Toward an Understanding of the Characteristics of Secondary School Cyberhate Perpetrators

Catherine Blaya[1,2]* and Catherine Audrin[1,3]

[1] Faculty of Teachers Education, Lasale, HEP du Canton de Vaud, Lausanne, Switzerland, [2] URMIS - University Nice Côte d'Azur, Nice, France, [3] Swiss Center for Affective Science, University of Geneva, Geneva, Switzerland

While the Internet offers many opportunities to access information, training and communication, it has created new grounds for risks, threats and harm. With the rise of populism and extremism, new forms of cyberbullying emerge, more specifically cyberhate. The Internet has become a privileged tool to disseminate hatred, based on racism, xenophobia, bigotry, and islamophobia. Organized groups use the internet as a dissemination tool for their ideas, to build collective identity and to recruit young people. The presence of these groups has been facilitated worldwide thanks to technology. Yet, little attention has been granted to the way the Internet eases the activities of individuals who promote and propagate hate online. The role they play in spreading racism, xenophobia and bigotry is paramount as they regularly comment online about news and events, interacting with like-minded people with impunity because the web prevents people from being easily identified or controlled. While literature on exposure to hateful contents and cyberhate victimization is growing, little is known about who the perpetrators really are. A survey with young people aged 12–20 ($N = 1,889$) was completed in France and forms the basis of this article. It provides an understanding of the characteristics and associated variables of cyberhate perpetration. The Structural Equation model shows that cyberhate perpetration is heavily related to time spent online, victimization, belonging to a deviant youth group, positive attitudes toward violence and racism. Results from the SEM further suggest that people who suffered from online victimization will themselves have a greater tendency to belong to deviant youth groups. Multiple mediation analysis further suggests that trust in institutions may however prevent young people from belonging to a deviant youth group and decrease positive attitudes toward violence, thus diminishing the tendency to perform hateful aggression.

Keywords: cyberhate, involvement, young people, perpetration, victimization, characteristics

INTRODUCTION

Young people are super-connected to the Internet and electronic communications are an integral part of their life (Boyd, 2014). France has the fourth largest number of Internet users of any country in Europe and is 17th worldwide. According to the last national poll on the digital practices IPSOS in France, teens widely use media services, spending on average 15 h per week online: 13–19 years-old spend over 15 h a week on the Internet while 7–12 years-old spend about 6 h. The vast majority of teens (81%) has a Smartphone. More than one in three young people (34%) use a tablet and it is common for them to own a personal device. Video game tablets are owned by 69% of the

surveyed young people. The applications that score the biggest success are YouTube (79%) followed by Facebook (77%) and Snapchat (57%). YouTube and Snapchat are the platforms that have shown the fastest growth in teens Internet usage lately.

The online world offers young people many opportunities to access information and knowledge, to explore their own identity as well as to communicate with others (Mishna et al., 2010; Boyd, 2014). However, the Internet and electronic communication tools can be used either positively or negatively. Notably, they can be used to convey antagonistic, hateful, racist and xenophobic content. In Europe, with the rise of populism and extremism, hate crimes and hate speeches have increased over the last decade (FRA (Fundamental Rights Agency), 2014; Penzien, 2017). While some findings may be controversial (Vitoroulis and Vaillancourt, 2015), research suggests that, indeed, ethnic minorities are subjected to hateful bullying and are more vulnerable than majority groups, both in the US and in Europe (Tynes, 2005; Hawdon, 2014; Llorent et al., 2016).

Social media have become a free, easy to use, and privileged tool for propaganda and victimization especially among young people (Blaya, 2019). Online hate speech and incitement have a potentially greater impact when spread in social media (Recommendation No. R (97) 20 of the Committee of Ministers of the Council of Europe on "hate speech"). Many countries have issued legislation to protect people from groups and organized individuals who use the web to propagate and incite hatred. Hatred is also spread by "ordinary" people. As stressed by Potok (2015), individuals have become more active than organized groups and produce hate that is widely disseminated through posts, comments and user-generated social media platforms. This should get full attention from part of decisionmakers, researchers and educators, due to the rise in anti-Semitism, Islamophobia and xenophobia throughout Europe and beyond. Such phenomenon may have dramatic consequences in terms of stigmatization and alienation on both individuals and society.

Research has increasingly investigated the exposure and victimization of individuals and communities to online hateful content (Oksanen et al., 2014; Hawdon et al., 2017; Blaya, 2019). However, only one research documents the involvement of the young people as perpetrators in the US (Costello and Hawdon, 2018). In Europe, there is a gap in research focusing on cyberhate perpetrators. This article is therefore attempting to address the current gap with an online survey completed in France by 1,889 young people aged 12–20.

RESEARCH BACKGROUND

Defining Cyberhate

Cyberhate is related to cyberbullying. However, although they use similar means and happen in similar context, there are differences between these two forms of aggression:

- In the case of cyberhate, individuals or communities are targeted because of supposed, specific or identified characteristics such as their physical appearance, religion or the language they speak.

- Even when individuals only are targeted, cyberhate expresses inter-group hostility (Hawdon, 2014). It can be the consequence of the competition for economic wealth or power, the feeling that one's identity is being threatened. Hatemongers thrive on this (Sherif and Sherif, 1969).
- The consequences of being exposed or a direct target of cyberhate not only generate individual or community unrest but also contribute to alter social cohesion and democracy/human rights.

Literature [be it journalistic (Knobel, 2012), from the associative sector (Messmer, 2009) or scientific] often refers to cyberhate as a virus that spreads like an infectious disease in our societies, affecting the most vulnerable people (Foxman and Wolf, 2013). The Council of Europe, in its Additional Protocol to the Convention on Cybercriminality of 28 January 2003, extends its scope to criminalize racist and xenophobic speech and propaganda via computer systems, states that:

"'Racist and xenophobic material' means any written material, any image or any other representation of ideas or theories, which advocates, promotes or incites hatred, discrimination or violence, against any individual or group of individuals, based on race, color, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors." (Art. 2-1).

For its part, the Anti-Defamation League defines cyberhate as:

"Any use of electronic communications technology to spread anti-Semitic, racist, bigoted, extremist or terrorist messages or information. These electronic communications technologies include the Internet (i.e., Web-sites, social networking sites, "Web 2.0" user-generated content, dating sites, blogs, on-line games, instant messages and e-mail) as well as other computer- and cell phone-based information technologies (such as text messages and mobile phones)." (Anti-Defamation League, 2010)

The Center for Equal Opportunities and Opposition to Racism in Brussels, defines cyberhate as "the propagation of hate speech on the Internet." It refers to hatred in the form of bullying, insults, discrimination on the Internet against individuals or groups of people on the grounds of their skin color, supposed race, ethnic origin, sex, sexual orientation or political or religious beliefs. It also refers to anti-Semitism and historical revisionism (p. 8, 2009).

We define cyberhate as electronic communication initiated by hate groups or individuals, with the purpose to attract new members, build and strengthen group identity; it aims at rejecting others' collective identity. The means used are the publishing of propagandistic messages, incitation to discrimination, violence, and hatred against individuals and their community with the view to potentially disaggregate social cohesion, on the ground of color of skin, religion, national or ethnic origin. In this research, we use the term "cyberhate" to refer to all hateful online forms of expression (text, images, videos, pictures, graphic representations) whose objective is to belittle, humiliate or ridicule a person or group of persons, by generating hatred or rejecting these persons or their communities who genuinely or

supposedly belong to a specific ethnic or religious background different from theirs.

Exposure to Cyberhate

A pioneer research in the United States and Europe, shows that 53% of respondents are exposed to online hate content and 16% are personally targeted (Oksanen et al., 2014; Hawdon et al., 2017). In Europe, the *Net Children Go Mobile* survey (Mascheroni and Ólafsson, 2014) reports that children aged 9–16 are becoming increasingly exposed to hateful comments. A research in France, concludes that one third of the surveyed youth are exposed to hate online (30%) (Blaya, 2019). Although exposure does not systematically lead to victimization, evidence suggests that being exposed to hateful content is linked to lower self-esteem, enhanced feeling of insecurity and fear as well as mental health issues such as mood swings (Tynes, 2005; Blaya, 2019). Exposed individuals are more likely to be associated with delinquent peers, to live alone and have higher levels of education (Schils and Pauwels, 2014; Hawdon et al., 2017). They also spend longer time on the Internet and are multi-users of online applications and services (Hawdon, 2014; Costello et al., 2016; Keipi et al., 2017). Being exposed to cyberhate is also associated to cyberbullying and violence (Leung et al., 2018) and to recruitment to extremist organizations (Foxman and Wolf, 2013). Finally, research reveals that exposure (put exposure to what) may be correlated with detrimental effects on a societal level: exposure is potentially linked to an increase in hate crimes offline, a lack of social trust, tougher discrimination, and prejudice against the targets (Näsi et al., 2015; Keipi et al., 2017), this including spreading extremist and violent ideology (Tynes, 2005; Foxman and Wolf, 2013).

Cyberhate Victimization

In Canada, the "*Young Canadians in a Wired World*" study (Taylor, 2001) indicates that 14% of Instant Messaging users had suffered threats. In the United States, a research among Afro-American, Latinos, Asians and Métis communities showed evidence of an increase in cyberhate victimization from 2010 to 2013 (Tynes et al., 2015). Statistical analyses do not identify any difference between male and female respondents nor between the involved ethnic groups. However, older respondents showed higher rates of victimization which might be due to longer hours spent online. Research shows that victims adopt more at-risk online behaviors such as spending long time online, being more active on social networking sites, and visiting potentially dangerous websites (i.e., promotion of self-inflicted violence) (Keipi et al., 2017; Costello et al., 2018). Being faced with hate speech can encourage people to adopt violent behaviors. Consequences range from lower self-esteem, mental health issues such as high levels of anxiety, identity erosion, anger, fear, to adopting violent behaviors (Leets, 2002; Tynes, 2005).

Cyberhate Perpetrators

A paper by Costello and Hawdon (2018) on a survey including a random sample of Americans aged 15–36, shows that one fifth of the respondents acknowledged producing and disseminating cyberhate on the grounds of ethnic origin, religion or color of skin. Their findings show that males are significantly more involved in online hate perpetration than females. They also highlight that perpetrators more often use specific sites such as Reddit, Tumblr and messaging boards. Belonging to an online community or visiting sites that spread hatred increases the probability of producing cyberhate. Having favorable attitudes toward violence and being submitted to some social pressure leads people feeling comfortable performing offline violent behavior (Ajzen, 1991). Online, the "filter bubble effect" as explained by Pariser (2011) virtually encloses in the same network, individuals who share similar ideas and live in a similar identity bubble. As for cyberhate, sharing subversive ideology, and positive attitudes toward xenophobia, racism, or bigotry can intensify the risk of perpetration (Costello and Hawdon, 2018).

Along findings showing that being a victim of "mainstream" cyberbullying increases the likelihood to be involved as a perpetrator, being the target of cyberhate is correlated with higher odds to become a perpetrator (Hawdon, 2014). However, unlike other forms of cyberbullying, spending much time online or playing first-person shooter games does not seem to influence the spread of hatred (Costello and Hawdon, 2018). Cyberhate exposure and victimization are negatively associated with trust in people in general (Keipi et al., 2017; Näsi et al., 2017). The link between trust in institutions and the perpetration of cyber hatred has not been studied yet.

As Perry and Olsson (2009) argue, spreading cyberhate contributes to consolidating hatred in real life. Racist individuals use the Internet to disseminate their ideas and to confirm their racist views by connecting with people sharing the same ideas on political blogs, games, forums, and chat rooms. They use this powerful communication tool to hurt, denigrate, humiliate people and communities. Victims of racism and discrimination in real life have increased risks of offending, as they develop hostile views of others (Burt et al., 2012). However, prejudice and racism are not always conscious and can be implicit. Implicit bias is likely to trigger discriminatory and hostile attitudes in interpersonal face-to-face or online interactions. This highlights the need to investigate attitudes toward racism while investigating cyberhate, in order to assess the potential association between declared racism and cyberhate perpetration.

Relying on the definition presented previously and the above research background, the objectives of this survey were two-fold:

(1) To assess the prevalence of the involvement of young people in cyberhate in France
(2) To examine the factors contributing to the involvement of young perpetrators of cyberhate

METHODS

The questionnaire was informed by several sources including questions on deviant youth groups (DYG) from the ISRD survey (Blaya and Gatti, 2010) and some questions on digital practices from the EU Kids Online survey (Livingstone et al., 2011). The questions related to cyberhate were informed by an extensive

literature review and some exploratory interviews that were conducted prior the designing of the survey.

Sample and Procedure

The study included 16 lower and upper secondary schools. 1889 students completed an online questionnaire survey. Participants ranged from 11 to 20 years of age (age mean = 14.631, sd = 2.053), 50.24% were females, 49.75% were males. The student population of these schools is diverse ranging from upper-class schools in the center of Paris to remote rural schools in the South West of France. All students from one class randomly selected per year group completed the questionnaire. Students under the age of 18 were asked to participate providing they submitted a parental consent form. Parents were sent an explanation letter informing them on the objectives of the research, the use and management of data and the associated potential risks. All students completing the questionnaire were informed about the study and provided their consent prior to participating. As a consequence, written consent was obtained from all adult participants and from the parents of all non-adult participants. The response rate was 90%. There is no mean to check on potential differences between students who participated and those who did not. Some of them had not provided the parental consent, others were out of school the day of data completion and there were a number of personal reasons why they could not take part in the survey.

Data were collected anonymously during the year 2016. To ensure children understood the questions, the wording of the questionnaire was improved after cognitive testing with children of different age groups and gender. It was then piloted to check on completion time.

In each school, questionnaires were administered online under the supervision of a research assistant in the school's IT room and no staff was present during data collection to ensure confidentially, trust and accuracy of the students' responses. Completion took no more than 45 min.

Measurement Tool

We used the questionnaire previously used by Blaya and Gatti (2010) and Livingstone et al. (2011), which yielded good reliability indices (α = 0.95, ω = 0.96). It was made up of general questions about the socio-demographic characteristics of respondents and their families, their digital practices (ICT use), their experiences of bullying in schools, their satisfaction with life. We also asked them questions regarding their religion and attitudes toward violence, their trust in institutions, the characteristics of their peer group, and their attitudes toward racism.

Finally, we asked them about their cyberhate experience as our main variables of interest were the prevalence of exposure, victimization and perpetration of hate online amongst young people.

Demographics

Questions regarding individual and family characteristics included gender (males vs. females); age (open question);

students' school grade; the person with whom they lived, if their parents had a professional occupation and their nationality.

ICT Use

Participants were asked to assess how much time they spent online every day (1) during the week (item 1) and (2) during the week-end (item 2, α = 0.72, ω = 0.72). They could answer on a scale ranging from never (1) to 4 h or more (5). They further were asked to assess which type of applications they preferred amongst Snapchat, Messenger, Facebook, Instagram, Twitter, and online games.

School Bullying

Participants were asked to answer three items referring to how they underwent psychological violence in school (α = 0.66, ω = 0.68). Specifically, participants were asked: (1) how many times they were excluded by other students, (2) how many times they were insulted or made fun of, and (3) how many times they were insulted during the last 12 months preceding the survey. Participants could answer on a scale ranging from 1 (never) to 5 (five times and more).

Life Satisfaction

Participants were asked how satisfied they currently were with their lives. They could answer on a scale ranging from not happy at all (1) to very happy (5).

Religion and Practice

Participants were asked (1) if they had a religion (yes or no), (2) if so, what their religion was, and (3) how often they practiced their religion. For this last question, participants answered on a scale ranging from 0 "I am not a religious person" to 4 – "I am very religious".

Attitudes Toward Violence

Attitudes toward violence were measured by seven items (α = 0.96, ω = 0.98). These items asked participants if they thought that fighting was legitimate (1) to defend oneself (item 1), (2) to defend someone else (item 2), (3) if participants were insulted (item 3). Other items assessed that using violence was legitimate (4) in the case where someone is made fun of because of their religion (item 4), (5) to defend one's country (item 5), (6) to fight back against racism (item 6), (7) to fight for one's ideas (item 7). One item (using violence is legitimate because it is funny) was removed from the analyses because it did not significantly load on the "attitude toward violence" factor. Participants were asked to answer on a scale ranging from 1 (I don't agree) to 3 (I agree).

Trust in Institutions

Institutional trust was initially measured by 11 items (α = 0.96, ω = 0.98). For each institution, participants were asked to assess what was their level of trust. Out of the eleven items, five were kept as they significantly load on the "trust in distant institutions" factor. These items were: the parliament, the prime minister, the President of the republic, the European Union, the United nations. To define the "trust in proximal institutions," we included the following items: parents, local politicians, schools, the police and the law. Participants were asked to answer how

much they trusted each institution on a scale ranging from 1 (I don't trust this institution) to 3 (I deeply trust this institution).

Peer Group(s)

This section comprises two scales: the first scale asked about the number of friends the participants had in real life (IRL) and online and the second scale aimed to evaluate if the participants belonged to a deviant youth group—DYG (Blaya and Gatti, 2010).

- Number of friends ($\alpha = 0.74$, $\omega = 0.75$): we measured the number of friends was measured by four items. Participants were asked to assess how many friends they had (1) at school, (2) in their neighborhood, (3) in their town. They were also asked to assess how many friends they had on the web, but this item did not significantly load on the factor. Participants were asked to answer on a scale ranging from 1 (no friends) to 7 (more than 300 friends).
- Belonging to DYG ($\alpha = 0.99$, $\omega = 0.99$): Belonging to a DYG was measured using 13 items. Among those, three were retained for our survey. Item one asked participants if they had to do particular things to be admitted into the gang. Item two asked participants if doing things prohibited by the law was accepted or tolerated in their gang. Item three asked participants if members of their groups did illegal things (i.e., things that were prohibited by the law). Participants were asked to answer if this was true ("Yes," 1) or not ("No," 0).

Racist Beliefs

Questions assessing racist beliefs were included in the survey ($\alpha = 0.88$, $\omega = 0.88$). These questions were asking (1) if the participants thought some "races" were superior to others (item 1: "do you think that some races are superior to others?") (2) if racism was the product of social intolerance (item 2: "do you think that racism is caused by social intolerance?"), (3) if the participants thought that racism could sometimes be justified (item 3), (4) if they thought that victims of racism were sometimes responsible for their own fate (item 4). Finally, they were asked if they thought that racism was (1) a long-existing problem with no solution, (2) a problem which could be solved if everyone works on it and (3) a situation less serious than what is claimed.

Cyberhate Exposure

Cyberhate exposure investigated whether the participants had been exposed to hateful online messages during the 6 months prior to the survey (Yes/no/I do not know). They were further asked if they had purposefully sought such messages (never to 4 times and more).

Cyberhate Victimization

Online victimization was measured by two items ($\alpha = 0.78$, $\omega = 0.79$). These items asked participants if during the last 6 months, they had been the target of hateful or humiliating messages, comments or images (1) on their cell phone (item 1), (2) on social media (item 2). The scale ranged from 1 (never) to 4 (4 times and more).

Cyberhate Perpetration

This is our dependent variable ($\alpha = 0.63$, $\omega = 0.64$). It was measured through two items. Participants were asked if they had (1) published (item 1) or (2) shared or transferred (item 2) humiliating or hateful messages, images or comments toward one specific person or a group of persons on the Internet. Participants were asked to answer on a scale ranging from 1 (never) to 4 (4 times and more).

Data Analyses

Data were analyzed with R and Mplus and consisted of two steps: descriptive statistics and a Structural Equation Model (SEM). We performed descriptive analyses on the prevalence of participants' involvement in cyberhate and the relationship with variables such as school violence, life satisfaction, cyberhate victimization and perpetration, as well as socio-demographics. In the Structural Equation Model analysis, we tested how the perpetration of cyberhate was related to (1) attitudes toward violence, (2) cyberhate victimization, (3) belonging to a DYG, (4) trust in proximal institutions, (5) trust in distal institutions, (6) social isolation (measured by offline number of friends), (7) attitudes toward racism, (8) school bullying, as well as (9) the time spent online and 10) the use of applications. As most of our variables were categorical or ordered data, we used the WLSMV estimator. This estimator does not assume normally distributed variables, and is recommended to analyze such kind of data (Brown, 2014). Multivariate Mardia coefficient reveals that our data were not normally distributed (Mardia Skweness = 24551.252, $p < 0.001$; Mardia Kurtosis = 55.192, $p < 0.001$). We kept at least two items for each latent variable, as recommended by Kenny (http://davidakenny.net/cm/identify.htm). Items were kept if (1) their loading were equal or higher than.3, and (2) if their R^2 were higher than 0.3.

Goodness-of-fit

To assess the model's goodness-of-fit, we relied on indices having different measurement properties, as recommended by Hu and Bentler (1999). Thus, we used the root mean-square error of approximation (RMSEA), the comparative fit indices (CFI) and the Tucker-Lewis index (TLI). Browne and Cudeck (1992) suggest that models with RMSEA below 0.05 are indicative of good fit, and that values up to 0.08 reflect reasonable errors of approximation. The CFI statistic (McDonald and Marsh, 1990) reflects the "distance" of the model from the perfect fit. It is generally acknowledged that a value >0.9 reflects an acceptable distance to the perfect fit. We also reported the Tucker–Lewis index (TLI; Tucker and Lewis, 1973), which accounts for the model complexity. The TLI indicates how the model of interest improves the fit in relation to the null model. As for the CFI statistic, a TLI value equaled or >0.9 reflects an acceptable distance to the perfect fit. However, we did not report SRMR indices, because it is not computed when performing SEM using WLSMV estimator in Mplus.

TABLE 1 | Numbers and proportions of perpetrators and victims of hateful content.

	Nb perpetrators (publish)	%	Nb perpetrators (share)	%	Nb exposed	%	Nb victims (cellphone)	%	Nb victims (social network)	%
No	1752	94.7	1752	94.9	842	56.7	1633	88.6	1659	90.4
Yes	98	5.3	94	5.1	642	43.3	211	11.4	176	9.6
Total	1850	100	1846	100	1484	100	1844	100	1835	100

RESULTS

Descriptive Statistics

Before analyzing our Structural Equation Model (SEM) results, we first provide descriptive analyses of how often participants reported being either author or victim of cyberhate and if they were victims of school bullying (**Table 1**). To do this, we analyzed if participants were involved at least once or never as (1) cyberhate perpetrator either by publishing or by (2) sharing hateful content, (3) exposed to cyberhate, (4) targets of cyberhate via their cell phone or (5) via social networks.

We then focused only on participants who reported being a perpetrator at least once, either by sharing or by publishing hateful content. This represents 146 participants out of the initial 1,889 students who completed the questionnaire. The study of socio-demographic characteristics shows that there were significantly more boys among perpetrators (59.3%, $n = 83$) than girls [40.7%, $n = 57$; $\text{Chi}^2_{(1)} = 4.828$, $p = 0.027$]. The vast majority of the perpetrators had both a working mother [73%, $n = 99$, $\text{Chi}^2_{(1)} = 27.161$, $p < 0.001$], and a working father [84.6%, $n = 110$, $\text{Chi}^2_{(1)} = 62.308$, $p < 0.001$]. Regarding their potential religion, a slight majority (51%) of perpetrators answered they belonged to a religion ($n = 73$), 34.5 % reported not belonging to any religion ($n = 49$) and the rest of the sample 14% ticked "other" ($n = 21$), [$\text{Chi}^2_{(2)} = 28.42$, $p < 0.001$]. Among those who stated they had a religion, 8% said that they were not active, 32% said they were little active, 38% they were rather active and 29% reported they practiced much [$\text{Chi}^2_{(3)} = 15.822$, $p = 0.001$]. Regarding the various types of reported religions, 19% of the sample answered they were Catholic ($n = 28$), 20% Muslim ($n = 30$), and 60% reported other religious affiliations [$\text{Chi}^2_{(2)} = 47.736$, $p < 0.001$].

We also focused on perpetrators' life satisfaction. Seventy percentage of reported being either very happy or relatively happy ($n = 99$); 14% reported being neither happy nor unhappy ($n = 20$); 14% of them finally reported being not really happy to really not happy ($n = 20$), thus revealing a significant difference between the proportions of life satisfaction [$\text{Chi}^2_{(4)} = 63.67$, $p < 001$].

Regarding school bullying, results reveal that most of the perpetrators were never victims of ostracism (61.7%, $n = 87$) nor threatened (51.4%, $n = 73$). However, many of them were insulted over five times (43.6%, $n = 61$, see **Table 2**) during the last 6 months prior the survey. Results reveal significant differences in the proportions of perpetrators being excluded [$\text{Chi}^2_{(4)} = 154.14$, $p < 0.001$], insulted [$\text{Chi}^2_{(4)} = 53.87$, $p < 0.001$] and threatened [$\text{Chi}^2_{(4)} = 88.85$, $p < 0.001$].

TABLE 2 | Reported school bullying.

	Excluded	%	Insulted	%	Threatened	%
Never	87	61.7	29	20.7	73	51.4
Once	13	9.2	21	15.0	19	13.4
Twice	10	7.1	16	11.4	21	14.8
3–4 times	14	9.9	13	9.3	13	9.2
5 times and more	17	12.1	61	43.6	16	11.3
Total	141	100	140	100	142	100

TABLE 3 | Proportion of perpetrators reporting a specific search for hateful content.

	Color of skin	%	Religion	%	Origin	%	Culture	%
Never	105	80.15	103	75.18	94	74.02	98	77.17
Once	7	5.34	11	8.03	15	11.81	12	9.45
2–3 times	10	7.63	10	7.30	9	7.09	8	6.30
More than 4 times	9	6.87	13	9.49	9	7.09	9	7.09
Total	131	100.00	137	100.00	127	100.00	127	100.00

We then checked whether perpetrators had previously been the target of online hateful messages. Results suggest that the vast majority had never suffered cyberhate aggression via social networks [71.4%, $n = 100$, $\text{Chi}^2_{(1)} = 25.712$, $p < 0.001$] or via cell phones [66.6 %, $n = 94$, $\text{Chi}^2_{(1)} = 15.67$, $p < 0.001$]. Interestingly however, most of them had previously been exposed to cyberhate [$n = 98$, 80.33%, $\text{Chi}^2_{(1)} = 44.88$, $p < 0.001$].

Perpetrators were then asked if they had intentionally searched for such content, notably by looking for websites which were targeting specific (groups of) people because of their religion, the color of their skin, their origin or their culture. In all these cases, more than 75% of the sample answered that they had never specifically searched for such content (see **Table 3** for an overview).

On average, perpetrators had a neutral attitude toward violence (mea $n = 2.01$, sd = 0.49). Interestingly however, most of them agreed with the idea of using violence to defend oneself [78.5 %, $n = 106$, $\text{Chi}^2_{(2)} = 124$, $p < 0.001$], or to defend someone else [70%, $n = 91$, $\text{Chi}^2_{(2)} = 82.82$, $p < 0.001$]. Moreover, perpetrators were equally distributed between not agreeing and agreeing with the fact that it is acceptable to use violence to defend oneself against racism ["I don't agree" = 34.5%, "I neither agree nor disagree" = 29.3%, "I totally agree" = 36.1%, $\text{Chi}^2_{(2)} =$

TABLE 4 | Descriptive statistics for the items included in the SEM analysis.

		Mean	SD	Median	Skew	Kurtosis
Cyberhate victimization	Target on cell phone (Item 1)	1.193	0.596	1	3.447	11.649
	Target on social network (Item 2)	1.194	9.597	1	3.239	9.916
Cyberhate perpetration	Publish (Item 1)	0.044	0.205	0	4.424	17.600
	Share/transfer (Item 2)	0.045	0.208	0	4.344	16.897
Attitude toward violence	To defend oneself (Item 1)	2.548	0.708	3	−1.243	0.081
	To defend someone else (Item 2)	2.533	0.672	3	−1.122	−0.007
	To answer toward insults (Item 3)	1.775	0.733	2	0.378	−1.077
	To answer if someone is made fun of because of their religion (Item 4)	1.583	0.725	1	0.818	−0.680
	To defend one's country (Item 5)	1.768	0.781	2	0.427	−1.244
	To fight back against racism (Item 6)	1.785	0.731	2	0.357	−1.077
	To fight for one's ideas (Item 7)	1.391	0.596	1	1.253	0.522
Trust in distant institutions	Parliament (Item 1)	0.918	0.888	1	0.635	−0.469
	Prime minister (Item 2)	0.898	0.852	1	0.720	−0.110
	President of the republic (Item 3)	0.926	0.826	1	0.638	−0.135
	European Union (Item 4)	1.053	0.985	1	0.494	−0.872
	United nations (Item 5)	1.089	1.012	1	0.438	−0.998
Trust in proximal institutions	Parents (Item 1)	2.282	0.509	3	−3.285	11.444
	Local political (Item 2)	0.837	0.765	1	0.608	−0.121
	Schools (Item 3)	1.848	0.947	2	−0.421	−0.749
	Police (Item 4)	1.677	0.967	2	−0.237	−0.913
	Justice (Item 5)	1.579	0.912	2	−0.195	−0.763
Belonging to DYG	Do particular things to be admitted into the group (Item 1)	0.065	0.247	0	3.501	10.271
	Tolerance and acceptance in doing prohibited things (by the law) by the group (Item 2)	0.361	0.480	0	0.578	−1.668
	Group members do illegal things (i.e., things that were prohibited by the law) (Item 3)	0.362	0.481	0	0.572	−1.676
Number of friends	At school (Item 1)	3.49	1.35	3	0.001	0.623
	In their neighborhood (Item 2)	2.49	1.49	2	0.001	0.755
	In their town (Item 3)	2.71	1.49	2	0.001	0.604
Racism	Some races are superior to others (Item 1)	0.058	0.234	0	3.754	12.116
	Racism is caused by social intolerance – reversed item (Item 2)	0.647	0.478	1	−0.616	−1.622
Time online	Time during the week (Item 1)	3.703	1.208	4	−0.368	−1.155
	Time during the week-end (Item 2)	4.155	1.083	5	−0.963	−0.300
	Youtube	0.917	0.275	1	−3.022	7.147
Applications	Snapchat	0.715	0.452	1	−0.948	−1.101
	Facebook	0.575	0.493	1	−0.343	−1.883

TABLE 5 | Regression coefficients on latent factor "cyberhate perpetration."

	Standardized estimate	95% lower	95% upper	Significance
Cyberhate victimization	0.414	0.299	0.530	0.001***
Attitude toward violence	0.104	−0.008	0.216	0.068
Trust in proximal institutions	−0.031	−0.200	0.137	0.718
Trust in distal institutions	0.012	−0.153	0.177	0.888
Number of friends	0.052	−0.062	0.166	0.369
Belonging to DYG	0.406	0.236	0.576	0.001***
Racism	0.240	0.061	0.427	0.009**
Time online	0.214	0.044	0.383	0.014*
Use of online applications	−0.159	−0.383	0.065	0.163

*** $p < 0.001$; ** $p < 0.01$, * $p < 0.05$.

1.00, $p > 0.05$] and to defend oneself if one is insulted ["I don't agree" = 27.1%, "I neither agree nor disagree" = 41.2%, "I totally agree" = 31.3%, $Chi^2_{(2)} = 3.9$, $p > 0.05$].

Regarding how trusting perpetrators were toward the institutions, results suggest that they have a low level of trust (mean = 1.30, sd = 0.44). Interestingly however, perpetrators strongly trusted their parents [94%, $n = 138$, $Chi^2_{(3)} = 225.5$, $p < 0.001$] and their friends [87%, $n = 118$, $Chi^2_{(3)} = 78.17$, $p < 0.001$]. In contrast, they showed less confidence in local politics [13.2%, $n = 65$, $Chi^2_{(3)} = 77.44$, $p < 0.001$], school [41%, $n = 54$, $Chi^2_{(3)} = 55.9$, $p < 0.001$], the parliament [12.8%, $n = 16$, $Chi^2_{(3)} = 77.67$, $p < 0.001$], the prime minister [15.6%, $n = 21$, $Chi^2_{(3)} = 71.27$, $p < 0.001$], and the President of the republic [17.1%, $n = 22$, $Chi^2_{(3)} = 79.25$, $p < 0.001$].

TABLE 6 | Factor loadings for each latent variable.

	Estimate	S.E.	Standardized estimate	p-value	Lower 95%	Higher 95%
CYBERHATE VICTIMIZATION						
Target on cell phone (Item 1)	1.000	0.000	0.910	0.001***	0.822	0.990
Target on social network (Item 2)	1.027	0.098	0.923	0.001***	0.832	1.019
CYBERHATE PERPETRATION						
Publish (Item 1)	1.000	0.000	0. 859	0.001***	0.780	0.937
Share/transfer (Item 2)	1.084	0.084	0. 931	0.001***	0.852	1.011
ATTITUDE TOWARD VIOLENCE						
To defend oneself (Item 1)	1.000	0.000	0.806	0.001***	0.771	0.840
to defend someone else (Item 2)	1.016	0. 033	0. 818	0.001***	0.788	0.844
To answer toward insults (Item 3)	0.867	0.026	0. 698	0.001***	0.655	0.731
To answer if someone is made fun of because of their religion (Item 4)	0.768	0.030	0. 619	0.001***	0.578	0.660
To defend one's country (Item 5)	0.826	0. 028	0. 665	0.001***	0.629	0.701
To fight back against racism (Item 6)	0.897	0. 028	0. 723	0.001***	0.689	0.757
To fight for one's ideas (Item 7)	0.740	0. 033	0. 596	0.001***	0.551	0.642
TRUST IN DISTANT INSTITUTIONS						
Parliament (Item 1)	1.000	0.000	0.838	0.001***	0.823	0.853
Prime minister (Item 2)	1.109	0.009	0.929	0.001***	0.921	0.938
President of the republic (Item 3)	1.062	0.008	0.890	0.001***	0.879	0.900
European Union (Item 4)	1.128	0.009	0.946	0.001***	0.938	0.953
United nations (Item 5)	1.109	0.010	0.930	0.001***	0.921	0.938
TRUST IN PROXIMAL INSTITUTIONS						
Parents (Item 1)	1.000	0.000	0.304	0.001***	0.219	0.388
Local political (Item 2)	2.086	0.305	0.633	0.001***	0.590	0.677
Schools (Item 3)	2.701	0.384	0.820	0.001***	0.797	0.844
Police (Item 4)	2.760	0.394	0.838	0.001***	0.816	0.861
Justice (Item 5)	2.263	0.323	0.687	0.001***	0.657	0.717
BELONGING TO DYG						
Do particular things to be admitted into the group (Item 1)	1.000	0.000	0.567	0.001***	0.435	0.690
Tolerance and acceptance in doing prohibited things (by the law) by the group (Item 2)	1.603	0.186	0.908	0.001***	0.859	0.957
Group members do illegal things (i.e., things that were prohibited by the law) (Item 3)	1.646	0.189	0.932	0.001***	0.883	0.981
NUMBER OF FRIENDS						
At school (Item 1)	1.000	0.000	0. 669	0.001***	0.623	0.714
In their neighborhood (Item 2)	1.365	0.080	0. 803	0.001***	0.755	0.850
In their town (Item 3)	1.204	0.078	0.655	0.001***	0.604	0.706
RACISM						
Some races are superior to others (Item 1)	1.000	0.000	0.415	0.001***	0.181	0.649
Racism is caused by social intolerance—reversed item (Item 2)	0.782	0.431	0.663	0.001***	0.297	1.028
TIME ONLINE						
Time during the week (Item 1)	1.000	0.000	0.838	0.001***	0.776	0.900
Time during the week-end (Item 2)	0.744	0.067	0.675	0.001***	0.619	0.731
APPLICATIONS						
Youtube	1.000	0.000	0.292	0.001***	0.176	0.408
Snapchat	2.613	0.562	0.763	0.001***	0.678	0.848
Facebook	1.990	0.427	0.581	0.001***	0.503	0.660

***p < 0.001.

When measuring the participant's attitudes toward racism, results showed that 47% of the perpetrators consider racism not to be justifiable ($n = 63$), while 30% agreed it is justifiable sometimes. Finally, 22.5 % ($n = 30$) of the perpetrators considered racism to be often justifiable. This suggests that a slight majority of cyberhate producers and disseminators consider that racism is justifiable [$Chi^2_{(2)} = 12.917, p = 0.001$].

Finally, when reporting their tendency to use specific apps, the vast majority of perpetrators reported using YouTube [$n =$

130, 89%, $\text{Chi}^2_{(1)} = 89.01$, $p < 0.001$], followed by Snapshot [$n = 107$, 73%, $\text{Chi}^2_{(1)} = 31.67$, $p < 0.001$] and Facebook [$n = 92$, 63%, $\text{Chi}^2_{(1)} = 9.89$, $p = 0.001$]. Regarding the remaining applications (i.e., Instagram, Viber, Messenger, WhatsApp, Twitter and online games), participants either did not use them [for example, 140 perpetrators did not use Viber, 95%, $\text{Chi}^2_{(1)} = 122.9$, $p < 0.001$ don't use Viber] or there were similar proportions of people who used them and did not use them (for example, 83 vs. 63 perpetrators used Instagram (56.8 vs. 43.1%, $\text{Chi}^2_{(1)} = 2.74$, $p = 0.09$].

Structural Equation Model

The model provided an acceptable fit (RMSEA = 0.049, CFI = 0.970, TLI = 0.965), and descriptive statistics for the variables kept in the model are reported in **Table 4** (mean, standard deviation, median which is especially relevant for our categorical variables, skewness, kurtosis). Regression coefficients are reported in **Table 5**, factor loadings in **Table 6**, and correlation between latent factors in **Table 7**.

Regression Coefficients

Regarding the model *per se* (**Table 5**, **Figure 1**), results show that producing cyberhate was significantly predicted by cyberhate victimization ($b = 0.414$, 95% CI = [0.299; 0.530], $p = 0.001$), revealing that the more participants reported being the target of cyberhate, the more they reported perpetration. Positive attitudes toward violence marginally predicted the tendency to be involved as a perpetrator ($b = 0.104$, 95% CI = [−0.008; 0.216], $p = 0.068$). Belonging to a DYG ($b = 0.406$, 95% CI = [0.236; 0.576], $p = 0.001$) and the time spent online ($b = 0.214$, 95% CI = [0.044; 0.383], $p = 0.014$) also significantly and positively predicted being a perpetrator. Surprisingly however, our results showed no significant link between perpetrators and (1) trust in proximal institutions ($b = −0.031$, 95% CI = [−0.200; 0.137], $p = 0.718$) nor with (2) distal institutions ($b = 0.012$, 95% CI = [−0.153; 0.177], $p = 0.888$), (3) the number of friends ($b = 0.052$, 95% CI = [−0.062; 0.166], $p = 0.369$) or with (4) the use of online applications ($b = −0.159$, 95% CI = [−0.0383; 0.065], $p = 0.163$).

Correlations Between Latent Factors

Regarding correlations between latent factors (**Table 7**), results reveal a significant and positive correlation between attitudes toward violence and confidence in distant institutions ($b = 0.080$, 95% CI = [0.026; 0.133], $p = 0.004$) but a negative correlation with proximal institutions ($b = −0.087$, 95% CI = [−0.143; −0.031], $p = 0.005$). Results also highlight a significant and positive correlation between attitude toward violence and the number of friends ($b = 0.264$, 95% CI = [0.206; 0.322], $p = 0.001$) and belonging to DYGs ($b = 0.432$, 95% CI = [0.366; 0.497], $p = 0.001$). Moreover, there was a significant and positive correlation between attitudes toward violence and the amount of time spent online ($b = 0.171$, 95% CI = [0.113; 0.228], $p = 0.001$) and the use of online applications (YouTube, Facebook and Snapshot) ($b = 0.064$, 95% CI = [0.007; 0.121], $p = 0.027$).

Trust in distant institutions was positively related to reliance in proximal institutions ($b = 0.585$, 95% CI = [0.552; 0.614],

TABLE 7 | Correlation coefficients between latent factors.

	Correlation	95% lower	95% upper	Significance
ATTITUDE TOWARD VIOLENCE				
Trust in distant institutions**	0.080	0.026	0.133	0.004
Trust in proximal institutions**	−0.087	−0.143	−0.031	0.002
Belonging to DYG***	0.432	0.366	0.497	0.001
Number of friends***	0.264	0.206	0.322	0.001
Cyberhate victimization	0.008	−0.073	0.088	0.850
Racism	0.012	−0.065	0.089	0.766
Time online***	0.171	0.113	0.228	0.001
Applications***	0.150	0.078	0.222	0.001
TRUST IN DISTANT INSTITUTIONS				
Trust in proximal institutions***	0.585	0.552	0.619	0.001
Belonging to DYG	0.023	−0.055	0.101	0.563
Number of friends**	0.087	0.025	0.148	0.006
Racism	0.021	−0.055	0.096	0.587
Cyberhate victimization	0.031	−0.055	0.118	0.481
Time online*	−0.064	−0.119	−0.009	0.023
Applications*	−0.075	−0.148	−0.015	0.041
TRUST IN PROXIMAL INSTITUTIONS				
Belonging to DYG***	−0.236	−0.309	−0.164	0.001
Number of friends	−0.032	−0.095	0.031	0.322
Racism	−0.054	−0.133	0.024	0.171
Cyberhate victimization	−0.074	−0.153	0.005	0.066
Time online***	−0.211	−0.266	−0.156	0.001
Applications***	−0.200	−0.272	−0.127	0.001
BELONGING TO DYG				
Number of friends***	0.312	0.241	0.383	0.001
Racism	−0.066	−0.172	0.041	0.226
Cyberhate victimization	0.075	0.043	0.247	0.005
Time online***	0.250	0.174	0.326	0.001
Applications***	0.357	0.261	0.453	0.001
NUMBER OF FRIENDS				
Racism	0.137	0.036	0.239	0.008
Cyberhate victimization	0.057	−0.027	0.140	0.185
Time online	0.089	0.028	0.150	0.005
Applications	0.297	0.219	0.375	0.001
RACISM				
Cyberhate victimization	0.053	−0.059	0.165	0.351
Time online	−0.020	−0.099	0.059	0.624
Applications	0.037	−0.963	0.139	0.477
TIME ONLINE*				
Cyberhate victimization***	0.267	0.180	0.354	0.001
Applications***	0.569	0.496	0.642	0.001
APPLICATIONS				
Cyberhate victimization	0.038	0.006	0.069	0.008

*** $p < 0.001$; ** $p < 0.01$, * $p < 0.05$.

$p = 0.001$) and to the number of friends ($b = 0.087$, 95% CI = [0.025; 0.148], $p = 0.006$). It was also negatively related to the time spent online ($b = −0.064$, 95% CI = [−0.119; −0.003], $p = 0.023$]
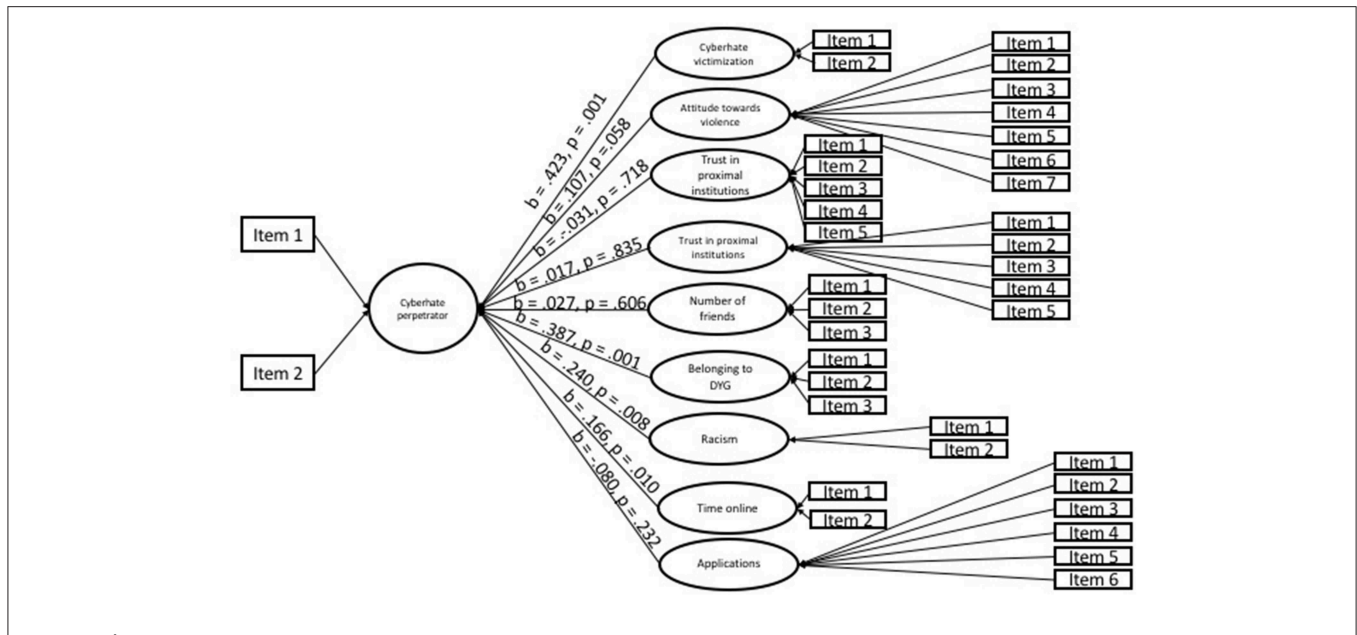
FIGURE 1 | Standardized regression estimates between latent factors.

TABLE 8 | Results of the mediation analysis.

	Standardized estimate	95% lower	95% upper	p-value
Total	−0.221	−0.325	−0.116	0.001***
Total indirect	−0.229	−0.303	−0.154	0.001***
Indirect via belonging to DYG	−0.182	−0.245	−0.129	0.001***
Indirect via attitude toward violence	−0.047	−0.069	−0.024	0.001***
Direct	0.008	−0.117	0.133	0.898

***p < 0.001; **p < 0.01, *p < 0.05.

and to the use of applications ($b = −0.073$, 95% CI = [−0.131; −0.015], $p = 0.014$).

Belonging to DYGs was positively correlated with the number of friends ($b = 0.312$, 95% CI = [0.242; 0.383], $p = 0.001$), cyberhate victimization ($b = 0.075$, 95% CI = [0.043; 0.247], $p = 0.005$), and negatively correlated with trust in proximal institutions ($b = −0.236$, 95% CI = [−0.308; −0.163], $p = 0.001$). It was also positively related to the time spent online ($b = 0.250$, 95% CI = [0.174; 0.326], $p = 0.001$) and to the use of applications ($b = 0.357$, 95% CI = 0.261; 0.453], $p = 0.001$).

The number of friends was positively correlated with racism ($b = 0.137$, 95% CI = [0.036; 0.239], $p = 0.008$), time spent online ($b = 0.089$, 95% CI = [0.028; 0.150], $p = 0.005$) and the use of applications ($b = 0.297$, 95% CI = [0.219; 0.375], $p = 0.001$). Finally, the time spent online was positively related to the use of applications ($b = 0.480$, 95% CI = [0.427; 0.533], $p = 0.001$). Cyberhate victimization was positively related with the time spent online ($b = 0.267$, 95% CI = [0.180; 0.354], $p = 0.001$), and with the use of application ($b = 0.038$, 95% CI = [0.006; 0.069], $p = 0.008$). Finally, time spent online was positively related with the use of applications ($b = 0.569$, 95%

CI = [0.180; 0.354], $p = 0.001$). No other significant correlation between latent factors was significant.

Multiple Mediation Analysis

Based on the observed results, we then sought to understand more clearly the mechanisms underlying the tendency to act as a cyberhate perpetrator. Notably, we wished to assess further if the level of trust in proximal institutions could actually predict acting as a cyberhate perpetrator. We hypothesized that the link between trust in proximal institutions and cyberhate perpetration could be mediated by (1) belonging to DYGs and (2) attitudes toward violence. In order to test these hypotheses, we performed a multiple mediation analysis in which we tested the direct and indirect link between trust in proximal institutions and cyberhate perpetration through belonging to a DYG and attitudes toward violence.

Results of the mediation analyses are reported in **Table 8** and depicted in **Figure 2**. The model provided an acceptable fit (RMSEA = 0.071, CFI = 0. 923, TLI = 0.908). Results reveal that there is a significant total effect of distrust in proximal institutions on cyberhate perpetration ($b = −0.221$, 95% CI = [−0.325; −0.116], $p < 0.001$). Individual mediation analyses revealed a negative indirect effect for both belonging to a DYG ($b = −0.182$, 95% CI = [−0.245; −0.129], $p < 0.001$) and attitudes toward violence ($b = −0.047$, 95% CI = [−0.069; −0.024], $p < 0.001$). However, as showed in **Figure 2** below, and confirming our previous analyses, no direct effect was found between the level of trust in proximal institutions and cyberhate perpetration ($b = 0.008$, 95% CI = [0.006; 0.069], $p > 0.05$). These results suggest that trust in institutions may indirectly prevent people from acting as cyberhate perpetrators since the more people trust their proximal institutions, the less they tend to become members of a DYG, and the more negative is their attitude toward violence.

FIGURE 2 | Multiple mediation analysis.
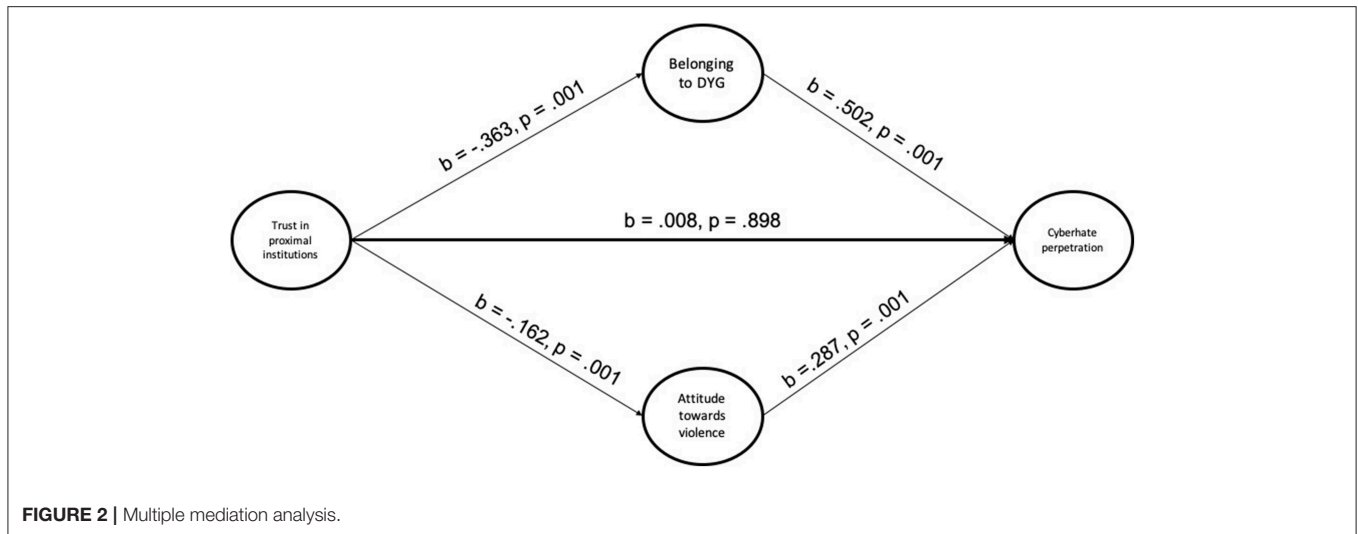
GENERAL DISCUSSION

This paper reports on a survey about the involvement of young people in cyberhate. It sets out to investigate the extent to which students are involved as perpetrators and to explore the characteristics of these young people. Descriptive findings show that out of 1,889 respondents to the survey, approximately one student in ten (10%) reported being victims of cyberhate and 5% that they had published or disseminated hateful content online. This percentage is much lower than the prevalence rates found by the survey by Costello and Hawdon (2018). There are probably contextual effects due to the cultural and legal differences as far as the freedom of expression and content regulations are concerned (Akdeniz, 2010). While in the United States, the First Amendment guarantees freedom of speech prevails, in Europe nations have made specific efforts to regulate cyberhate (Gagliardone et al., 2015). This is also probably due to the discrepancy in the age of the participants since the American survey included older participants (aged 15–36). Descriptive analyses show that male perpetrators are more numerous although females are also active (40%). As a whole, perpetrators are relatively happy with their lives. This contradicts research suggesting that young people adhering to extremist ideologies are unhappy and frustrated individuals (Khosrokahvar, 2014) and the fact that both parents work goes against the idea that extreme right youngsters are part of the "white trash" (Patricot, 2013). Apart from being a male, socio-demographics did not seem to be predictors of producing or disseminating cyberhate. A majority of perpetrators (70%) are actively involved in religion and spend longer time online as shown in previous research (Costello and Hawdon, 2018). This last point meets previous findings from research on cyberbullying, showing that the longer the time online, the higher the odds to become a victim. Exposure to cyberhate is associated with producing cyberhate as in the findings of Leung et al. (2018) for cyberbullying and Costello and Hawdon (2018).

While descriptive analyses suggest that perpetrators are more exposed than victims of cyberhate, regression analyses reveal that cyberhate perpetration is significantly predicted by cyberhate victimization. This goes along previous studies on bullying and "general" cyberbullying dedicated to the identification of the victims and aggressors' characteristics, showing that there is a strong overlap between victimization and perpetration (Vandebosch and Cleemput, 2009; Mishna et al., 2012). Perpetrators report more often being insulted at school than the other participants and thus being in a vulnerable position within their peer group. Having a group of friends usually is a testimony that one has positive social skills; it acts as a protective factor against victimization (Aoyama, 2010). However, some groups of friends are inadequate and have a negative influence on the way their members behave. As we could see in this survey, belonging to a DYG is positively correlated with having a high number of friends, contributes to a lower trust in the institutions and has a direct effect on producing or disseminating cyberhate. Latent factors analyses show that positive attitudes toward violence are correlated to belonging to a DYG and that they positively influence cyberhate behaviors. These attitudes are justified as a mean to defend oneself or someone else. As we could check with the SEM, cyberhate is strongly linked with being a victim of cyberhate, which might explain positive attitudes toward violence as a way to defend oneself or to defend others. Cyberhate and belonging to a DYG go along with the assumption that violence is a social construct and adolescents who associate with deviant or antisocial peers are more at risk to be involved in such behaviors themselves (Rugg, 2013), seeking peer acceptance or being under domination. The mere perception that deviance, delinquency or violence are accepted within the group can lead to the adoption of such behaviors (Petraitis et al., 1995). As Huang et al. (2014) show, peers' online risky behaviors is a risk factor for adolescents who will end up acting in a similar fashion. These findings

lead us to think that cyberhaters are particularly vulnerable since their involvement is associated with school bullying and cyberhate victimization.

The TUI Foundation (2018) showed that low levels of trust toward institutions are common among youth. One third of the participants to this survey thought that a radical change is needed and 7–23% showed populist attitudes. Quite a few youngsters experienced a feeling of loss of boundaries and that there was a demoralization process toward politics and institutions. This feeling affects the process of building stable identities and contributes to making them feel threatened, as stressed by (Rosa, 2017). Anomia leads some of them to adhere to populist ideologies. Discriminating and alienating others gives them a sense of order that they perceive, can only be achieved if they are hostile toward other groups (Schaafsma and Williams, 2012).

Perpetrators in our survey have lesser confidence in institutions than the other respondents. Latent factors analyses show that trust in institutions can act as a preventive factor as it prevents young people from belonging to deviant groups and decreases positive attitudes toward violence. We shall note that some perpetrators as well as young people repeatedly victimized reported low levels of trust toward school. This leads us to stress the urgency to reverse this situation and not only restrain the misuse of the Internet but also to promote attitudes toward tolerance and against racism and violence. As some previous studies show, teaching and fostering open discussions on racism, Islamophobia or anti-Semitism have positive incidence on students' attitudes toward racism and intolerance in general, both online and offline (Bergamaschi and Blaya, 2019). Schools can mediate intolerance through intercultural dialogue and education diversity. Although they cannot bear the responsibility to solve this societal problem on their own, schools can play an important role in counteracting hostile and abusive behaviors that stigmatize students and their communities both offline and online. As Foxman and Wolf (2013) argue, this would benefit not only this specific group but society as a whole as cyberhate affects the coherence of society, feeds hatred offline and spreads violent and extremist ideology.

LIMITATIONS

Some limitations of this study should be noted. The sample is a convenience sample, as we could only perform the study in schools who accepted to participate. Consequently, we cannot rely on a nationally representative sample. Questionnaires are self-reported, and answers are potentially biased like any survey of this type. However, it contributes to a better understanding of the characteristics of youth involved in cyberhate and thus can inform intervention against a phenomenon that can have heavy social consequences on witnesses, victims and perpetrators themselves.

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the supplementary files.

ETHICS STATEMENT

The survey was approved by the ethics committee of the Université Nice Sophia Antipolis and fulfills the requirements from the CNIL.

AUTHOR CONTRIBUTIONS

All authors made substantial contribution to the theoretical framework, design, data analyses or interpretation of this article.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211. doi: 10.1016/0749-5978(91)90020-T

Akdeniz, Y. (2010). *Racism on the Internet*. Strasbourg: Council of Europe Publishing.

Anti-Defamation League (2010). *Responding to Cyberhate, Toolkit for Action*. Available online at: https://www.adl.org/sites/default/files/documents/assets/pdf/combating-hate/ADL-Responding-to-Cyberhate-Toolkit.pdf (retrieved November 1, 2016).

Aoyama, I. (2010). Cyberbullying: What are the psychological profiles of bul-lies, victims, and bully-victims? *Diss. Abstr. Int. Section A: Humanit. Soc. Sci.* 71, 3526.

Bergamaschi, A., and Blaya, C. (2019). *Interactions Entre Jeunes: Lien Social, Relations Amicales et Confiance Envers les Institutions ?* Research Report, University Nice Côte D'azur/CAF, Nice.

Blaya, C. (2019). *Cyberhaine, Les Jeunes et la Violence sur Internet*. Paris: Nouveaux Mondes.

Blaya, C., and Gatti, U. (2010). Deviant youth groups in Italy and France: prevalence and characteristics. *Eur. J. Crim. Policy Res.* 16, 127–144. doi: 10.1007/s10610-010-9124-9

Boyd, D. (2014). *It's Complicated: The Social Lives of Networked Teens*. New Haven: Yale University Press.

Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research*. New York, NY: Guilford Publications.

Browne, M. W., and Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociol. Methods Res.* 21, 230–258. doi: 10.1177/0049124192021002005

Burt, C. H., Simons, R. L., and Gibbons, F. X. (2012). Racial discrimination, ethnic-racial socialization, and crime: a micro-sociological model of risk and resilience. *Am. Sociol. Rev.* 77, 648–677. doi: 10.1177/0003122412448648

Costello, M., and Hawdon, J. (2018). Who are the online extremists among us? Sociodemographic characteristics, social networking, and online experiences

of those who produce online hate materials. *Violence Gender* 5, 55–60. doi: 10.1089/vio.2017.0048

Costello, M., Hawdon, J., Ratliff, T., and Grantham, T. (2016). Who views online extremism? Individual attributes leading to exposure. *Comput. Human Behav.* 63, 311–320. doi: 10.1016/j.chb.2016. 05.033

Costello, M., Rukus, J., and Hawdon, J. (2018). We don't like your type around here: regional and residential differences in exposure to online hate material targeting sexuality. *Deviant Behav.* 40, 1–17. doi: 10.1080/01639625.2018.1426266

Foxman, A. H., and Wolf, C. (2013). *Viral Hate: Containing its Spread on the Internet*. New York, NY: Macmillan.

FRA (Fundamental Rights Agency) (2014). *Fundamental Rights: Challenges and Achievements in 2013*. FRA Annual Report. Office des Publications de l'Union Européenne.

Gagliardone, I., Gal, D., Alves, T., and Martinez, G. (2015). *Countering Online Hate Speech*. Unesco Publishing. Retrieved from: http://unesdoc.unesco.org/images/ 0023/002332/233231e.pdf

Hawdon, J. (2014). "Group violence revisited: Common themes across types of group violence," in *The Causes and Consequences of Group Violence: From Bullies to Terrorists*, eds J. Hawdon, J. Ryan, and M. Luch (Lanham, MD: Lexington Books), 241–254.

Hawdon, J., Oksanen, A., and Räsänen, P. (2017). Exposure to online hate in four nations: a cross-national consideration. *Deviant Behav.* 38, 254–266. doi: 10.1080/01639625.2016.1196985

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equation Model. Multidiscipl. J.* 6, 1–55. doi: 10.1080/1070551990 9540118

Huang, G. C., Unger, J. B., Soto, D., Fujimoto, K., Pentz, M. A., Jordan-Marsh, M., et al. (2014). Peer influences: the impact of online and offline friendship networks on adolescent smoking and alcohol use. *J. Adolesc. Health* 54, 508–514. doi: 10.1016/j.jadohealth.2013. 07.001

Keipi, T., Näsi, M., Oksanen, A., and Räsänen, P. (2017). *Online Hate and Harmful Content: Cross-National Perspectives*. Abingdon and New York: Routledge.

Khosrokhavar, F. (2014). *Radicalisations*. Paris: Editions de la Maison des Sciences de l'Homme.

Knobel, M. (2012). *L'Internet de la Haine: Racistes, Antisémites, Néonazis, Intégristes, Islamistes, Terroristes et Homophobes à l'assaut du Web*. Paris: Berg International.

Leets, L. (2002). Experiencing hate speech: perceptions and responses to anti-semitism and antigay speech. *J. Soc. Issues* 58, 341–361. doi: 10.1111/1540-4560.00264

Leung, A. N., Wong, N., and Farver, J. M. (2018). You are what you read: the belief systems of cyber-bystanders on social networking sites. *Front. Psychol.* 9:365. doi: 10.3389/fpsyg.2018.00365

Livingstone, S., Haddon, L., Görzig, A., Olafsson, K., and with members of the EU Kids Online Network. (2011). *EU Kids Online: Final Report*. London: EU Kids Online Network, London School of Economics and Political Science.

Llorent, V. J., Ortega-Ruiz, R., and Zych, I. (2016). Bullying and cyberbullying in minorities: are they more vulnerable than the majority group? *Front. Psychol.* 7:1507. doi: 10.3389/fpsyg.2016.01507

Mascheroni, G., and Ólafsson, K. (2014). *Net Children Go Mobile: Risks and Opportunities, 2nd Edn.* Milano: Educatt.

McDonald, R. P., and Marsh, H. W. (1990). Choosing a multivariate model: noncentrality and goodness of fit. *Psychol. Bull.* 107:247. doi: 10.1037/0033-2909.107.2.247

Messmer, E. (2009). Racism, hate, militancy sites proliferating via social networking: websense sees tripling of such active sites and "pockets" over last year. *Netw.* Retrieved from: http://www.networkworld.com/news/2009/ 052909-hate-sites.html (accessed September 20, 2018).

Mishna, F., Cook, C., Gadalla, T., Daciuk, J., and Solomon, S. (2010). Cyber bullying behaviors among middle and high school students. *Am. J. Orthopsychiatry* 80, 362–374. doi: 10.1111/j.1939-0025.2010. 01040.x

Mishna, F., Khoury-Kassabri, M., Gadalla, T., and Daciuk, J. (2012). Risk factors for involvement in cyber bullying: victims, bullies and bully–victims. *Child. Youth Serv. Rev.* 34, 63–70. doi: 10.1016/j.childyouth.2011.08.032

Näsi, M., Räsänen, P. Hawdon, J., Holkeri, E., and Oksanen, A. (2015). Exposure to online hate material and social trust among Finnish youth. *Information Tech. People* 28, 607–622. doi: 10.1108/ITP-09-2014-0198

Näsi, M. J., Räsänen, P., Keipi, T., and Oksanen, A. (2017). Trust and victimization: A cross-national comparison of Finland, the U.S., Germany and UK. *Res. Finn. Soc.* 10, 119–131.

Oksanen, A., Hawdon, J., Holkeri, E., Nasi, M., and Räsänen, P. (2014). "Exposure to online hate among young social media users," in *Soul of Society: A Focus on the Lives of Children and Youth*, ed M. N. Wareheim (Emerald: Bingley), 253–273.

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding From You*. London: Penguin Press.

Patricot, A. (2013). *Les petits Blancs: Un Voyage dans la France D'en Bas*. Paris: Plein jour.

Penzien, E. (2017). *Xenophobic and Racist Hate Crimes Surge in the European Union*. Human Rights Brief. Retrieved from: http://hrbrief.org/2017/02/ xenophobic-racist-hate-crimes-surge-european-union/ (accessed September 20, 2018).

Perry, B., and Olsson, P. (2009). Cyberhate: the globalization of hate. *Inf. Commun. Technol. Law* 18, 185–199. doi: 10.1080/13600830902814984

Petraitis, J., Flay, B. R., and Miller, T. Q. (1995). Reviewing theories of adolescent substance use: organizing pieces in the puzzle. *Soc. Psychol. Bull.* 117, 67–86. doi: 10.1037/0033-2909.117.1.67

Potok, M. (2015). *The Year in Hate and Extremism*. Intelligence Report 151. Retrieved from: https://www.splcenter.org/intelligence-report (accessed October 21, 2015).

Rosa, H. (2017). *Beschleunigung: Die Veränderung der Zeitstrukturen in der Moderne*. Berlin: Suhrkamp Verlag.

Rugg, M. (2013). Teenage smoking behaviour influenced by friends and parents smoking habits. *Eurek Alert J. Adolesc. Health* 143, 120–125.

Schaafsma, J., and Williams, K. D. (2012). Exclusion, intergroup hostility, and religious fundamentalism. *J. Exp. Soc. Psychol.* 48, 829–837. doi: 10.1016/j.jesp.2012.02.015

Schils, N., and Pauwels, L. (2014). Explaining violent extremism for subgroups by gender and immigrant background, using SAT as a framework. *J. Strateg. Secur.* 7, 27–47. doi: 10.5038/1944-0472.7.3.2

Sherif, M., and Sherif, C. W. (eds.). (1969). "Ingroup and intergroup relations: experimental analysis," in *Social Psychology*, (New York, NY: Harper & Row), 221–266.

Taylor, A. (2001). Young Canadians in a wired world: how Canadian kids are using the internet. *Educ. Canada* 41, 32–35.

Tucker, L. R., and Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 38, 1–10.

TUI Foundation (2018). *Young Europe 2018 – The Youth Study of TUI Foundation*. Retrieved from: https://www.tui-stiftung.de/en/ (retrieved January 15, 2019)

Tynes, B. (2005). "Children, adolescents and the culture of online hate," in *Handbook of Children, Culture and Violence*, eds N. E. Dowd, D. G. Singer, and R. F. Wilson (Thousand Oaks, CA: Sage), 267–289. doi: 10.4135/9781412976060.n14

Tynes, B., Del Toro, J., and Lozada, T. (2015). An unwelcomed digital visitor in the classroom: the longitudinal impact of online racial discrimination on academic motivation. *Sch. Psychol. Rev.* 44, 407–424. doi: 10.17105/SPR-15-0095.1

Vandebosch, H., and Cleemput, K. V. (2009). Cyberbullying among youngsters: profiles of bullies and victims. *New Media Soc.* 11, 1349–1371. doi: 10.1177/1461444809341263

Vitoroulis, I., and Vaillancourt, T. (2015). Meta-analytic results of ethnic group differences in peer victimization. *Aggress. Behav.* 41, 149–170. doi: 10.1002/ab.21564