to access any information about the users who reacted/commented to Facebook content on public pages. Such information was available only for the *Corriere della Sera* Facebook page, where we have administrative privileges. In that case, we could access the Facebook IDs of users interacting with the content of the page. Facebook IDs were encrypted via the hash algorithm SHA-512. The algorithm allows a unique message of fixed length, called a digest, to be obtained from a starting message of variable length. Currently, obtaining the users' ID from the digest is hardly possible, given the high level of technical expertise and massive computational effort required. For this reason, the adopted encryption algorithm should be considered, within the context of this project, as an adequate and proportionate means to ensure users' anonymity and discourage any data breach attempt.

⁴² <https://developers.facebook.com/docs/graph-api>.

⁴³ <https://developers.facebook.com/docs/graph-api/reference/v2.10/comment>.

Data collection was performed at three levels: on *Corriere's* posts, on reactions to *Corriere's* posts, and on comments responding to *Corriere's* posts. The specific data we gathered related to *Corriere's* posts were: the numeric ID and the name associated with the publishing account (*Corriere*), the message contained within the post, the date and time at which the post was initially published, the type of post (link, photo, video etc.), the link attached to the post, the post ID, the "story" description associated with the post, the aggregated number of likes, comments and shares, and the numeric ID associated with the page on which the post is published.

Moreover, users' reactions were collected for each post published by *Corriere*. The specific data gathered for each reaction were: the numeric ID of the user (encrypted with SHA-512 and thus anonymous), the type of reaction ("angry", "hah", "like", "sad", "wow"), and the numeric ID of the post to which the reaction refers.

Finally, users' comments were collected for each post published by *Corriere*. The specific data gathered for each comment were: the commenter's numeric ID (encrypted with SHA-512), the comment's message, the date and time at which the comment was published, the aggregated number of likes and sub-comments, and the numeric ID of the comment.

1.2 - Twitter Data Collection

The entire data collection process on Twitter was performed exclusively by means of the Twitter API,⁴⁴ which is publicly available and can be used through one's personal Twitter user account. We used only publicly available data. Users with privacy restrictions were not included in our dataset. Data were downloaded from Twitter accounts that are public entities. We abide by the terms, conditions, and privacy policies of Twitter.

Data include public tweets made on the timelines of the specific accounts established in our list of Italian news outlets as well as their number of retweets, replies and favourites. The total list of followers of each account were also downloaded and, in some cases, a random sample of retweeters, where the users' IDs were encrypted with SHA-512. In this case also, the encryption makes it hardly possible to obtain the users' ID from the digest, owing to the high level of technical expertise and the massive computational effort required.

2- Data Analysis

The adopted data-elaboration techniques include: overall text refinement, text filtering and topic extraction procedures. For the experimental data, the data-analysis techniques include: annotation task, frequency analysis, audience engagement analysis, sentiment analysis and stance analysis techniques. The results of the analysis are presented in aggregation and no personal information is disclosed.

3 - Annotation Task

The annotation task was carried out with the external support of Dr Petra Kralj Novak, researcher in the [Department of Knowledge Technologies](#) of the Jožef Stefan Institute (JSI) of Ljubljana. This external support was scrutinised and approved by all the parties

⁴⁴ <https://developer.twitter.com/en/docs/tweets/search/api-reference.html>.

involved in the project. The annotation task was performed on the dedicated platform Goldfinch,⁴⁵ managed by JSI, and employed 26,886 anonymised Facebook users' comments from the Experimental Dataset, inclusive of the encrypted numeric IDs and the associated textual comments. JSI commits to delete data and task results immediately after task completion, and in any case no later than the project end date.

4 - Managing and Storing Data

4.1 - Data Storage

Anonymised raw data and results are stored indefinitely by Ca' Foscari for research needs and purposes.

4.2 - Data Management and Storage Facilities

Data destination is EU27. Data are stored in two machines of the Laboratory of Data Science and Complexity (DSCLab) located in the Department of Environmental Sciences, Informatics and Statistics (Building ZETA, Room Z.B03) at Ca' Foscari University. Access to data on the storage is subject to authentication using username and password. Only the researchers involved in the project have access to the data.

Both machines are connected to the Internet. The network is administered by the Ca' Foscari IT team following the best practices and standards available. Machines are not accessible externally, but through VPN.

4.3 - Data Preservation Strategy and Standards

Data have been shared only among the project participants; they are therefore closed access.

4.4 - Main Risks to Data Security

No significant risk is expected. Backup copies of data have been made according to the best security procedures and practices.

5 - Responsibilities

Data collection, processing, management and storage are carried out by the Laboratory of Data and Complexity (DCLab) of Ca' Foscari University. Data are collected, processed and managed under the responsibility of the project coordinator Dr Walter Quattrociocchi.

⁴⁵ http://annotate.ijs.si/goldfinch/kt/_home/Index.



**Journalism In The Age of
Populism and Polarisation:**
Insights from the Migration
Debate in Italy

June 2019