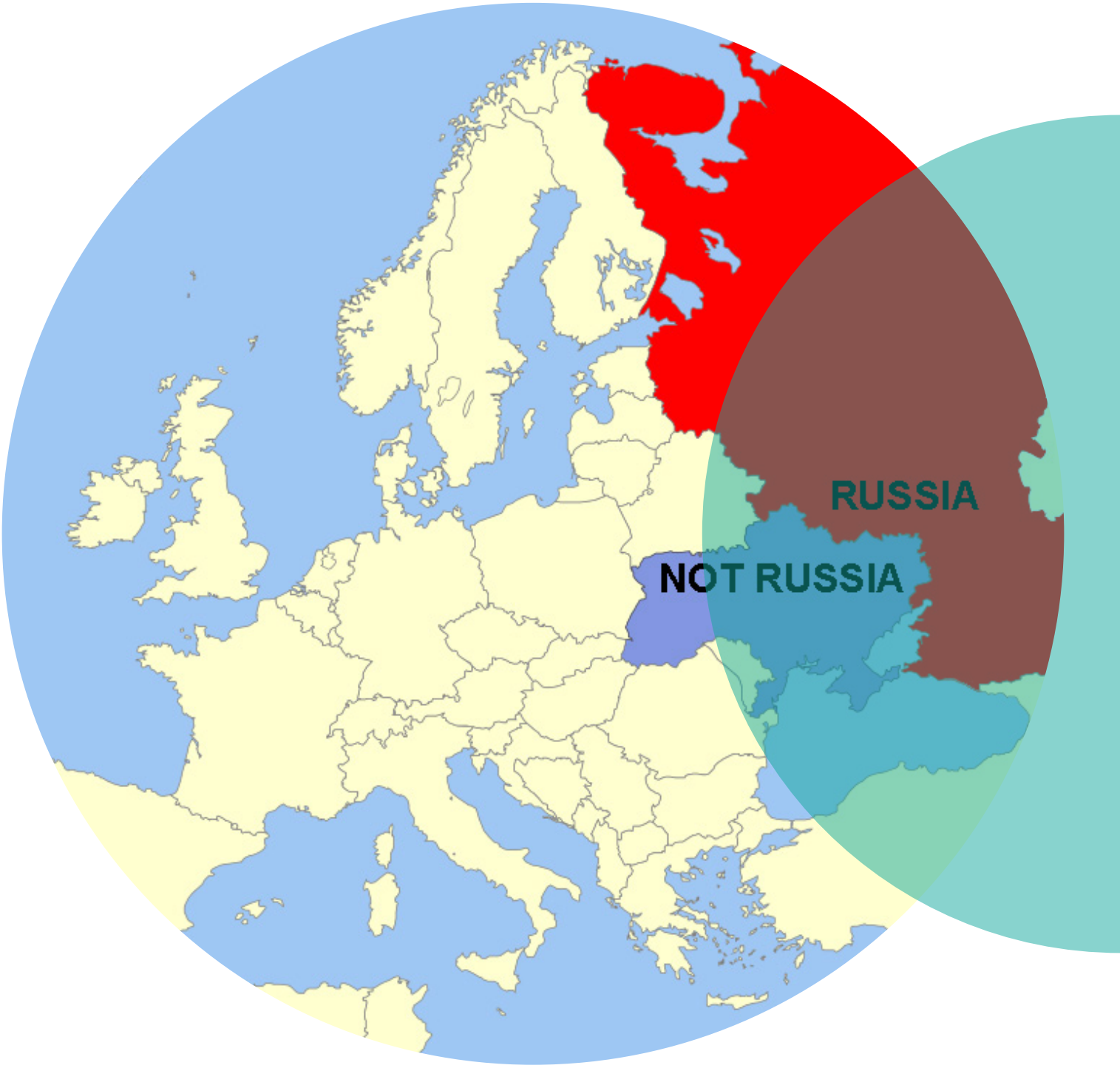
Hybrid CoE Working Paper 26

Humour in online information warfare: Case study on Russia's war on Ukraine



Hybrid CoE

Keir Giles – November 2023

Hybrid CoE Working Papers cover work in progress: they develop and share ideas on Hybrid CoE's ongoing research/workstrand themes or analyze actors, events or concepts that are relevant from the point of view of hybrid threats. They cover a wide range of topics related to the constantly evolving security environment.

The European Centre of Excellence for Countering Hybrid Threats
tel. +358 400 253800 | www.hybridcoe.fi

ISBN (web) 978–952–7472–84–2
ISBN (print) 978–952–7472–85–9
ISSN 2670–160X (web)
ISSN 2814–7235 (print)

November 2023

Cover photo: Geigle / Wikimedia Commons:
https://creativecommons.org/licenses/by-sa/4.0/deed.en

Hybrid CoE's mission is to strengthen its Participating States' security by providing expertise and training for countering hybrid threats, and by enhancing EU-NATO cooperation in this respect. The Centre is an autonomous hub for practitioners and experts, located in Helsinki, Finland.

Contents

Summary

State-backed parody and mockery of the enemy in conflict are nothing new. What is new is the way social media has democratised access and audience. This has both opened the playing field to self-motivated private individuals and facilitated their joining forces in informal collectives for greater effect. The result has been to impose costs on adversaries – if only by making it harder for them to achieve their aims.

In particular, humour-based responses to Russian actions in the information space and in the physical domain have been found to deliver multiple clear benefits for the defending side, even if the direct impact on Russia itself is hard to measure. This has led to recognising the value of adopting similar methodologies on a more formal basis. Success in this endeavour is difficult to measure objectively. Nevertheless, authoritative studies conclude that ridicule has been shown to work with general audiences.

This Hybrid CoE Working Paper considers instances of humour put to effective use to counter disinformation and propaganda in online spaces, using Russia's war on Ukraine as a primary case study, and Twitter (now X) as the main medium under consideration. It is a practical review seeking to identify examples of best practice from both government and civil society.

In Ukraine, humour was adopted as a coping mechanism in the earliest stages of the armed conflict after 2014. The specific potential of memes as a method of serious persuasion for political purposes has also been noted. Russia's actions during and after 2014 also prompted a response by private citizens on social media in the form of parody accounts mimicking prominent Russian individuals or organisations. This paper highlights a series of examples.

As effective a counter-disinformation tool as fact-checking and debunking is, this remains a very labour-intensive process. Ukraine combines humour with a range of different measures to combat Russian information operations, including debunking but also discrediting and blacklisting the sources and conduits for Russian disinformation.

Rather than tackle the disinformation challenge head-on through debunking, humour-based tactics often take an indirect approach to achieving their objectives, discrediting the messengers of disinformation and their narratives through ridicule rather than deconstruction.

This approach also reverses the roles: instead of trolls and propagandists sucking their opponents into futile arguments that achieve nothing, if the conversation in itself is ridiculous, it is instead against their interests to engage in it.

As an aide to delivering maximum effect, this paper offers several criteria for achieving impact. They include:

- appreciating the hypersensitivity of the adversary to ridicule as a key enabler;
- the importance of knowing one's audience, or in other words cultural sensitivity as an essential component of successful humour, in contrast to Russian counterparts' repeated failure at attempted humour;
- government communicators' willingness to accept risk as they endeavour to embrace humour;

- and the fact that, as argued here, achieving success with humour-based engagement is easiest when it leverages the agility of organic, unregulated organisations of private individuals.

A key advantage of communities of interest built around humour is their accessibility, with no membership criteria or joining mechanism for the movement, and an inbuilt incentive to acquire greater mass through numbers. Humour, this paper argues, is the glue that holds a community of interest together, allowing a mass response to disinformation. Self-starting communities of interest built around humour have also been credited with shoring up support for Ukraine since the humour and creativity help stave off war fatigue.

The overall effect of a community built around humour has been to turn the tables on social media platforms. The agents of influence and other servants of authoritarian regimes, who for so long held the advantage, are turned into the targets rather than the deliverers of mockery and abuse.

While humour cannot replicate the effects of formal monitoring, debunking and user education, neither can these achieve the effects delivered by humour. As such, humour – whether wielded by official government communications agencies, or individuals acting on their own initiative – constitutes an important additional tool for bolstering the resilience of the information environment.

Introduction

Since the widespread recognition of Russian disinformation as a problem, Western governments and expert communities have considered a range of possible solutions and countermeasures over the years, none of which has proved fully effective in isolation. These have included fact-checking and rebuttals (found to be labour-intensive and of limited impact among some audiences); trying to reach Russia's own population with the truth (largely ineffective due to substantial and long-term Russian preparations to prevent such efforts); and hitting back at Russia using its own methods (rejected as incompatible with Western values).

However, the period following Russia's full-scale invasion of Ukraine in February 2022 saw a notable reduction in the reach of Russian and pro-Russian disinformation and propaganda on social media platforms, including but not limited to Twitter. This is a result not only of Russia discrediting its narratives through its own actions, but also of the emergence of informal groups countering these narratives through the use of humour, parody and satire.

While state-backed parody and mockery of the enemy in conflict are nothing new, the democratisation of access and audience offered by social media has opened the playing field to self-motivated private individuals, as well as facilitated their joining forces in informal collectives for greater effect. The result has been to impose costs on adversaries – if only by making it harder for them to achieve their aims. Moreover, these humour-based responses to Russian actions in the information space and in the physical domain have been found to deliver multiple clear benefits for the defending side even if the direct impact on Russia itself is hard to measure. This has led to recognition of the value of adopting similar methodologies on a more formal basis. Adam Kinzinger, CNN commentator and former US Congressman, assesses that: "This will be a future case study in how to counter online propaganda successfully."[1] Furthermore, academic studies have established "a need to develop memetic insurgencies that are participatory for end users to not only support, but actively take part in themselves, and for state actors wishing to deploy them as capabilities in their own right to design them accordingly".[2]

The impact on government communication strategies is already visible. This includes an acceptance – for now, unevenly spread across Western partner nations – that responses to Russia in the information space achieve the best effect when discarding the pretence that engagement with Russian and pro-Russian entities can take the form of a civilised diplomatic conversation. The UK Foreign, Commonwealth & Development Office responded to Russia's Ministry of Foreign Affairs releasing its new Foreign Policy Concept on 31 March 2023 by simply quote tweeting that "April Fool's Day is TOMORROW" – which did more to discredit it than any detailed academic deconstruction could achieve.[3]

This Hybrid CoE Working Paper considers instances of humour in effectively countering disinformation and propaganda in online spaces, using Russia's war on Ukraine as a primary

1 Adam Kinzinger (@AdamKinzinger), Twitter, 27 August 2022, https://twitter.com/AdamKinzinger/status/1563711015348998144.
2 Callum Harvey and Jack Goldsmith, "Building a meme war machine: A comparative analysis of memetic insurgencies in cyberspace", forthcoming.
3 Foreign, Commonwealth & Development Office (@FCDOGovUK), Twitter, 31 March 2023, https://twitter.com/fcdogovuk/status/1641852678106906630.

case study, and Twitter (now X) as the main medium under consideration. It is a practical review rather than an academic study, seeking to identify examples of best practice from both government and civil society that may offer successful models of how to discredit and expose the absurdity of disinformation activities, to disseminate awareness about disinformation beyond audiences already attuned to it, and to impose costs on those who are spreading disinformation by making their work more difficult and their objectives harder to attain.

The paper begins with a discussion of the difficulty of achieving objective measures of success, before considering the role of humour as part of a range or toolkit of counter-disinformation measures. It then describes some of the apparent essential criteria for achieving positive impact through the use of humour, and the positive side effects delivered by this approach. Finally, the paper concludes with a brief consideration of lessons learned.

Measuring effect

The seizure of Crimea in 2014 refocused Western minds on the information threat due to Russia's highly successful exploitation of the information environment both to accomplish the operation with little resistance, and to escape the consequences of it. Since that date, one of the key questions that has repeatedly been posed to students and practitioners of information warfare is that of metrics: how, exactly, is it possible to measure the impact both of information activities and countermeasures against them? The answers are notoriously complex, and difficult and expensive to implement (one key reason why policymakers continue to ask the question nine years later, perhaps in the hope that a simpler or cheaper answer might eventually emerge).

When considering the role of humour in countering disinformation, the problem becomes even more intractable. Assessments of humour are subjective – jokes can fall flat, and conversely some things described as failed attempts at humour in this paper might be uproarious to some readers. Consequently, the difficulty of measuring whether something is 'funny' presents an additional complexity layered upon the already existing challenge of measuring whether it has any impact on attitudes or behaviours. For instance, an authoritative guide to combatting conspiracy theories concludes that "ridicule has been shown to work with general audiences", but does not offer any practical and scaleable means of measuring just how well it works.[4] In the absence of objective measures of either phenomenon, it becomes necessary to fall back on criteria for success or failure that are imperfect or entirely subjective. These include simple (but potentially misleading) counts of views and likes to measure the virality of online content, and personal impressions and anecdotal evidence for the effect of humour.[5]

Examples of anecdotal evidence include personal reporting by owners of influential humour or parody accounts. The Darth Putin Twitter account collects examples of Russian propaganda tweets which they conclude from extensive observation were deleted after being mocked. Their theory is that this may be due to internal pressure to perform. "We have a feeling that when they have budget meetings, if it looks like all they are doing is getting mocked, that's not great. They have to show progress," Darth observes. Darth also thinks that a new habit of Russian official accounts putting out "on this day" anniversary tweets three days early and at odd times of the night is a response to the widespread ridicule.[6] "They have a flunky tweet at midnight to try to own hashtags."[7] Darth's conclusion is that "we've yet to hear another way of doing this that is as effective as what we do".[8] When this subjective

4 Stephan Lewandowsky & John Cook, *The Conspiracy Theory Handbook* (2020). https://skepticalscience.com/conspiracy-theory-handbook-downloads-translations.html.

5 Researcher Olga Tokariuk plans to examine the impact of humour in information warfare in detail in a forthcoming paper. See Monique Camarra, "Olga Tokariuk, Humour vs Disinfo: How Ukraine is resisting the Information War", Substack, 20 March 2023, https://camarra.substack.com/p/mar-20-olga-tokariuk-humour-vs-disinfo.

6 For example, the Russian MFA tweet regarding the anniversary of the Molotov-Ribbentrop pact on 23 August, posted on 20 August. See https://twitter.com/mfa_russia/status/1428369500555018243.

7 Author interview with @DarthPutinKGB, 13 August 2021.

8 Author interview with @DarthPutinKGB, 13 August 2021.

and circumstantial evidence of direct impact is combined with readily observed side effects of the use of humour, its standing as an effective measure becomes stronger – especially in combination with other possible lines of effort like fact-checking, debunking, or working with platforms to remove malicious content.

Another criterion that can at least be measured objectively is whether a given information activity succeeds in crossing barriers from its original environment. This is referred to by leading scholar of disinformation Ben Nimmo as the "Breakout Scale", a measure of impact ranging from category one for an event that passes practically unnoticed, through five for operations that are amplified in real life by high-profile influencers, and six for operations that lead to a discernible policy response or concrete action.[9] While Nimmo devised the scale to categorise the impact of disinformation, the scale can equally well be applied to efforts at humorous countering of disinformation – and when this is done, it can be seen that this "breakout" effect is regularly achieved. This report contains multiple instances of humorous interventions on social media being reported in mainstream media, and each of these represents a crossing of boundaries – a "breakout" of at least category four in Ben Nimmo's terminology. But gaps that can successfully be bridged by means of humour include not just those between platforms and media, but between audiences. One successful example has been the Lithuanian television show *Laikykitės ten* (*Hang in there*) with comedian Andrius Tapinas, lampooning Russian propaganda and disinformation not only for its original Lithuanian-speaking audience but now also in a Russian-language version.[10]

Nevertheless, none of these measures of effect provides a clear and objective answer to the question of whether a given counter-disinformation initiative has "worked" – still less of whether it is funny. This presents a key challenge for policymakers considering humorous approaches to tackling their disinformation problem, since in the absence of empirical metrics or evidence for what works, they are faced with the need to take risks and invest trust based on criteria that are intangible or unquantifiable. The effects of this challenge on "official" attempts at humour will be discussed further below.

9 Ben Nimmo, "The Breakout Scale: Measuring the impact of influence operations", Brookings, September 2020, https://www.brookings.edu/articles/the-breakout-scale-measuring-the-impact-of-influence-operations/.

10 Daniel Schearf, "Baltics' Russian Media Use Online Humor to Combat Propaganda", VOA, 27 March 2017, https://www.voanews.com/a/russian-media-located-in-baltics-use-online-humor-to-combat-propaganda/3784002.html.

Humour as part of the counter-disinformation toolkit

Old ideas, new impact

There is of course nothing new about the use of humour in conflict in and of itself, either as a morale-booster for one side or as a means of attacking the other. From the UK, the "Wipers Times" series of trench newspapers produced close to the frontline during the First World War represents a well-preserved example of the use of humour to cope with horror.[11] Later, in the Second World War, both overt parody and covert satire were employed to subvert Nazi Germany.[12]

In Ukraine itself, humour has also been adopted as a coping mechanism from the earliest stages of the armed conflict after 2014, providing "solidarity and stress-relief".[13] Academic studies built on this observation to consider the potential validity of humour as an online strategic communications tool, noting that "humour is not only entertaining, funny, satirical, and joke-laden, but is also among the foundations of group identity, and can therefore be a tool for strategic communication".[14] More recently, humour had already been highlighted during the coronavirus pandemic as a useful tool for raising awareness and expanding reach to new audience sectors.[15]

Study of the specific potential of memes as a method of serious persuasion for political purposes also predates the current war on Ukraine.[16] But information activities during the conflict have emphasised how "memes may serve to structure and shape perceptions of an event or action... [showing] the value in information warfare of a memetic, ironic and highly participatory quality".[17] This was demonstrated clearly by one of the first viral instances of a formal government communications channel using ridicule to deliver its message: a tweet by the Joint Delegation of Canada to NATO in August 2014 in response to reporting that Russian soldiers entering Ukraine were "lost".[18] The map of "Russia" and "Not Russia", with the caption "Geography can be tough. Here's a

11 "The Wipers Times: The soldiers' paper", National Army Museum (UK), https://www.nam.ac.uk/explore/wipers-times.

12 "How Britain fought Hitler with humour", BBC, 31 August 2019, https://www.bbc.com/culture/article/20190829-how-britain-fought-hitler-with-humour.

13 See Maksym Kyiak, "Case study: Use of humour for solidarity, denigration and stress-relief in the Ukrainian media during the Russian aggression in 2014–2016", in Žaneta Ozoliņa et al., *StratCom laughs: in search of an analytical framework*, NATO Strategic Communications Centre of Excellence, 2017, https://stratcomcoe.org/publications/stratcom-laughs-in-search-of-an-analytical-framework/201.

14 Žaneta Ozoliņa et al., *StratCom laughs: in search of an analytical framework*, NATO Strategic Communications Centre of Excellence, 2017, https://stratcomcoe.org/publications/stratcom-laughs-in-search-of-an-analytical-framework/201.

15 Jakub Kalenský, "How to Defend Against Covid Related Disinformation", in Ritu Gill and Rebecca Goolsby (eds.), *COVID-19 Disinformation: A Multi-National, Whole of Society Perspective* (Springer, 2022), https://link.springer.com/chapter/10.1007/978-3-030-94825-2_7.

16 "A virtual army of impish cartoon pooches is waging war on Russia", *The Economist*, 31 August 2020, https://www.economist.com/europe/2022/08/31/a-virtual-army-of-impish-cartoon-pooches-is-waging-war-on-russia.

17 Callum Harvey and Jack Goldsmith, "Building a meme war machine: A comparative analysis of memetic insurgencies in cyberspace", forthcoming.

18 Canada at NATO (@CanadaNATO), Twitter, 27 August 2014, https://twitter.com/CanadaNATO/status/504651534198927361.

guide for Russian soldiers who keep getting lost & 'accidentally' entering #Ukraine" was novel and effective – in part because at the time it seemed a daring and outré departure from staid diplomatic norms.[19] By 2023, the trend set by the Canadian tweet is far more commonplace and accessible to governments, with even the German Foreign Office making creditable attempts at humour to undermine Russian talking points.[20]

Russia's actions during and after 2014 also prompted a response by private citizens on social media in the form of parody accounts mimicking prominent Russian individuals or organisations. This represented a transition from a long tradition of domestic political parody to addressing geopolitics, and was novel in that it was not a government propaganda effort but an entirely spontaneous response to world events by concerned citizens, now armed with a voice and reach that would have been inconceivable in prior centuries. Examples in the pre-2022 period included @SovietSergey, parodying Russian Foreign Minister Sergey Lavrov, and the @Sputnik_Not account, posting in the style of straight news headlines with accompanying images to ridicule Russia's Sputnik propaganda outlet.[21] Among these, the Darth Putin account – @DarthPutinKGB – achieved a breakthrough to virality and rapidly increasing follower numbers precisely as a result of Russian attempts in May 2016 to close it down – indicating that even at this early stage this was a challenge taken seriously by the targets of the parody.[22]

Darth Putin describes the techniques used by these accounts as holding a mirror up to Russian state propaganda, with the effect of exposing its ludicrousness.[23] "Their statements get more and more ridiculous. And we just shine a torch on that... we shine a light on what they are doing that lets a wider audience understand it from a different perspective," Darth Putin explains.[24] The approach can be to "wittily compare historical fallacies created by Russia with primary sources that refute their claims", or on occasion simply to repeat statements by Russian sources verbatim in order to highlight their ridiculousness.[25] The wide range of examples of implementation of humour as a defensive weapon in information and influence derives from its universality as a tool. This also means that humour-based campaigns deliver a number of positive side effects as well as their primary intended impact, to be explored further below.

19 Brett LoGiurato, "Canada's NATO Delegation Just Epically Trolled Russia With This Map Of Ukraine", *Business Insider*, 27 August 2014, https://www.businessinsider.com/canada-nato-russia-ukraine-putin-2014-8?r=US&IR=T.

20 See e.g., tweet by @GermanyDiplo on 24 January 2023, https://twitter.com/GermanyDiplo/status/1617834372232785920?s=20.

21 See https://twitter.com/Sputnik_Not.

22 "Twitter restores Putin parody account", BBC News, 1 June 2016, https://www.bbc.co.uk/news/blogs-trending-36429074.

23 Author interview with Twitter user @DarthPutinKGB, 13 August 2021.

24 Author interview with Twitter user @DarthPutinKGB, 13 August 2021.

25 Connor Forberg, "NAFO: A Masterclass on Upending the Russian Propaganda Machine in the Modern Era", course paper, Georgetown University, 13 March 2023.

Debunk or deride?

The period since 2014 has seen proposals from a wide variety of government bodies and think tanks calling for serious efforts to counter Russian propaganda online through fact-checking and debunking. In this context, Ukrainian journalist and former presidential media advisor Iuliia Mendel thinks voluntary online groups (such as NAFO, described further below) have a key role in countering disinformation through achieving critical mass in disseminating truth. "The only way to fight fake news is to be ready to tell the truth as many times as the fake news is repeated," she believes, and thus volunteers can make a difference by "repeat[ing] truth consciously, fighting Russian propaganda".[26]

Multiple academic studies have argued that debunking is an effective counter-disinformation tool among specific audiences, although few of them argue that this means resources should be devoted to rebutting every single piece of false information that is encountered, and second-order effects including an overall reduction of trust in media output have been identified.[27] Logic-based, fact-based and source-based debunking each have roles in mitigating the effects of disinformation among audiences that will be receptive to these methods.[28] Ukraine combines humour with a range of different measures to combat Russian information operations, including a heavy emphasis on debunking but also discrediting and blacklisting the sources and conduits for disinformation, and proactive steps like shutting down Russian and Russian-backed media outlets in the country. Reliance on monitoring and debunking false information as a foundation for these other efforts comes despite recognition of its labour-intensive nature, a concern offset by the widespread involvement of volunteers and civil society in contributing manpower.[29]

However, a debunking or fact-checking approach to countering disinformation relies on the audience applying the time and effort to read the "true" version of events, and having an interest in discovering it in the first place. Thus debunking is effective primarily among audiences that could be expected to possess all of these attributes, such as government officials, analysts, academia, and most journalists. On social media, by contrast, the interest of the majority of audience members is likely instead to be both superficial and fleeting, meaning that the impact of a counter-disinformation measure has to be far closer to instantaneous than a lengthy exposition of why a given piece of information is wrong.

Consequently, multiple studies have noted the limited utility of contradiction and debunking in countering disinformation specifically on social media.[30] In this context, debunking

26 Speaking at "NAFO and winning the information war: lessons learned from Ukraine", CSIS, 5 October 2022, https://www.csis.org/analysis/nafo-and-winning-information-war-lessons-learned-ukraine.

27 See e.g, https://e.america.gov/t/ViewEmail/i/8D41EB2341B3EE972540EF23F30FEDED/F2AB8F86D-C5635A4AF060D6555554232; https://journals.sagepub.com/doi/10.1177/20563051231179694; https://www.nature.com/articles/s44159–021–00006-y.pdf; https://www.climatechangecommunication.org/wp-content/uploads/2020/10/DebunkingHandbook2020.pdf.

28 Lewandowsky & Cook, *The conspiracy theory handbook*.

29 Author interviews, April–July 2023.

30 E.g., Puneet Bhargava et al., "How effective are TikTok misinformation debunking videos?" *Misinformation Review*, Harvard Kennedy School, 29 March 2023, https://misinforeview.hks.harvard.edu/article/how-effective-are-tiktok-misinformation-debunking-videos/.

remains a labour-intensive process but with far less likely return on investment of resources; and the associated measure of reporting hostile profiles relies on the platform hosting them being sufficiently responsive and interested to take action, which is far from guaranteed. (The undermining of enforcement on Twitter, now rebranded as X, makes that platform in particular an environment that is even more permissive for Russia and hostile for its critics.[31]) In this environment, although debunking is not redundant, humour-based approaches have a number of specific attributes that can render them more effective.

Rather than tackle the disinformation challenge head-on through debunking, humour-based tactics often take an indirect approach to achieving their objectives, discrediting the messengers of disinformation and their narratives through ridicule rather than deconstruction. In academic studies, ridicule of the cognitive dissonance required to subscribe to conspiracy theories has been identified as an alternative approach to logical argument in attempting to debunk them.[32] One example of these tactics was the UK's response in 2014 to Russia's implausible yet dogged denials of its troops' presence in Ukraine's southeast.[33] Russia's distinctive T-72BM tank was featured in a 'guide' to help Russia spot its own tanks in Ukraine, tweeted by the British Embassy in Kyiv on 19 November 2014.[34] The embassy also tweeted this in Ukrainian, and the tweet garnered international attention.[35]

In practical observation too, Darth Putin thinks discrediting is a more effective approach to countering Russian state announcements than soberly cataloguing the lies they contain. Humour "shows a different way of dealing with the Russian authorities' gaslighting and propaganda and lies [because] arguing with it gives it credibility it doesn't deserve," they note.[36] This points to a key weakness of fact-checking and debunking approaches, namely responding to disinformation arguments or false facts as though they had merit. In addition, the nature of social media platforms leads to debunkers operating at a substantial disadvantage when trying to counter mass information campaigns, including the use of trolls deploying a range of well-established techniques to exhaust time and energy by tying them down in arguments they cannot win.

Furthermore, dismissing propaganda by laughing at it has the further advantage over fact-checking and debunking that it is – if the humour resonates with the target audience – instantly comprehensible, and thus appeals both

31 Caroline Orr Bueno (2023), "Twitter exec says 'hundreds of thousands' of Russian disinformation accounts still active on Twitter", Weaponized Spaces, 13 February 2023, https://weaponizedspaces.substack.com/p/twitter-exec-says-hundreds-of-thousands.

32 Gábor Orosz et al. "Changing Conspiracy Beliefs through Rationality and Ridiculing", *Frontiers in Psychology*, Volume 7, (2016), https://doi.org/10.3389/fpsyg.2016.01525.

33 "Analysis: Russian arms and forces in eastern Ukraine", BBC Monitoring, 9 December 2014, https://monitoring.bbc.co.uk/product/f1c83o3o [requires subscription].

34 See https://twitter.com/UKinUkraine/status/535014678842593280.

35 Agence France-Presse in Kyiv, "British embassy in Ukraine tweets guide to Russian tanks", *The Guardian*, 19 November 2014, https://www.theguardian.com/world/2014/nov/19/british-embassy-ukraine-tweets-guide-russian-tank.

36 Author interview with Twitter user @DarthPutinKGB, 13 August 2021.

to neutral observers and to those with no time or energy to compare competing versions of the truth. This approach also reverses the roles: instead of trolls and propagandists sucking their opponents into futile arguments that achieve nothing, if the conversation in itself is ridiculous, it is instead against their interests to engage in it because "the moment somebody's replying to a cartoon dog online, [they've] lost".[37] Graham Brookie, Senior Director of the Atlantic Council's Digital Forensic Research Lab, notes that disinformation is a trap: "We don't need to fight fire with fire, we need to find out how to get above that."[38] Humour offers one means of avoiding the trap.

Importantly, the two approaches of contradicting propaganda and ridiculing its exponents are not mutually exclusive. Online activist "Jen Bones", posting on Twitter as @gnucontrol, explains:

> *Some of the things I create actually do challenge the Russian narratives directly, and with those I can argue "I'm directly participating in an information war effort". Most of the other stuff would fall into the category of simple mockery. I still think that's also worthwhile. It keeps people interested in and engaged with what's going on, and it promotes coherence in this group of people united by a common enemy.*[39]

37 Matt Moores, speaking at "NAFO and winning the information war: lessons learned from Ukraine", CSIS, 5 October 2022, https://www.csis.org/analysis/nafo-and-winning-information-war-lessons-learned-ukraine.
38 Speaking at the Riga StratCom Dialogue, Riga, 7 June 2023.
39 Email exchange with "Jen Bones" (posting on Twitter as @gnucontrol), 1 May 2023.

Criteria for achieving impact

Adversary self-image as an enabler

A key enabler for the efficacy of humour-based approaches is the hypersensitivity of the adversary to ridicule. Multiple prominent online parodists and volunteers have noted that Russian and pro-Russian individuals and entities are intensely sensitive to mockery and show an inability to cope with being the object of derision. This is in part because they appear entirely devoid of a sense of humour about themselves; but even more significantly, Russia's self-image – perhaps to an even greater extent than that of other authoritarian regimes – carries within it a need to be taken seriously. Ridicule has been found to be effective by information warfare theorists because it attacks Russia's central myth of strength and "greatness".[40] It is particularly effective in undermining the intended effect of Russia's nuclear rhetoric – as with widely distributed memes based on animated series, showing classrooms where Bart Simpson is encouraged to "say the line" or Vladimir Putin is told by a schoolteacher "this is the third time you've brought a nuclear weapon to show and tell this week". This trend has gone beyond social media, with the Australian state broadcaster mocking Putin as a "nuclear pervert" craving ever more outlandish ways of satisfying his nuclear urges.[41]

In parallel, Russia's agents of influence, useful idiots and trolls are also hyper-sensitive to mockery, often because of the same personality defects that lead them to serve authoritarian regimes against their own countries in the first place.[42] This theory of Russian individuals' and entities' sensitivity to ridicule is borne out by consistent practical observations by online volunteers. Darth Putin notes that, "They take themselves desperately seriously and they want us to as well. What they can't stand is being laughed at."[43] NAFO founder Kamil Dyszewski notes that this builds on another key vulnerability of Russian information campaigns, namely their predictability:

> *Once you nail down their habits and patterns, you can just sweep them off their legs. But what always works is ridicule. They all see themselves as incredibly serious and important, hence why they struggle with being made fun of. And this doesn't really work against us because we make enough fun of ourselves.*[44]

The lack of self-awareness of many official Russian social media accounts does lend itself to ridicule, like the claim by the Russian Embassy in London that the "Royal Navy abuses UK's geographic position to intimidate and spy on

40 David R. Shedd and Ivana Stradner, "Waging Psychological War Against Russia", *Politico*, 9 July 2022, https://www.politico.com/news/magazine/2022/09/07/waging-psychological-war-against-russia-00054995.

41 Riley Stuart, "What 'nuclear pervert' Vladimir Putin's latest missile threats could mean for humanity", Australian Broadcasting Corporation, 6 October 2023, https://www.abc.net.au/news/2023–10–07/what-vladimir-putin-missile-threats-could-mean-for-humanity/102946826.

42 See discussion in "Radicalized: Truth Survives" podcast, Episode 55, 10 January 2023, https://www.youtube.com/watch?v=EE0j87ElnMU.

43 Author interview with @DarthPutinKGB, 13 August 2021.

44 "Founder of NAFO Reveals Identity, Discusses Raison D'être", *Kyiv Post*, 14 November 2022, https://www.kyivpost.com/post/204.

Russian ships in lawful transit", prompting Twitter users including former Estonian president Toomas Ilves to wonder if they were asking the UK to move.[45] Jen Bones also sees this as a key weakness for Russia.

> *As far as I can see, they and their cheerleaders are uniquely humourless. They take themselves so seriously, and this magnifies the effectiveness of making fun of them. But at the same time, the Russian army is a literal clown show. They video their own friendly fire incidents and post the footage online. They loot low-value hard-to-carry white goods and then abandon them in fields. They film propaganda that is so bad we can't even tell if it's real or parody. Just when you think they are getting the hang of being a military superpower, they're like 'what if we steal a raccoon?'*[46] *It's just extremely hard to resist making jokes about all this.*[47]

Knowing one's audience

Cultural sensitivity is an essential component of successful humour. This has to include recognition that not all senses of humour are equally robust, and target audiences will vary widely in their responses to it across different national contexts, even with the benefit of a notionally shared language. For example, the British Embassy in Washington triggered widespread outrage in the US in 2014 with a light-hearted tweet "Commemorating the 200th anniversary of burning the White House" – demonstrating clearly how humour is not a universal language and attempts at it need to consider the specific cultural matrix of both intended and unintended audiences.[48]

Throughout Russia's war on Ukraine, Ukraine has been highly successful in creating and leveraging messages of heroic defence – aided, of course, by the fact that there is no shortage of genuine material to work with.[49] The international resonance of the death in a flying accident in late August 2023 of pilot Andriy Pilshchykov, callsign "Juice", highlighted the manner in which this material could be leveraged for advantage beyond the information space: Pilshchykov's reputation and communication skills were credited with being instrumental in Ukraine's campaign to be supplied with F-16 combat aircraft.[50]

45 Tweet by @RussianEmbassy, 6 December 2020, https://twitter.com/RussianEmbassy/status/1335557456307556352.

46 "Russian soldiers steal racoon as they flee Kherson, Ukraine officials claim", YouTube, 14 November 2022, https://www.youtube.com/watch?v=tHntEfzBp9o.

47 Email exchange with "Jen Bones" (posting on Twitter as @gnucontrol), 1 May 2023.

48 Ed Pilkington, "British embassy sparks anger for tweet celebrating 1814 White House burning", *The Guardian*, 25 August 2014, https://www.theguardian.com/world/2014/aug/25/british-embassy-tweet-white-house-burning-1814.

49 Sofia Romansky, Lotje Boswinkel and Michel Rademaker, "The parallel front: An analysis of the military use of information in the first seven months of the war in Ukraine", The Hague Centre for Strategic Studies, October 2022, https://hcss.nl/wp-content/uploads/2022/10/The-Parallel-Front-HCSS-2022.pdf.

50 Andy Gregory, "Who is 'Juice'? The 'mega talent' Ukrainian pilot killed in mid-air plane crash", *The Independent*, 27 August 2023, https://www.independent.co.uk/news/world/europe/juice-pilot-killed-ukraine-war-who-b2400243.html.

Humour has also been a key tool for Ukraine to achieve resonance for its other messages. The ability of Ukrainian government agencies, especially the Ministry of Defence, to achieve virality and engagement through humour has also drawn widespread admiration,[51] and Ukraine appears to have comprehensively overcome the "bureaucratic virality paradox" where government communications tend by default to be too stilted, clumsy, or simply boring to be widely shared.

Ukraine has utilised this effect for direct messaging to its backers. In January 2023, at the height of the discussion over whether Western powers would supply Ukraine with main battle tanks, Ukraine's Ministry of Defence gently poked fun at the tone of the conversation, with a tweet that read "Western countries are so worried about sending tanks to Ukraine, they're arguing about what is and isn't a 'tank'. We offer our humble suggestion" – followed by a spoof commercial for the M1A2 Abrams tank rebranded as a "recreational utility vehicle" in order not to offend Russia.[52]

The success of initiatives like this presents an obvious lesson to other government communications entities around the world, and some, as exemplified throughout this report, have risen to the challenge. Others, however, even in the third decade of the 21st century, are struggling to adapt to the nature of the online information environment.

Perhaps unexpectedly, the EU bodies tasked with the lead role in countering disinformation have largely not succeeded in overcoming the paradox, and their attempts at humour remain at times ponderous and heavy-handed.[53] An unfortunate secondary effect of being unfunny is to fuel criticism of mirroring the style and manner of the propaganda outlets the EU bodies seek to document and discredit. That squanders any claim to distinctiveness, and by extension to the moral high ground.[54] It may be that this could be overcome by a more robust mechanism for impact evaluation and feedback. Responding to public criticism, a staff member from EUvsDisinfo commented that "this is a bit of a surprise for me".

> *The feedback from our partners in official circles (like NATO colleagues) has not suggested that... We push out different kinds of pieces, some more polemic, some more analytical, not forgetting humour every now and then. Think tank and academic community and official EU institutional communications push out more 'dry' content than us. We have more latitude, but we also get censored a bit internally.*[55]

The example of successful engagement by Ukrainian government voices, enabled by cultural understanding of the target audiences, also

51 Mehul Srivastava, Christopher Miller and Roman Olearchyk, "'Trolling helps show the king has no clothes': how Ukraine's army conquered Twitter", *Financial Times*, 14 October 2022, https://www.ft.com/content/b07224e1–414c-4fbd-8e2f-cfda052f7bb2.

52 Max Hauptman, "Ukraine riffs on classic Chevy commercial to ask for M1 Abrams tanks", 13 January 2023, https://taskandpurpose.com/news/ukraine-russia-abrams-tank-video/.

53 See e.g., "Kremlin Bovine Droppings Bingo", tweet by @EUvsDisinfo on 21 February 2023, https://twitter.com/EUvsDisinfo/status/1627947159675887616.

54 See e.g., critique of @EuvsDisinfo's 'tone' and 'hyperbole' in tweet by @MarkGaleotti, 7 July 2022, https://twitter.com/MarkGaleotti/status/1545006232284585985.

55 Interview with anonymous EUvsDisinfo staff member, July 2022.

contrasts with repeated failure by their Russian counterparts.[56] Russia's own ham-fisted attempts at humour have routinely fallen flat, achieving penetration only among Russia's existing supporters rather than reaching vitality through their own merits.[57] This does often appear to follow a broader pattern of Russian failure to consider or understand the audiences for its messaging, in contexts not limited to attempts at humour. This has been demonstrated in places so apparently familiar as neighbouring Belarus. In the months following the fraudulent re-election of Aleksandr Lukashenko in August 2020, Belarusian journalists resigned en masse. To replace them, propaganda teams were drafted in from Russia. A series of self-inflicted injuries in Belarusian government information campaigns included confusing Belarusian with Ukrainian, and baffling audiences with "homophobic rhetoric tying EU integration to same-sex marriages", although anti-Lukashenko protesters had not even been calling for closer ties with the EU.[58] Even at home, Russia can spectacularly misjudge how its messages will be received, especially among a young online audience. Moscow correspondent for the *Independent* newspaper Oliver Carroll described how the "Russian Foreign Ministry opened TikTok account to try and win back da kids – [but] judging by sarcastic, pro-Navalny comments, they weren't too impressed".[59]

This lack of comprehension of audiences – and apparently institutionalised lack of a sense of humour – means that when Russian government voices are apparently given free rein to vent their creativity, the results can be bizarre. For a number of years, the Twitter account of Russia's Embassy in London was notorious for extraordinary output, some of which appeared to attempt irony or other forms of humour, but which overall created an impression of being deranged rather than amusing.[60] These examples, together with that of the British Embassy in Washington above, demonstrate that while creative licence is essential for government agencies to deliver engaging content, it needs to be tempered with a common sense check, and in particular verification based on culturally-informed expertise that the output is appropriate to its aims rather than counter-productive.

Responding to Russian actions with humour can present a particularly hazardous grey area given the appalling human suffering caused by those actions; such humour will inevitably at times be deeply black, with potentially even less cross-cultural acceptance. Attempts at mockery of Russia's actions before February 2022 were criticised on the grounds that "trivializing a war criminal's actions and the harm which befalls his victims has no critical or cathartic value. It's sophomoric and insensitive at best".[61]

56 See Keir Giles, "Russian cyber and information warfare in practice: Lessons observed from the war on Ukraine", Chatham House, forthcoming.

57 Keir Giles, "Why Russia keeps laughing at the world", CNN Opinion, 6 April 2018, https://edition.cnn.com/2018/04/06/opinions/why-russia-laugh-world-giles-opinion-intl/index.html.

58 Tatyana Fedosyuk, "The comic tale of an epic propaganda fail that confirmed Russian intervention in Belarus", Great Power, 2 September 2020, https://www.greatpower.us/p/the-comic-tale-of-an-epic-propaganda.

59 Oliver Carroll (@olliecarroll), Twitter, 7 February 2021, https://twitter.com/olliecarroll/status/1358326004352188416.

60 Alec Luhn, "Russian embassy's Twitter account vents barbs against west", *The Guardian*, 11 April 2017, https://www.theguardian.com/world/2017/apr/11/russian-embassys-twitter-account-vents-barbs-against-west.

61 Political consultant Ariana Gic, by email, August 2021.

But this criticism, while still present, has become less visible with the enormous increase in the numbers engaging in this mockery following the transition to full-scale war. Some social media users engaging in the mockery are sufficiently distant from the war, both mentally and physically, to be able to laugh at Russia's actions as an abstraction. Others fully recognise the contradiction between the horror and their response – but are encouraged by the approach of Ukrainian private and government sources themselves. Jen Bones explains:

> *Why are we making jokes about this war? It still feels weird to me sometimes, and some people are quick to take objection to it. There are many reasons I think it's okay. One is that Ukrainians themselves seem to have a superbly bleak and irreverent sense of humour even in the face of an existential threat. Zelenskyy's own background as a comic actor probably helps here. So long as Ukrainians are laughing, we should laugh with them.*[62]

Accepting risk

In addition to being culturally attuned to the audience, the second common factor in successful attempts by government communicators to embrace humour is willingness to accept risk. The examples in this paper of overcoming the bureaucratic virality paradox all depend on the willingness of media departments to empower the managers of their social media accounts to attempt humour without a guarantee of success, as opposed to being constrained to staid and dry output by a fear of failure.

Official efforts can not only be mired in risk-averse bureaucracy, which puts a brake on reaction speeds, but more crucially, constrained by legislation and regulations that impose a uniquely one-sided handicap in information confrontation with unscrupulous adversaries.[63] This appears to affect NATO allies unevenly, with some more concerned about centralised control than others: in 2020, a component commander from a major NATO ally observed that on operations in Syria, his government regulations meant "media communications requiring approval at General level, while US counterparts placed this level of responsibility with a staff sergeant".[64] By contrast, non-governmental, non-structured entities "swim in online waters that governments would struggle to enter".[65] The overall effect is that viral humour-based campaigns have as a rule had little in common with the methodology for government operations recommended by NATO's Strategic Communications Centre of Excellence.[66]

This approach to risk is a key reason why private individuals and civil society have, in many cases, been more adept at maximising the use

62 Email exchange with "Jen Bones" (posting on Twitter as @gnucontrol), 1 May 2023.

63 Lee Fang, "Twitter aided the Pentagon in its covert online propaganda campaign", *The Intercept*, 20 December 2022, https://theintercept.com/2022/12/20/twitter-dod-us-military-accounts/.

64 Speaking at an event held under the Chatham House Rule, June 2020.

65 Mark Scott, "The shit-posting, Twitter-trolling, dog-deploying social media army taking on Putin one meme at a time", *Politico*, 31 August 2022, https://www.politico.eu/article/nafo-doge-shiba-russia-putin-ukraine-twitter-trolling-social-media-meme/.

66 Žaneta Ozoliņa et al., *StratCom laughs: in search of an analytical framework*, NATO Strategic Communications Centre of Excellence, 2017, https://stratcomcoe.org/publications/stratcom-laughs-in-search-of-an-analytical-framework/201.

of humour to get their messages across. But government communicators can, if permitted, achieve this too, as the following section explains.

Agility and innovation

Achieving success with humour-based engagement appears easiest when it leverages the agility of organic, unregulated organisations of private individuals. In many cases, these appear far more capable than government-directed entities both of rapid responses to incidents and of evolution in the face of changing circumstances and environments. In particular, swift and effective interventions by individuals or groups stand in stark contrast to "responses and refutations to Russian misinformation from official sources [that] can often be colorless, pedantic, and repetitive".[67] In this respect, developments in crowdsourcing humour show multiple parallels with the evolution of information warfare in and around Ukraine overall[68] – as well as the widespread engagement of volunteer organisations in supporting the war effort in other domains too.[69] Not only are private citizens involved and delivering effect to an extent that would have been difficult to conceive in previous conflicts, but their rapid advances in sophistication and methodology have far outstripped Russia's range of information warfare techniques, which now appear degraded and obsolete.[70]

The success of individuals in countering disinformation with humour results in part from the fact that resource constraints that might previously have limited information operations to well-funded organisations have largely fallen away. This applies both to access to audiences, and to the creation of material. The democratisation of production tools has meant a corresponding increase in the range of individuals capable of producing high-quality content without the need for specialist facilities or training.[71] Memes can still on occasion be relatively crude, but the strong trend is towards sophisticated still image memes and videos with high production standards, now well within the capabilities of private citizens operating independently and pro bono. In addition, private individuals can operate at a speed and in a manner that government agencies largely cannot. This also applies to private corporations: as described by Gen. (Rtd) John Allen, describing information support for operations in Ukraine, "The private sector moves at the necessary speed because war is a competition in time. Government

67 Connor Forberg, "NAFO: A Masterclass on Upending the Russian Propaganda Machine in the Modern Era", course paper, Georgetown University, 13 March 2023.

68 See Keir Giles, "Russian cyber and information warfare in practice: Lessons observed from the war on Ukraine", Chatham House, forthcoming.

69 Peter Guest, "Ukraine War: How to Win With Trucks, Trolls, and Tourniquets", *Wired*, 6 July 2023, https://www.wired.com/story/ukraine-war-trucks-trolls-tourniquets/.

70 Thomas Brewster, "Russian Information Warfare Used To Be Sophisticated. Meta Says It Now Looks Like Basic Spam", *Forbes*, 23 February 2023, https://www.forbes.com/sites/thomasbrewster/2023/02/23/meta-victories-in-russia-information-war/.

71 "Jen Bones", posting on Twitter as @gnucontrol, is a clear example of how high-quality content can be produced by someone self-educated in graphic and media design and production. See https://twitter.com/gnucontrol.

moves at the speed of government, which is too slow."[72]

However, these rules are not universal. Well before 2022, outliers in government communications were demonstrating that responsive engagement is possible, and earning a reputation for influence at times out of proportion to their notional importance or seniority. One classic example is the Twitter account of the Canadian military representation in the United States.[73] Now apparently dormant, this account had the key advantage of being run by a single individual who was empowered both to take risk and to have a distinctive voice. This, plus the agility and cultural and situational awareness of the single account manager, were credited with the result that "anecdotal evidence suggests [the account] has had more success reaching certain segments of the U.S. population than an expensive public-relations campaign conducted by the Canadian Embassy".[74] Similarly, the US National Parks Service is a highly successful example of a government entity developing an online "voice" centred around irreverent humour and appalling puns, and leveraging this to impart useful and important information to a broader audience, including groups that could otherwise potentially find it dry and forbidding.[75]

Reproduced on a much larger scale, these same principles have been instrumental in the relative success of Ukrainian strategic communications campaigns in the context of the country's defence against Russia, including those using humour for effect. This success was built on the foundations of a substantive overhaul of Ukrainian institutions after 2014, with assistance in both education and resources from foreign partners, including the UK.[76] This overhaul has been credited with creating a relatively youthful cohort of strategic communications professionals, who as well as being granted the essential leeway to make mistakes, may have a greater inherent aptitude for risk and experimentation.[77] But since 2022, their efforts have been greatly augmented by assistance from the public relations and advertising industries, volunteered as part of the war effort. This, facilitated by the predominance of skilled communicators in senior positions in the Ukrainian government, has enabled agile, proactive and engaging strategic communications making full use of modern media tools, in stark contrast to Russian – and on occasion Western – efforts.[78]

Few Western government agencies will be in a position to replicate the conditions that Ukraine has made use of. But all should be observing the methods on display and drawing appropriate conclusions: in particular, the style

72 Gen. (Rtd) John Allen, speaking at the Lennart Meri Conference, Tallinn, Estonia, on 14 May 2023.
73 See https://twitter.com/CAFinUS.
74 Lee Berthiaume, "Trust, awareness key to success of Canadian Forces' most famous tweeter", CBC News, 16 October 2020, https://www.cbc.ca/news/canada/ottawa/canada-military-forces-twitter-1.5764786.
75 See https://twitter.com/natlparkservice.
76 FOI, "War of words – how Ukraine uses strategic communication to beat Russia on the information front", Swedish Defence Research Agency, 21 April 2023, https://www.foi.se/en/foi/news-and-pressroom/news/2023–04–21-war-of-words---how-ukraine-uses-strategic-communication-to-beat-russia-on-the-information-front.html.
77 Author interviews, April–July 2023.
78 FOI, "War of words".

of engagement, of which humour is an integral component, that achieves virality.[79]

Accessibility of communities of interest

Another key advantage of communities of interest built around humour is their accessibility. With no membership criteria or joining mechanism for the movement, and an inbuilt incentive to acquire greater mass through numbers, leading parodists broadly welcome support and emulation (although not direct plagiarism).

Versatile catchphrases and slogans also ease engagement, as with the campaign of mockery of Russia's air defence capabilities with the tag "what air defense doing?" – instantly adoptable by third parties and adaptable to other instances of Russian deficiency and failure.[80] Thus, they achieve greater virality among broader audiences than previous catchphrases that have entered the common language of Russia-watchers because they sum up key aspects of the Russia problem, such as Darth Putin's "Never believe anything until the Kremlin denies it", and "I remain a master strategist".

Here too, it is Russia's own actions that have spurred the democratisation of resistance, with February 2022 serving as a catalyst for many formerly uninterested or disinterested individuals to take whatever action they could. Darth Putin's consistent display of insights into Russian behaviour and policy has led to their being credited with well-honed skills in foreign policy analysis. But they counter that this is accessible to anybody: "It doesn't take a lot of time thinking about it. You don't have to be a deep Russia expert to notice and understand these things… it's all blatantly obvious; you can get that with the most basic understanding of the world today."[81] This applies in full to many individuals who have become active in this space more recently. Jen Bones adds:

> *I don't have any personal connection to Ukraine. But I care about the world I live in, and I had followed the situation in Ukraine on and off since the 2014 invasion in as much as I read or listened to mainstream coverage when it presented itself. I joined Twitter not long after the 2022 invasion, just to try and get closer to the breaking news. I never expected to post anything, much less get involved in any sort of activism. But it made me so angry to see all this distorted information being thrown around, and I gravitated towards the people who were making an effort to challenge and debunk it.*[82]

Nevertheless, engagement in activities even as apparently innocuous as mockery of Russian propaganda, and even at a great distance from Ukraine, is not without a degree of risk. Those who oppose Russia, China or other hostile regimes publicly face consequences, delivered

79 Classic examples of blending humour with cultural awareness are evident in the series of videos released by the Ukrainian Ministry of Defence thanking donor countries for arms supplies, set to fitting musical accompaniments, of which the version for the UK was endorsed by the British Ambassador to Kyiv as "the best video @DefenceU ever made". See tweet by Dame Melinda Simmons, 24 August 2023, https://twitter.com/MelSimmonsFCDO/status/1694694898841141757.

80 Connor Forberg, "NAFO: A Masterclass on Upending the Russian Propaganda Machine in the Modern Era", course paper, Georgetown University, 13 March 2023.

81 Author interview with @DarthPutinKGB, 13 August 2021.

82 Email exchange with "Jen Bones" (posting on Twitter as @gnucontrol), 1 May 2023.

by the hostile regime itself or by its agents and sympathisers.[83] These reprisals can range in severity from common abuse and threats online,[84] through doxxing, to direct and dangerous interventions in real life.

The exposure to risk is not on the same level as civilians in direct proximity to the fighting who engage in information collection and processing, blurring the lines between combatants and non-combatants.[85] But repeated incidents show the consequences can still be personally devastating.[86] Individuals taking on authoritarian propaganda structures conduct their own risk assessment, sometimes with an incomplete picture of the potential hazards. Those who recognise the dangers take precautions to protect their identity accordingly. According to Jen Bones:

> *Russia and its supporters have shown themselves to be some of the worst people in the world, and here I am deliberately trying to wind them up and discredit them. What could possibly go wrong? One healthy dose of paranoia later and I'm doing everything with pseudonymous burner accounts and avoiding posting things that would make me easily identifiable... I've seen people get doxxed and harassed and it doesn't look fun. I want my family to be safe.*[87]

Nevertheless, the emergence of a mass movement which opponents of Russia can call on for support has to some extent mitigated the previous chilling effect of harassment and intimidation. Private citizens have been emboldened to speak out by the realisation that they are not – or no longer – alone. As described by disinformation researcher and activist Joohn Choe, "Usually, trolls are *less famous* than the people they are trolling... Together, though, we are several orders of magnitude more famous than what we troll."[88] In an open forum, this enables them to compete on more equal terms with organised groups that might otherwise be better able to exploit both their mass and the rules of the platform. Rather than being specific to Russia, this principle is universally applicable. As Choe points out, it applies in equal measure to confrontations with aggressively-inclined supporters of former US President Donald Trump.[89]

83 See e.g., tweet by Kit Klarenberg (@KitKlarenberg), 4 June 2023, https://twitter.com/KitKlarenberg/status/1665409400436150272.
84 See e.g., tweet by Peter Jukes (@peterjukes) describing harassment, 3 June 2023, https://twitter.com/peterjukes/status/1665110823331569665.
85 See also Keir Giles, "Russian cyber and information warfare in practice: Lessons observed from the war on Ukraine", Chatham House, forthcoming.
86 As described in detail in Keir Giles, "Russia's War on Everybody", Bloomsbury, November 2022.
87 Email exchange with "Jen Bones" (posting on Twitter as @gnucontrol), 1 May 2023.
88 Joohn Choe (@JoohnChoe), Twitter, 23 February 2023, https://twitter.com/joohnchoe/status/1628875590278447104.
89 As described by @JoohnChoe, 7 March 2023, https://twitter.com/JoohnChoe/status/1633212588191547392.

Impact and side effects

Darth Putin's assessment, based on long-term observation of the behaviour of propaganda accounts in response to parody, suggests that humour does at least impose costs and friction on their operations. But in addition, effective humour campaigns have been shown to directly alter the behaviour of Russian propagandists.[90] This is because a humour-based community of interest provides a broad and supportive audience that will be receptive to evidence of the personal hypocrisy of leading propagandists – evoking a reaction that plainly diverts them from their core mission.[91]

Similar impacts can be observed through counters to Russian information activities that are not necessarily humorous, but that achieve effect simply by not according Russian diplomats the gravitas they crave, as in March 2022 when the Canadian UN delegation posted on Twitter a letter by their Russian counterparts that had been "edited" in red to bring it more in line with reality.[92] In this case, the primary measure of effect, beyond widespread dissemination of the "edited" version, was Russian diplomats snapping back with crude insults, indicating that the message had indeed hit home.

Other side effects of targeted humour come through its role as the glue holding together a community of interest. According to a study on NAFO compiled in early 2023, this allows a mass response to disinformation: "Russian propaganda is screenshotted and meme'd; other older memes pile on as well and are reused and refined – with the funniest and wittiest efforts floating to the top of the feed."[93] A community of interest is a step towards granting counter-disinformation measures the virality that is so much more easily achieved by disinformation itself.

Crowdsourcing, and the pooling and sharing of observations and situational awareness that comes with it, also leads to greater shared situational awareness of the information environment; in effect, a practical implementation by civilians on social media of the "every soldier a sensor" ambition commonly expressed by Western militaries in the first decade of this century.[94] This awareness in turn leads to a number of secondary and tertiary effects, all of which are disruptive for disinformation and propaganda efforts. It allows the countering of disinformation that otherwise would remain unchallenged in the online space because it was undetected – particularly when that disinformation is laundered through content farms and individual propagators.[95] In doing so, it also enables the

90 See e.g., tweet by @RasReload, 16 December 2022, https://twitter.com/RasReload/status/1603841019820351488.

91 Sinéad Baker, "Pro-Kremlin propagandist loses it on live TV after his son's patriotism was called into question for not fighting in Ukraine", *Business Insider*, 15 February 2023, https://www.businessinsider.com/russia-propagandist-erupts-after-son-accused-dodging-ukraine-fight-2023-2?r=US&IR=T.

92 Jake Epstein, "Canada trolls Russian diplomat by marking up his letter denouncing a UN resolution condemning attacks on Ukrainian civilians", *Insider*, 17 March 2022, https://www.businessinsider.com/canada-trolls-russian-united-nations-diplomat-edits-letter-resolution-ukraine-2022-3.

93 Connor Forberg, "NAFO: A Masterclass on Upending the Russian Propaganda Machine in the Modern Era", course paper, Georgetown University, 13 March 2023.

94 Stew Magnuson, "Army wants to make 'every soldier a sensor'", *National Defense*, 1 May 2007, https://www.nationaldefensemagazine.org/articles/2007/5/1/2007may-army-wants-to-make-every-soldier-a-sensor.

95 Michael Weiss, "Inside the Kremlin's disinformation war against Ukraine", Yahoo! News, 16 February 2023, https://news.yahoo.com/inside-the-kremlins-disinformation-war-against-ukraine-183128480.html.

detection and investigation of networks and agents, both public and private, and potentially their subsequent neutralisation through public exposure.

Self-starting communities of interest built around humour have also been credited with shoring up support for Ukraine since "the humour and creativity help stave off war fatigue".[96] Memes exploiting the comic potential of the technological gap between Russian and Western armaments serve as a reminder that the war is finite, as well as a short-term morale booster.[97] A classic example came in September 2023, when satellite imagery revealing that Russia was protecting its strategic bombers against drone attack by putting car tyres on them spawned a vast number of mocking memes.[98]

Interaction between online humourists is mutually reinforcing, and also sees them receiving affirmation from the objects of their support. Jen Bones notes:

> *I have a fair few Ukrainian followers and a like/RT/comment from one of them is super motivational... if I were in their place, I do think it would be at least a little bit reassuring to see that people around the world were actively giving up their spare time to try and make a difference.*[99]

This aligns with explicit objectives laid out by Ukrainian government officials, who name building the psychological resilience of the population among one of the objectives of using humour, alongside imposing costs and friction on the aggressor and disseminating information among audiences that might not be reached by content that is dry or academic.[100]

The overall effect of a community built around humour has been to turn the tables on social media platforms. The agents of influence and other servants of authoritarian regimes, who for so long held the advantage, are turned into the targets rather than the deliverers of mockery and abuse.[101] This causes both Russian officials and their extended network of influencers, enablers and trolls to realise that if they choose to serve a criminal regime, they expose themselves to mass ridicule and mockery.

A key point is that in the case of support for Ukraine, large sections of that community already existed, but were isolated and weak in the face of the much more cohesive and organised opposition (essentially, the constellation

96 Mehul Srivastava, Christopher Miller and Roman Olearchyk, "'Trolling helps show the king has no clothes': how Ukraine's army conquered Twitter", *Financial Times*, 14 October 2022, https://www.ft.com/content/b07224e1–414c-4fbd-8e2f-cfda052f7bb2.

97 See e.g., Instagram post by Saint Javelin (@saintjavelin) on 25 January 2023, https://www.instagram.com/p/Cn1q2yRKpBc/.

98 Miriam Burrell, "Russia covers nuclear bombers with tyres in apparent bid to shield them from Ukraine drones", *Evening Standard*, 6 September 2023, https://www.standard.co.uk/news/world/russia-ukraine-bomber-tyres-drones-tu95-engels-airbase-b1105326.html.
See also https://twitter.com/Rickslimfit85/status/1699115220709314920.

99 Email exchange with "Jen Bones" (posting on Twitter as @gnucontrol), 1 May 2023.

100 Author interviews, April–July 2023.

101 See e.g., tweet by @LeSuderer, 24 January 2023, https://twitter.com/lesuderer/status/1617771340022546432.

of Russian and pro-Russian trolls, influencers, and useful idiots). In other words, rather than a movement coalescing out of nowhere, shared humour to some extent gave a latent community an identity, and unity and the power that went with it. That community can now rally for mutual support on demand.[102]

102 See e.g., tweet by @NelltheWeaver, 23 May 2023, https://twitter.com/nelltheweaver/status/1661103057084510210.

Case study: NAFO

The North Atlantic Fella Organization (NAFO) provides a clear example of an online community of interest organically responding to disinformation from governments and counterfactual communities with methods including humour. Significantly, this was not the group's original intent: NAFO grew from a fundraising initiative in mid-May 2022 by founder Kamil Dyszewski, designing cartoon dogs to thank individuals who donated to the Georgian Legion fighting in Ukraine. But the establishment of a mutually-aware group with largely common objectives and mindset provided the ideal structure and format for a secondary result of countering disinformation, largely through the application of humour and mockery.

NAFO's structure and format is loose and open. As a decentralised network, NAFO took a long time to acquire a website, a form of administration and coordination, a means of contacting that administration, or even "official" accounts on social media.[103] One description of NAFO explains that "While there is no central governing authority, responses are coordinated generally across platforms including Twitter, Telegram and Discord".[104] But this is not universal, and coordination (just like participation) is entirely voluntary. Jen Bones, for example, works as a lone operator not coordinating efforts with other content creators, and yet created some of the earliest viral content that was highly influential in shaping the NAFO identity and ethos.

This breadth of community and open participation by a wide range of contributors with varied locations, backgrounds and skills gives NAFO key advantages in crowdsourcing information activities. For instance, NAFO includes individuals engaged in combat operations in Ukraine. They are able to provide the base material for memes in the form of first-hand images and videos of Russian failures or Ukrainian determination, adding to the responsiveness and impact of campaigns overall.[105]

The transformative effect of the mass that NAFO provides in online interactions can be seen from the case of Russia's Permanent Representative to International Organisations in Vienna, Mikhail Ulyanov. Previously, efforts to counter Ulyanov's offensive propagation of Russian state narratives took the form of relatively ineffectual individual engagements.[106] But in June 2022 a mass response by the still-nascent NAFO to Ulyanov's attempts to justify Russian actions led to his slightly off-key English[107] achieving virality and memification, and spawning a range of fund-raising merchandise.[108] Ambassador Ulyanov was subsequently described as having been "chased offline" by these efforts.[109] However, at the time of writing, this does not appear to be correct, as Ulyanov has repeatedly shown himself to be sufficiently lacking in self-awareness to be impervious to near-universal derision, blunting its effect. Nevertheless, the overall impact has been to discredit Ulyanov – and by extension the messages he delivers – among a wide sector of online audiences, who now perceive him as a figure of fun rather than a serious diplomat.

Scholar of information warfare P. W. Singer notes that through aligning efforts to dismiss and discredit Russian propaganda, NAFO has denied Russia the ability "to run wild across the info-war landscape the way they used to".[110] In addition to extensive serious news coverage,[111] this impact has been recognised and endorsed at senior government level in the frontline states; not only by the Ukrainian Defence Ministry,[112] but also for example by President of the Republic of Lithuania Gitanas Nausėda, who tweeted a welcome to NAFO "Fellas" visiting Lithuania at the time of the NATO Summit in July 2023.[113]

103 Now available at https://twitter.com/Official_NAFO.

104 Callum Harvey and Jack Goldsmith, "Building a meme war machine: A comparative analysis of memetic insurgencies in cyberspace", forthcoming.

105 Connor Forberg, "NAFO: A Masterclass on Upending the Russian Propaganda Machine in the Modern Era", course paper, Georgetown University, 13 March 2023.

106 See e.g., tweet by Keir Giles (@KeirGiles), 17 April 2020, https://twitter.com/KeirGiles/status/1251168295169884160?s=20.

107 Tweet by Mikhail Ulyanov (@Amb_Ulyanov), 19 June 2022, https://twitter.com/Amb_Ulyanov/status/1538562863199141889.

108 Primarily through the Saint Javelin online store, which raises funds for aid to Ukraine. See https://www.saintjavelin.com/products/fellas-you-pronounced-this-nonsense-sticker.

109 Matthew Gault, "Shitposting Shiba Inu Accounts Chased a Russian Diplomat Offline", Motherboard, 12 July 2022, https://www.vice.com/en/article/y3pd5y/shitposting-shiba-inu-accounts-chased-a-russian-diplomat-offline.

110 "A virtual army of impish cartoon pooches is waging war on Russia", *The Economist*, 31 August 2020, https://www.economist.com/europe/2022/08/31/a-virtual-army-of-impish-cartoon-pooches-is-waging-war-on-russia.

111 Adam Taylor, "With NAFO, Ukraine turns the trolls on Russia", *The Washington Post*, 1 September 2022, https://www.washingtonpost.com/world/2022/09/01/nafo-ukraine-russia/.

112 See e.g., tweet by Defense of Ukraine (@DefenceU), 28 August 2022, https://twitter.com/DefenceU/status/1563851548643426304.

113 Gitanas Nauseda (@GitanasNauseda), Twitter, 20 May 2023, https://twitter.com/gitanasnauseda/status/1659878208068624384.

Conclusions

This Hybrid CoE Working Paper has sought to present examples of best practice in adopting humour as a counter-disinformation tool. But in doing so it has demonstrated the difficulty of objectively measuring best practice when the only metrics available are highly subjective. This underscores some of the key considerations that emerge from examining the cases above:

- Bureaucratic processes that are risk-averse are innately unsuited to being funny. This is because success in achieving virality appears to depend on investing trust and empowering account managers to take risks.
- Cultural sensitivity is a key ingredient for success, and cannot be invested in too heavily.
- Robust feedback mechanisms and monitoring of effect – however subjective – are essential. This too militates against success for organisations that are culturally unable to accept the risk of failure, and when it occurs to treat it as a learning opportunity.

It follows that government departments wishing to utilise the full power of humour and ridicule for countering disinformation need to embrace a culture where innovation and experimentation are not feared, and neither are the mistakes that may result from them. Conversely, a culture of zero tolerance for errors and embarrassment is unlikely to be agile enough to overcome the bureaucratic virality paradox.

When humour is adopted and applied by private citizens for the same purpose, a different picture emerges. In a reversal of the usual pattern, the engagement of citizens makes humour one of the few tools in information warfare that is more readily available for use by liberal democracies than by authoritarian regimes. In this way, the proactive use of humour, both as a practical countermeasure against disinformation and as a means of building a network of shared interest, augments previous coordinated volunteer efforts such as the "elves" that have relied primarily on mass debunking and reporting of suspect accounts and posts.[114] But in addition, the use of memes and humour has the advantage of exploiting social media platform algorithms that otherwise favour disinformation operations.

Rather than attempting to fight the inherent challenges of an online environment that is neutral or in some cases actively favours the adversary, a crowdsourced humour approach makes use of that environment and its algorithms for advantage.[115] It is therefore ideally suited to mitigating the effects of propaganda and disinformation campaigns on shared platforms. This in itself makes humour an essential element of the counter-disinformation toolkit, simply because it reaches the parts that other countermeasures – like fact-checking or media user education – cannot.

Besides the direct effect of discrediting the sources of disinformation (and hence the disinformation itself that emanates from them), humour achieves three vital side effects: building the resilience of your own audience, imposing costs on the aggressor, and spreading the

114 Lisa Abend, "Meet the Lithuanian 'Elves' Fighting Russian Disinformation", *Time*, 6 March 2022, https://time.com/6155060/lithuania-russia-fighting-disinformation-ukraine/.

115 Joseph Menn, "Twitter removes labels from state-controlled media, helping propaganda", *The Washington Post*, 21 April 2023, https://www.washingtonpost.com/technology/2023/04/21/twitter-russia-china-state-media-propaganda/.

message to audiences that are not inclined to consume "boring" products. Here too the primary impact is on social media, since it is there that the speed of access and propagation of information maximises the advantages that humour brings. While humour cannot replicate the effects of formal monitoring, debunking and user education, neither can these achieve the effects delivered by humour. As such, humour – whether wielded by official government communications agencies, or individuals acting on their own initiative – constitutes an important additional tool for bolstering the resilience of the information environment.

Author

Having briefly tried to earn a living as a stand-up comedian, **Keir Giles** knows what it feels like when a joke falls flat. More recently, Keir has enjoyed slightly greater success as an internationally recognised expert on the Russian approach to information warfare, including the subdomains of computational propaganda and of cyber conflict. He is the author of a significant number of groundbreaking studies on Russian theory, doctrine, and structures for engaging in information and cyber confrontation. He is currently a Senior Consulting Fellow with the Russia and Eurasia Programme at Chatham House in London, and also works with the Conflict Studies Research Centre, a group of subject matter experts focused on the range of security threats posed by Russia.

Hybrid CoE

The European Centre of Excellence
for Countering Hybrid Threats