



# Prevention of Youth Radicalisation Through Self-Awareness on Cognitive Biases

## Scientific report

### Contributors

**Prof. Dr. Catherine Bouko**  
Universiteit Gent, Belgium

**Julie Carton**  
Universiteit Gent, Belgium

**Dr. Brigitte Naderer**  
Ludwig Maximilians Universität München, Germany

**Prof. Dr. Diana Rieger**  
Ludwig Maximilians Universität München, Germany

**Pieter Van Ostaeyen**  
Katholieke Universiteit Leuven, Belgium

**Pierre Voué**  
Textgain, Belgium





# **PRECOBIAS**

**#thinkaboutit**



This report was funded by the European Union's Internal Security Fund — Police. Project n.: 867186

The content of this report represents the views of the author only and is his/her sole responsibility.

The European Commission does not accept any responsibility for use that may be made of the information it contains.

# About the project

PRECObIAS is a campaign project that centres around the mental processes and cognitive biases at play when youngsters are faced with extremist or terrorist discourses on social media.

By enhancing the digital resilience and critical thinking of our target audience, we want to counter radicalisation in the long term. PRECObIAS helps youngsters to understand themselves better by revealing the underlying mental processes and cognitive biases that shape their interpretations and analyses.

To reach these objectives, PRECObIAS targets:

- vulnerable and radicalized youngsters.

HOW? Through a social media campaign with videos, an Instagram contest and a self-test to increase awareness of cognitive biases.

- social workers and teachers who are in charge of the target youngsters.  
HOW? By offering a MOOC on cognitive biases and radicalisation, as well as two toolkits with ready-to-use activities.

PRECObIAS involves partners from 6 EU countries (Belgium, Germany, Hungary, Italy, Poland, Slovakia).

PRECObIAS comprises scientific research on cognitive biases and radicalisation on social media, conducted by researchers from Universiteit Gent and LMU, in collaboration with researchers from KUL and Textgain. This report sums up our main findings.



# Precobias – report

## 1) Why does PRECOBIAS focus on radicalization and cognitive biases?

Radicalization is typically described as an increasing extremity of beliefs, feelings, and behaviours in support of conflict and, potentially, violence.<sup>1</sup> The radicalization process is based on a set of complex factors that need to coincide in order to actually produce extremist outcomes.<sup>2</sup> Yet, confrontation with extremist propaganda material or hate speech is commonly seen as potentially conducive.<sup>3</sup> Therefore, it seems important to first investigate what extremist communication looks like before we can draw conclusions about its potential effects.<sup>4</sup> Rather than focusing on what extremists post online, it is key to understand how they communicate and how we process that information, in order to understand extremists' use of psychological forces to radicalize potential sympathizers. Cognitive biases play a very practical and important role in the way we process information. They shape the way we think and behave, and they act on an unconscious level. Most people are unaware of the fact that their thinking or behavior is steered by cognitive biases. These biases are not uncommon, nor are they bugs in our cognitive functioning. They are necessary and automatic processes that allow us to process information and make sensible decisions in our day-to-day lives.<sup>5</sup> In PRECOBIAS, we hypothesize that online extremist narratives can trigger cognitive biases in users.

## 2) What was the focus of our research questions?

The project wants to examine the cognitive biases that play a role in exposure to extremist communication, and how they might push people into holding more radicalized opinions. The project targets radicalization in the political as well as in the religious sense. This is why our strategic focus lies on both non-official Islam-based extremist communication (see Question 3) and on right-wing extremists (see Question 8). We analyzed the specific patterns and narratives that are used in extremist communication and that might activate biases when people process information.

In an effect-based examination (see Question 12) we then built on these identified patterns to examine 1) how extremist communication can trigger cognitive biases and 2) how the activation of those cognitive biases might make people more vulnerable to extremist ideas.

### **3) Why does PRECOBIAS focus on non-official Islam-based extremist communication?**

Europol deals solely with official propaganda content posted by designated terrorist organizations. Content published by private IS supporters, who play a major role in extremist propaganda on social media, falls beyond the scope of official monitoring and is left for other stakeholders to address. The same goes for academic research, the majority of which covers only official propaganda (IS magazines, videos, etc.). Research on the effect of cognitive biases in radicalization is typically centred around established IS frames (e.g. ingroup versus outgroup) and rarely addresses how unofficial communication may trigger specific cognitive biases.

We argue that, in addition to official IS communication, Salafist extremists' propaganda on social media can also play a key role in Islam-based radicalization. In particular, these extremist profiles can be “dormant cells”<sup>6</sup>, who share mostly non-violent content on their public profiles and might additionally reach potential recruits or sympathizers through private messages. Since user-generated content can be uploaded by non-official sources, it is relevant to focus on this communication as well in order to grasp the potential of extremist communication to trigger cognitive biases.

### **4) What social media posts did we analyze?**

We analyzed over 3000 posts from eight Facebook and five Instagram profiles of Salafist extremists that contain text and image. These profiles are supporters of the globally active Jihadist movement, but they do not express adherence to any extremist organization. The profiles vary in the size of the community, which is either unknown (i.e., private friends lists on Facebook) or ranges from 155 to 52 000 followers. We covered the period from October 2016 to October 2019, i.e. after 2015, the turning point when social media content regulations on extremist content increased significantly.

## 5) What did we observe about topics and frames in Islam-based extremist propaganda?

The first step of PRECOBIAS's research was to gain an insight into the topics and frames used in Islam-based extremist propaganda (i.e., the content level, the 'what'). To analyze the content level, we built our research design on Ingram's publications in which he explains how the concepts of crisis, identity and solution interplay in IS propaganda.<sup>7</sup>

### 5.1. CRISIS

#### 5.1.1. The low salience of anti-outgroup posts

It is striking that most of the posts (between 76% and 100%) do not oppose ingroup and outgroup. The profiles do not appear to be fuelled by anti-outgroup frames as a key topic. In most profiles, these frames are just one of many topics and covered even less frequently than others.

#### 5.1.2. Attacks on the ummah vs. the dilution of true Islam and/or the breakdown of religious traditions

The **dilution of true Islam** and/or the **breakdown of traditions** is a key topic in some of the profiles. Interestingly, the profiles that emphasize attacks against the ummah (i.e. the Muslim community) tend not to focus on the dilution of Islam and the breakdown of the traditions, and vice versa.

**Victimhood** is a key frame in both official and unofficial propaganda. We observed several images of lethal victims (e.g. members of the prosecuted Rohingya and Uighur minorities, as well as numerous pictures of injured citizens, particularly children. Posts that frame **Muslim prisoners as victims** also fall within the category of attacks against the ummah. Their prison conditions are particularly denounced, with mentions of torture, medical negligence or detailed descriptions that create shocking visual representations in the reader's mind (e.g., a Muslim prisoner forced to eat pork and drink toilet water). The stories about these prisoners are often taken out of their context. In many cases, the post first introduces a particular prisoner and then extends to Muslim prisoners in general. One prisoner acts as a symbol for a more general ingroup vs. outgroup confrontation, rather than a particular example that needs to be discussed. This strategy of **decontextualization and positive framing** allows the authors of the posts to circumvent Facebook's terms of service. Most posts do

not praise the prisoners' actions or beliefs, which is not allowed, but nine posts out of ten do consider their arrest and detention to be illegitimate.

## 5.2. IDENTITY

Extremism can also be fuelled by an identity construct, often through the framing of certain extremist figures as inspiring, usually in an implicit way. Extremist profiles use decontextualized quotes from extremist figures that contain vague fight rhetoric. They subtly endorse those figures and put them in a positive frame (i.e. by listing the figure's intellectual qualities).

For example, mujahid Emir Khattab's call to "free ourselves from the Arab leaders" (see quote below in French) might sound rather **violent**, but the violence is **only implicit**. These quotes are transformed into decontextualized empty shells that are ready for hosting new implicit meanings according to the contexts in which they are used.



**Image 1:**  
Reference to mujahid Emir Khattab in our corpus

We observed this same lack of explicitness in posts with identity images and symbols. There was never any explicit visual content referring to IS (flag, etc.); only implicit symbols that occur regularly in extremist communication, but not exclusively. (e.g. lion imagery).

## 5.3. SOLUTION

### 5.3.1. Caliphate as Islamic utopia

Not a single post of our corpus praises the caliphate, which is forbidden on Facebook and Instagram. However, it is praised implicitly in two posts

that lament one specific Muslim's development, who used to be "pro-caliphate" and "pro-sharia."

### 5.3.2. Violence as an explicit solution

We observed **two types of posts** in which violence is explicitly advocated as a solution against outgroups. The first type contains extracts from **religious texts** that are prone to violent jihad. Example 1: "He believes that Jihad continues from the time Allah sent Muhammad to the last group that will fight the dajal. They will not suffer any harm from the misdeeds and wrongdoers." Example 2: "Hatred towards these polytheists, their criticism, takfir, disavowal, is the foundation of Islam and the greatest means of access to the Lord of the Worlds. Moreover, the life of a Muslim will have no pleasure except with the jihad against them, against their opposition and their takfir, and with approaching Allah with that, hoping for His reward." **Jihad** is not framed positively through frames of adventure or camaraderie, but only evoked via religious texts. The second type of posts explicitly advocates certain **terrorist acts**.

Apart from these explicit cases that violate Facebook's terms of service, many other cases manage to **circumvent** them, through **five discourse strategies**.

The first strategy is a pseudo-factual **explanation and justification** of the emergence of terrorist groups (e.g., the thousands of deaths in Syria, Iraq or Afghanistan, the tortures in American or secret prisons). Here, the post does not support terrorist acts, but does not condemn them either. They are presented as defensive violence, triggered by faults of the West.

The second strategy is the use of **vague fight rhetoric**, which leaves the door open for violent interpretations.

The third strategy is the use of extracts with only limited signs of endorsement. The author of the post reveals his approval in a very subtle way. Calls to divine violence are the fourth strategy: e.g., "May Allah hasten his punishment to all those tyrants who humiliate, torture, imprison and oppress Muslims on earth." These excuses for divine violence do not violate Facebook's terms of service, because as they remain abstract, whereas Facebook removes content only when "there is a genuine risk of physical harm or direct threats to public safety" (Terms of service). Advocating hypothetical future violence is another strategy that does not constitute a threat to security.

Interestingly, the salience of the two frames that fuel extremist content the most (i.e. inspiring extremist figures and violence as solution) is very low: in nearly all the profiles, they take up around 1% of the posts, and never more than 4%.



## 6) Are there differences between Facebook and Instagram for Islam-based extremist propaganda?

Both within and across Facebook and Instagram, the profiles differ greatly: they oppose different outgroups and combine the crisis-identity-solution frames in different ways. Nevertheless, we observed four trends, which were relevant for both Facebook and Instagram:

- Profiles without or with hardly any ingroup vs. outgroup frames. They focus on non-extremist religious propaganda.
- Predominantly political profiles: the majority of posts is about attacks against the ummah; institutions are the main outgroup and the main visual strategy lies in the use of pictures, to denounce these attacks through ideological realism.
- Predominantly religious profiles: the majority of posts is related to religious issues. Disbelievers and moderate Muslims are more often addressed as outgroups than in the type above.
- Profiles without thematic prevalence.

Visual content was not very different across Facebook and Instagram. What is striking is that apologies for violence were very often expressed in long texts (and thus could easily remain undetected) rather than through visual content.

## 7) What did we observe about extremists' communication styles/discourse patterns?

In a second phase, we identified communication styles (i.e., the discourse level, the 'how') of Islam-based extremist propaganda. These are our findings:

### 7.1. Adhesion and identification

- **Storytelling** was particularly prevalent in several profiles. Typical instances were of oppression that serve as personified illustrations of more general ideological opinions. These stories may fuel the reader's sense of **adhesion and identification**, and thereby strengthen their connectedness to the propaganda.

## 7.2. Ingroup vs outgroup

- Extremist profiles use many negative Islam-based terms to express opposition against outgroups. These terms are sometimes left untranslated, possibly to increase legitimacy and brotherhood, which can fuel the psychological force of **social bonding**.
- These profiles do not particularly promote the **ingroup's identity in positive terms**, which might favour identification or fuel the superiority bias.

## 7.3. Emotion and stress

- **Emotion and stress** are potential key mental processes, conjured by **intensification. Up-scaling adverbs** like 'never' or 'always' in posts about the outgroup emphasize the unsolvable incompatibility between the ingroup and the outgroup, avoiding any grey zone.
- Visual intensification took the form of **close-ups of the victims' injuries**.

## 7.4. Preference for simplicity

- Intense crisis-related posts contrast with **presuppositional, solution-related statements**, which often take the form of short, religious principles (e.g., "Nationalism is scum") or imperatives (e.g., "Don't vote"), which can further reinforce the polarization between the ingroup and the outgroup. This taken-for-grantedness might contribute to the audience's **projection bias, false consensus effect, bandwagon effect and/or Occam's razor effect**. In case of projection bias, the degree to which other people agree with oneself is overestimated. This bias is close to false-consensus effect, where one's own opinions are overestimated as the normal and typical ones. This **social norming** can also strengthen a bandwagon effect, whereby one increasingly adopts views and behaviours the more they have already been adopted by others, as well as Occam's razor, which is a preference for simplicity.
- **Taken-for-grantedness** can be expressed through various textual arrangements, especially presupposition. For example, posts with literal religious doctrine, like "Halloween is shirk [disbelief]", are formulated as presuppositional statements: they are not up for discussion and can be treated as a 'given'.<sup>10</sup> When there is some implicit room for discussion, the writer may oppose alternative positions by presenting his views as highly warrantable through "proclaim" strategies <sup>11</sup>, i.e., emphasizes on their credibility (e.g., naturally, obviously, there can be no doubt that, of course) or endorsement formulations by which external sources are construed as undeniable (e.g., this document shows, demonstrates). These patterns

to state or increase **credibility** can also strengthen **projection bias**, **false consensus effect** and **bandwagon effect**, as well as **Occam's razor**.

- The purpose of the vast majority of these posts is not to convince or to win the debate of ideas; their function is to **denounce** situations that are framed as presuppositional and taken for granted, with much recourse to intensification patterns.

### 7.5. Social bonding and social norming

- The use of the inclusive pronouns **'we' and 'our'** is prevalent. They help construct the ummah as an **imagined homogenous ingroup community**, as opposed to the outgroup. This and the frequent address to the reader with the 'you' pronoun may also respond to the reader's **quest for social status** within one's collective).

- Some posts were built upon the **us vs. them opposition**, often strengthened by the outgroup's gaze to the external viewer. This binary opposition can foster **polarization** and the **outgroup homogeneity effect**, where individuals perceive the ingroup members as diverse but the outgroup members as more similar to one another.

- Only one quarter of the pictures contain a visual **address** to an external reader. Pictures including gaze address (i.e. someone is looking into the lens) can be seen as a demand of virtual relations with the reader.<sup>12</sup> The **picture superiority effect** is mostly used as intense visual evidence of oppression rather than as a tool for social norming (e.g. the famous "We want you..." slogan).



**Image 2:**

Us vs. them opposition in a post without visual address to an external viewer ("Yemen. Sweets for your children and bombs for us! Thank you, France.")

## 7.6. Authoritarianism

- Between one third and a half of the profiles' posts were composed of quotes. Quotes written in sacred texts or by inspirational extremist figures can trigger **authoritarianism**.
- The high frequency of addresses to the reader with **imperatives** emphasizes how many of these posts are constructed as social norming, which can also fuel **authoritarianism**. For example, the post below uses a positive ingroup visual nomination in a call to "surround yourself with those who have the same mission as you".



**Image 3:**  
Reference to the lion  
as ingroup's positive  
nomination strategy

## 7.7. Confirmation of existing beliefs

- The low frequency of **confirmation patterns** reveals that the vast majority of these posts are not constructed with linguistic patterns that might trigger **confirmation bias**, wherein people prefer messages that align with their existing beliefs. This is not surprising, given that confirmation patterns are very specific (compared to more common anti-outgroup terms, for example).

## 8) Why did PRECOBIAS focus on right-wing extremism?

Like for Islam-based extremism, we sought to investigate to what extent and through what strategies extreme-right social movements or individuals have managed to share violent extremist content in text-image posts on Facebook and Instagram in recent years (2017- 2020).

We analyzed 500 extreme-right Facebook and Instagram posts, all including text and visual content.

Since research focuses on the most established extreme-right political parties or social movements, we decided instead to analyze practices carried out on a smaller scale, by relatively small groups or individuals, whose communication practices are not designed by political communication strategists but who, nevertheless, enjoy some popularity on Facebook or Instagram, and publish content quite regularly:

- Facebook : Action Nationale et Radicale/ANR <sup>13</sup>
- Facebook : Le Parti nationaliste français – branche de Lyon /PNF
- Instagram: Europa Invicta/EI
- Instagram: Patriotен\_nrw/NRW
- Instagram: one account run by an extreme-right individual/I3.

## 9) What did we learn about topics and frames in extreme right propaganda?

### 9.1. The high salience of anti-outgroup posts

The vast majority of the posts contain anti-outgroup feelings. Sometimes, they only use their logos to express such feelings. The frequency of anti-outgroup feelings varies greatly across the profiles.

Very often, posts cover a specific topic but then also contain a disparate multitude of hashtags, some covering mainstream ideas, while others referring to several outgroups.

### 9.2. Crisis, identity and solution constructs

Unsurprisingly, **migration** remains a core issue. Migration as a threat is often covered in posts which contain two or more outgroups (i.e. outgroup mix).

We observed an **ideological frame extension** <sup>14</sup> that often addresses new themes while maintaining hostility against migrants and Muslims

as a key issue. This frame extension is mainly created through the use of the “piggybacking” tactic,<sup>15</sup> where the user mixes trending hashtags with some extremist ones as a means of infiltrating mainstream trending topics.

**Extremist leaders** were regularly framed as inspiring figures in three accounts. Positive ingroup values and ways of life (European identity, culture, heterosexuality, etc.) were observed in every profile.

Lastly, our analysis delivered results for three solutions constructs. The **apology for non-violent reaction** (e.g., participation in demonstrations or meetings) was observed in every account. **Ambiguous calls to actions** were also identified in all accounts. Posts expressing **apologies for violence** were observed in the two French accounts (ANR and PNF). With I3, these accounts are the same ones that frame extremist leaders as inspiring figures.

## 10) Are there differences between Facebook and Instagram for extreme right propaganda?

The profiles differ greatly: they oppose different outgroups and combine the crisis-identity-solution frames in different ways. Interestingly, we observed that the visual strategies across the two platforms were not significantly different from each other. While EI takes advantage of Instagram’s visual nature, with hipsterish images and vague rhetoric (often taking the form of narcissistic ideas of self-love,<sup>16</sup> the other two Instagram profiles, namely EI and NRW, prefer to publish the same types of content that we found on their Facebook pages.

## 11) What did we learn about extreme right communication styles/discourse patterns?

Apologies for violence and inspiring banned extremist figures, both constituting a breach of Facebook’s terms of service, were rather uncommon. At the other end of the spectrum, ambiguous calls for action that play cat-and-mouse with Facebook’s content rules were prevalent in some cases. In total, we observed six cat-and-mouse discourse patterns and five explicit apologies for violence, some of which are similar to those

used in Islam-based extremist propaganda:

- Extremist figures as inspiring:
  - Promotion of extreme-right figures through future or past political events
  - Quotes from an extremist figure
  - Visual portrait-tribute with years of birth and death, coupled with the organisation's logo in the image, and quote in the text section
  - Logo of a banned organisation coupled with explicit support in the text section
  
- Ambiguous calls for action:
  - Vague rhetoric of fight or defence
  - Inclusiveness of all means of defence
  
- Apologies for violence:
  - Endorsement of violent acts in text and in image (e.g., the Chemnitz incident)
  - Textual call for death in extended crisis descriptions, coupled with an image illustrating the crisis
  - Violent imagery against a specific outgroup (e.g., antifa)
  - Violent imagery coupled with a textual call for violence against a specific group
  - Decontextualization of the third Reich's army's pictures

**Doublespeak**, which allows ideological innocence,<sup>17</sup> was particularly observed. The message is vague and implicit enough to circumvent hate speech regulations but is, at the same time, clear enough to be understood as extremist ideas. While hardcore extremism has long since switched from public accounts to private ones or other platforms, Facebook and Instagram's public profiles can still represent points of entry to extremist ideology. El perfectly illustrates this **mainstreaming process**. While most of the posts are graphically more reminiscent of perfume ads than of extremist propaganda, posts like the one with the slogan "Hail Europe" with a European village in the background reminds us that the coding game with the Nazi repertoire is not far off.

## 12) How do radical posts affect cognitive biases?

The analysis of the social media posts revealed that cognitive biases might be triggered by certain stylistics and narratives. To shed light on the radicalization process and identify what measures need to be taken to lever out effects of radicalization, we conducted a survey experiment. In this, we included three biases that were particularly prevalent in extremist content and presented them in the style of Instagram-like Internet memes, as they would typically be presented on social media.

A sample of 393 people participated in our study. They started out to answer some questions about their general values and worldviews, then they saw social media posts that we created for the purpose of this experiment. We decided to test three different, commonly encountered narratives as they might relate differently to cognitive biases and the radicalization process. All three narratives related to the migrant situation in Europe, which was prevalent in the public discourse at the time of the experiment, because of the conflict situation between the EU and Turkey. The chosen narratives for the experiment were based on statements that could be encountered in comment sections of newspapers and all respondents were extensively briefed about the content of the questionnaire to ensure an ethical process. The following three narratives were tested:

(A) Radicalization often draws on the concept of an enemy that is seen in established structures and systems. We therefore worked with a narrative that referenced **elite-critique**. We presented three postings that talked about how the political elite has “failed the people” in the eye of the migrant situation in Europe and how it was time to look for alternative solutions.

(B) For the **ingroup-outgroup** narrative, we again showed very similar posts, but we explicitly referred to migrants as the outgroup that the people needed to join against.

(C) For the **violence** narrative, we included explicit invocations of violence against migrants.

We compared these three extremist narratives to posts that either discussed a topic totally unrelated to politics, or we showed them posts that presented political views on the migrant situation but in a very objective and calm manner. After our respondents saw the posts, they answered questions to assess whether the posts triggered the three cognitive bia-



ses. Afterwards, we asked for their extremist attitudes and assessments of the presented posts. Since we know from extant research that not all humans are equally vulnerable to extremist attitudes, we also analyzed whether people with high levels of authoritarianism reacted more strongly to extremist messages.

At the end of the study, respondents were educated about the source of the posts. We distanced ourselves from the radical content, and we provided additional information on extremism online as well as responsible regulatory organizations. This is a so-called debriefing process that ensures there are no long-lasting effects on the participants.

### **1. Ingroup-outgroup bias**

The so-called *ingroup-outgroup* bias describes a superior view of one's own ingroup while simultaneously downgrading outgroup members.<sup>18</sup> The analysis of extremist posts described under point 3-11 has shown that radical messages strongly use narratives that try to show a positive image of the ingroup and build a strong sense of community, while highlighting the negative characteristics of the outgroup. Moreover, a strong sense of community with the ingroup and conflicts with outgroup members can foster radicalization processes.

As far as our results are concerned, we found that a single exposure to radical materials did not increase the ingroup-outgroup bias. However, if people in general held this belief, the ingroup-outgroup bias was positively associated with people's extremist attitudes. That is, the higher the ingroup-outgroup bias was present, the higher the extremist attitudes. Holding the belief of a better ingroup was more common in people who were also more authoritarian. Hence, people who longed for a strong leader and a more authoritarian system were more likely to believe that their ingroup was superior.

### **2. Negativity Bias**

As also shown in the analysis of radical material online, these posts often reference very negative narratives by showing victims of war, expressing threats and violence, as well as focusing on the fall of an existing, hostile social order. We thought that these negative associations could maybe raise more awareness in Internet users and could activate extremist attitudes. This is formalized in the so-called *negativity bias* which outlines that humans tend to respond more strongly to negative content.<sup>19</sup> This relates to the explanation stemming from evolution that humans are inclined to watch out for threats to ensure survival.<sup>20</sup> Thus, humans tend to respond more strongly to negative stimuli, which is showcased in closer attention to, and better memory for, negative information.

In our results, when posts tackled a political issue or when they were rather explicit in their radical ideas, they were also perceived as more negative. Not surprisingly, then, violent posts showed the highest effect on negativity bias, followed by elite critique and lastly the ingroup-outgroup narrative. When a person was less authoritarian, he or she reacted with a more pronounced negativity bias to the posts. However, a bias towards negativity did not influence the extremist attitudes in a person. The intention to like or share the posts was lower when it triggered a higher negativity bias. Hence, while extremist narratives certainly gained attention, they did not contribute to extremist attitudes or sharing this content online.

### **3. Just-world hypothesis**

The analysis of extremists' posts has indicated that victimization of members of the ingroup and targeting outgroup members with violent invocations and threats is a relevant frame. If readers are confronted with such presentations, they need to deal with the depictions of injustice. A common bias that can work as a coping mechanism is the just-world hypothesis. This bias describes the belief that the world is a just place and, therefore, people generally get what they deserve.<sup>21</sup> To obtain this world view, a victim's circumstances often will be connected to her own actions and/or character. The belief that a victim of cruelty and violence might have played a role in these actions shields the observer from the reality of an unjust world and thus allows a preservation of their own well-being.<sup>22</sup>

With our study, we found that the violence narrative activated the just-world belief more than the other posts. Hence, when confronted with invocations of violence, the belief of the world as a just place increased. This is probably a coping mechanism to deal with the presented content. Holding just-world views, then, can translate into extremist attitudes, because beliefs of victims being in part responsible for violence and threats they experience as normalization of extremist views. People who hold authoritarian beliefs were also more inclined to holding just-world views.

## **13) What are the main results of the effect perspective?**

- Extremist narratives on social media can trigger a negativity bias in just one exposure. The negativity bias, however, plays no role in holding extremist attitudes.

- While an ingroup-outgroup bias cannot be built through a single exposure to extremist posts, holding beliefs of one's ingroup being superior to an outgroup positively relates to higher levels of extremist attitudes.
- Violence narratives in posts increased just-world beliefs, which are related to higher extremist attitudes.
- Holding authoritarian attitudes made just-world beliefs and ingroup-outgroup biases more likely, while they decreased the negativity bias to some extent.
- In addition, our results indicate that violence narratives directly increase extremist attitudes. This might lead back to the normalization of extremist opinions through violent social media posts.

## 14) What is the project take-away?

Cognitive biases shape how humans process information and are relevant in the decision-making process of our day-to-day lives. Extremists have found ways, however, to take advantage of these biases and subconsciously, or sometimes consciously, include them in their employed stylistics and narratives. If these biases are activated, that also makes us more vulnerable to extremist attitudes, particularly if people are already drawn to authoritarian attitudes. Educating the public about the ways cognitive biases impact media perceptions and information processing and how they can be enforced in content we are confronted with online, therefore, seems to be an important media literacy endeavour.

### Want to know more?

These findings are further developed in scientific papers that will be published on our website [www.precobias.eu](http://www.precobias.eu) in the upcoming months. Don't hesitate to download them!



# Notes

<sup>1</sup> McCauley, C., & Moskalenko, S. (2008). Mechanisms of political radicalization: Pathways toward terrorism. *Terrorism and Political Violence*, 20(3), 415-433.

<sup>2</sup> Jensen, M. A., Atwell Seate, A., & James, P. A. (2018). Radicalization to violence: A pathway approach to studying extremism. *Terrorism and Political Violence*, 32(5), 1067-1090.

<sup>3</sup> Frischlich, L., Rieger, D., Morten, A., & Bente, G. (2018). The power of a good story: Narrative persuasion in extremist propaganda and videos against violent extremism. *International Journal of Conflict and Violence (IJCV)*, 12, a644-a644.

<sup>4</sup> See e.g. Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, 63(2), 221-243.

<sup>5</sup> Ehrlinger, J., Readinger, W. O., & Kim, B. (2016). Decision-making and cognitive biases. *Encyclopedia of Mental Health*, 12.

<sup>6</sup> Bindner, L., & Gluck, R. (2019). Jihadist extremism. In *Extreme Digital Speech: Context, Responses and Solutions* (pp. 19–26). VOX-Pol Network of Excellence.

<sup>7</sup> Ingram, H. J. (2016). Deciphering the Siren Call of Militant Islamist Propaganda: Meaning, Credibility & Behavioural Change. *The International Centre for Counter-Terrorism – The Hague*, 7(9); Ingram, H. J. (2018). Islamic State's English-language magazines, 2014-2017: Trends & Implications for CT-CVE Strategic Communications. *The International Centre for Counter-Terrorism – The Hague*, 8(15).

<sup>8</sup> Baker, P., & Vessey, R. (2018). A corpus-driven comparison of English and French Islamist extremist texts. *International Journal of Corpus Linguistics*, 23(3), 255–278.

<sup>9</sup> Bazzi, S. (2019). The non-translation strategy in translating ISIS radical discourse. *Journal of Language Aggression and Conflict*, 7(2), 133–155.

<sup>10</sup> Martin, J. R., & White, P. R. R. (2005). *The language of evaluation. Appraisal in English*. Palgrave Macmillan, 102.

<sup>11</sup> Idem, 121-132.

<sup>12</sup> Kress, G., & van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design (2nd Edition)*. Routledge, 118-149.

<sup>13</sup> Action Nationale et Radicale<sup>13</sup> /ANR, meaning 'National and Radical Action' is a group of French nationalist activists who advocate "racial nationalism as opposed to pseudo-patriotism that is multi-coloured" (Action Nationale et Radicale, 2020). They intend to commit to the growing invasion of Europe, which threatens the same blood that runs in the veins of all Europeans. The preservation of their race and social justice are their priorities (ANR, 2020).

- Le Parti nationaliste français – branche de Lyon /PNF, meaning the French nationalist Party –City Lyon branch, is not a political party but a French political movement that notably considers that "each human being can only find their magnitude, strength and fulfilment in the group from which they come" (PNF, 2020).

- Europa Invicta/EI, meaning 'Undeclared or unconquered Europe', is an Instagram account in English followed by 7,317 users which was turned into a private group in July 2020. His creator, who notably praises Marine Le Pen and Viktor Orbán, explained how Europa Invicta was created to celebrate Europe's landscapes, cultural heritage and technological advances in response to the negative feelings caused by the immigration threat to the Old Continent, a.o.

- Patrioten\_nrw/NRW is a little German group of extreme-right activists, who a.o. organize actions and demonstrations with neo-Nazis, partly in North Rhine-Westphalia.

- The third Instagram account in our corpus is an account in German run by an extreme-right individual. Unlike organizations, we did not record the user handle to maintain confidentiality and we re-labeled it using the anonymized identifier IN3. This account defends the theory of the Great Replacement and uses an anti-migrant rhetoric.

<sup>14</sup> Berger, J. M. (2016). *Nazis vs. ISIS on Twitter: A Comparative Study of White Nationalist and ISIS Online Social Media Networks*. Program on Extremism at George Washington University. [https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/downloads/Nazis\\_v.\\_ISIS.pdf](https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/downloads/Nazis_v._ISIS.pdf); Berger, J. M. (2018). *The Alt-Right Twitter Census. Defining and Describing the Audience for Alt-Right Content on Twitter*. VOX-Pol Network of Excellence ;

Schedler, J. (2014). The devil in disguise: Action repertoire, visual performance and collective identity of the Autonomous Nationalists. *Nations and Nationalism*, 20(2), 239–258.

<sup>15</sup> Graham, R. (2016). Inter-ideological mingling: White extremist ideology entering the mainstream on Twitter. *Sociological Spectrum*, 36(1), 24–36.

<sup>16</sup> Back, L. (2002). *When Hate Speaks the Language of Love. Racism in the Age of Information* [The Social Movement Studies Conference].

<sup>17</sup> Tuters, M. (2019). LARPing & Liberal Tears. Irony, Belief and Idiocy in the Deep Vernacular Web. In M. Fielitz & N. Thurston (Eds.), *Post-Digital Cultures of the Far Right. Online Actions and Offline Consequences in Europe and the US* (pp. 37–48). Transcript Verlag.

<sup>18</sup> Brauer, M. (2001). Intergroup perception in the social context: The effects of social status and group membership on perceived out-group homogeneity and ethnocentrism. *Journal of Experimental Social Psychology*, 37(1), 15-31.

<sup>19</sup> E.g., Zillmann, D., Chen, L., Knobloch, S., & Callison, C. (2004). Effects of lead framing on selective exposure to Internet news reports. *Communication Research*, 31, 58-81.

<sup>20</sup> Soroka, S., & McAdams, S. (2015). News, politics, and negativity. *Political Communication*, 32, 1-22.

<sup>21</sup> Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, 85(5), 1030.

<sup>22</sup> Strelan, P. (2007). The prosocial, adaptive qualities of just world beliefs: Implications for the relationship between justice and forgiveness. *Personality and Individual Differences*, 43(4), 881-890.



# PRECObIAS

**#thinkaboutit**



This report was funded by the European Union's Internal Security Fund — Police. Project n.: 867186

The content of this report represents the views of the author only and is his/her sole responsibility.

The European Commission does not accept any responsibility for use that may be made of the information it contains.