

Moral struggles in social media discussion: the case of sexist aggression.

Marinella Paciello
Psychology Faculty
Uninettuno University
Rome, Italy

m.paciello@uninettunouniversity.net

Francesca D'Errico
Fil.Co.Spe Department
Roma Tre University
Rome, Italy
francesca.derrico@uniroma3.it

Giorgia Saleri
Psychology Faculty
Uninettuno University
Rome, Italy
g.saleri@students.uninettunouniversity.net

Abstract—The online interactions have become increasingly negative and aggressive, in particular when stereotyped topics are discussed. Based on extensive literature on aggression, the present study investigates the forms, the frequencies of the negative comments and the related both emotive and cognitive processes in response an offensive post toward Carola Rackete. A total of 1.249 comments on Twitter were analyzed by coding basic position toward the tweet, moral argumentation, underlying moral process, emotions and their intensity. Overall the results showed that online communication is largely denoted by negative tone and aggressive forms, even when a commenter would support the target of an aggressive communication. In particular, the gender aggression is carried out by the same women who became aggressive and much less prosocial than what one would expect

Keywords—discrimination, sexist aggression, prosocial, emotion, cognitive processes

I. INTRODUCTION

Over the last years the online interactions have become increasingly negative and aggressive (Pew Research Center, 2017). It is always more frequent to observe negative online phenomena such as hate speech, flaming and trolls in particular related to sensible ethical topics that can polarize the public opinion. For instance, some of the most online discussed topics concern immigration and stereotypes that could lead to explicit forms of cyber-racism and discriminations [1; 2]. Moreover, some groups are most likely to become target of online aggression, as in the case of girls and woman [3].

If this framework is typically applied to racists discussions, recent studies have also shown that regardless the ethical positions on discussions on ethical topics can be characterized by a high degree of hostile emotions [4]. Specifically, also people who support potential victims of discriminatory and offensive can react with the same aggressive modalities. Moreover, the spread of negative online emotion is strictly related to cognitive processes that maintain and exacerbate the circle of online incivility, such as in the case of dehumanization and attribution of blame to potential victims for their situation [4; 5]. Despite this form

of aggression is relatively new, the phenomena of online aggression can be understood based on the traditional models developed in the extensive literature on aggression.

First based on social cognitive theory [6], the hypothesis such as the presence of environmental cues that can increase aggression through priming mechanisms [7; 8] and more specific literature on media and aggression [9], the online aggressive phenomena can be understood if we consider the characteristics of online environment. The online interactions occur in a context in which some technological affordance hinder the possibility to empathize with a potential victim (e.g. lack of emotional cues) and the sense of personal responsibility with respect one's own action (e.g. anonymity), and at same time foster the toxic online disinhibition [10]. Moreover, the selective media attention on aggressive situations that could backfire by providing an advantageous comparison point to reduce the severity perception of one's own online aggressive behaviors [11]. In these situations, aggressive behavior can be trigger and aggressive social modeling can exacerbate negative online manifestations [11; 12]. Indeed, by observational learning processes the exposure to violence leads to the acquisition of cognitive beliefs (es. aggressive behavioral scripts) and processes (e.g. cognitive desensitization" to violence) that facilitate the adoption of aggressive behaviour. Secondly, based on the F-A hypothesis [13], the transfer of excitation [14], and the dissipation of aggression over time [15], the exposure to violence can be considered a case of frustrating condition that increase negative emotional arousal encouraging reactive aggressive response. Hostile emotions such as distress, frustration, anger, contempt, disgust and hate potentially promote aggressive behaviours [15; 16] because they feature medium/high arousal that, in the presence of other aversive stimuli, can progressively increase. In addition, hostile emotions, such as anger, could increase the attention towards provoking events and hostile interpretation of neutral situations [17]. Third, aggression can be resulted by dysfunctional coping strategies [18; 19]. The strategies used by defenders of potential victims range from an empathic support victim to angry attack toward a

perpetrator [20]. Thus, also in the case a prosocial position, it is possible to find hostile emotions, in particular, the so-called third-party anger [21; 22] or empathetic anger [23]. This type of anger is caused not by personal concern but by the desire to avoid injustice and restore social equity.

Moreover, the study of stimulus-thought-emotion-reaction sequence for helping versus aggressive response could be particularly interesting in the case of stereotyped topic, as in the case of woman. Several studies have attested that member of stereotypes and prejudice can become more easily target of hate-motivated aggression [3; 24]. In this case the aggression, usually acted by multiple perpetrators, has been motivated not only by the intent to hurt one victim but all members of victim's social category. Previous findings stressed the importance of considering social category as well the gender of target and agent of discriminatory and stereotyped forms of aggression [25]. Literature has suggested that the online verbal aggression against girl and woman is related to primary intention of intimidate and control them [3; 26], especially in the case of traditional 'male domains' such as publicly relevant contexts. Based on these premises, the present research questions are:

1. which are the emotional features of prosocial and proself stance toward an exhibition of a clear attack to woman image?
2. Are the prosocial stances toward a woman public aggression featured by a third-party anger?

Specifically, in the present study we investigate the forms, the frequencies of the negative comments in response to a post, and one of its comments, that intends to declare/exhibit an offensive message toward a specific woman who actively defends the right of a minority group (immigrants), and the women more in general. Based on above theoretical approaches and empirical studies, we expected to find aggressive comments also from commenters in defense of the attacked woman. We expected that both prosocial and proself stance will be featured by high levels of negative emotional activation.

II. METHOD

A. Online Communicative Scenario

The online communicative scenario relevant to this study is related to a message posted by Selvaggia Lucarelli, a popular Italian journalist and TV commentator, on July 2nd, 2019. The tweet was intended to condemn cruel and rude criticisms among women after the Sea-Watch case: at the end of June, the civil rescue vessel Sea-Watch 3, captained by Carola Rackete, entered the port of Lampedusa without permission with 42 migrants on board. (<https://www.theguardian.com/world/2019/jun/29/sea-watch-captain-carola-rackete-arrested-italian-blockade>) These facts have sparked heated debates in public opinion. The tweet was composed of text and an image. The image was a screenshot representing a dark humor meme shared by

another woman on Facebook to insult Carola Rackete by making assumptions about her sexual habits. On the contrary, the text expressed indignation towards use of sexist hate speech online (with the aim of stigmatize this sexist conduct):

'Donne che se la ridono condividendo 'sta roba. Ho esaurito le parole.'

'Women who laugh at sharing this stuff. I've run out of words.'

B. Data and Procedure

A total of 1,249 comments made in response to Selvaggia Lucarelli's post, distributed from the 2nd to the 3rd of July 2019, were extracted through Twitter API. Authors can privately share extracted data, by guaranteeing commenters' anonymity, in the interest of repeatability of analysis. All comments were manually filtered in order to better understand users' point of view and exclude tweets that no longer exist. First of all, coders identified a level of agreement or disagreement with Lucarelli's statement. In addition to this coding, the valence and intensity of emotions were examined considering both semantic and formal elements. Data coding carried out in this way led to the identification of basic position toward the tweet (pro and contro), moral argumentation/reasoning (aggressive, proself and prosocial), underlying moral process based on Bandura and Darley and Latané models [27; 28] (see Table 1) and four emotions (joy, calm, anger and sadness) and their intensity (low vs high). Another control element in this investigation was the presence or absence of gender stereotypes in analyzed tweets.

TABLE I. PROSOCIAL AND PROSELF ARGUMENTS

Loci	Cognitive Processes	
	Prosocial Processes	Proself Processes
Victim	Humanization	Dehumanization
	Value Attribution	Blame Attribution
Behaviour	Principle Recalling	Moral Justification
	Realistic Labelling	Euphemistic Labelling
Consequences	Structured Argumentation	Advantageous Comparison
	Positive Consequences	
	Disconfirm of Negative Consequences	Consequences Distortion (of hosting)
Agency	Exaggerate Negative Consequences	
	Agentic Trigger	Diffusion of Responsibility
	Assumption of Responsibility	Displacement of Responsibility

III. RESULTS

A Chi Square analysis [$\chi^2(1296) = 87,4; p < .001$] pointed out how commenters who openly support the main tweet are featured by negative emotions, mainly anger (49,3%) and sadness (35,1%), while commenters against the main tweet

express their positions in an angry (66,5%) or happy way (19.9%).

TABLE II. POSITION TOWARD THE MAIN TWEET*EMOTIONS

	Emotions			
	<i>Anger</i>	<i>Sadness</i>	<i>Joy</i>	<i>Calm</i>
Pro	49.3%	35.1%	12.8%	2.8%
Against	66.5%	6.8%	19.9%	6.8%
Neutral	16.7%	50.0%	16.7%	16.7%

Rarely commenters in support of the journalist Lucarelli expressed their opinion by using a prosocial argumentation but they are mainly aggressive (63,5%); while people against the act of public stigmatization express their comment with a proself (56,3%) or aggressive act (33,5%) but in a very less extent compared to the pro position [$\chi^2(1296)= 162$; $p < .001$].

TABLE III. POSITION TOWARD THE MAIN TWEET*MORAL STANCE

	Moral Stance		
	<i>Aggressive</i>	<i>Proself</i>	<i>Prosocial</i>
Pro	63.8%	17.6%	18.6%
Against	33.5%	56.3%	10.2%
Neutral	57.5%	25.4%	17.0%

In particular when we consider the processes emerging from the commenters words we unexpectedly found that people who support Lucarelli act of public indignation for the sexist tweet report a greater percentage of dehumanization (26%)- more than the unsupportive commenters (9,9%) - and guilt attribution (21,8%) toward the woman who make fun of Carola. In this sense the woman became another victim in the 'pro' commenters. Another frequent process activated by the Lucarelli supporters is the responsibility displacement (10,7%), when commenters tend to attribute the responsibility of this 'cultural level' to politicians as in the case 'Il fatto che il ministro del food blogging non condanni mai questa violenza verbale e scritta, di fatto, la sdogana'; on the other side commenters in favour of Lucarelli assume to themselves the shame of belonging to the female genre (Mi vergogno di essere donna...': *I am ashamed of being a woman*) A minimal part of the coded processes are devoted to principle recalling, women who ask if it is possible to denounce. Differently the commenters against the Lucarelli public act are mainly focused to the guilt attribution to the journalist or to Carola (25,7%) or to labelling euphemistically the indignation act as a political strategy toward the vice minister Matteo Salvini or comparing to the deviant act of the woman to the Carola 'aggression' to Italian law (Ma che cosa stai dicendo??? Ora speronare una

nave della guardia di finanza è un'azione giustificata; Transl: 'But what are you saying ??? Now ramming a finance guard ship is justified'). The difference between pro and against position are significant to the chi square analysis [$\chi^2(697)= 164$; $p < .001$]. The difference between Pro and Against the denouncing tweet of Lucarelli is very close to the gender difference, in the sense that women tend to dehumanize or attribute guilt or on the other hand assuming their responsibilities' or recall to a legal principle.

TABLE IV. PRO VS AGAINST POSITION*TYPE OF PROCESSES

Proself Processes	<i>Pro</i>	<i>Against</i>
	Dehumanization	26.0%
Blame Attribution	21.8%	25.7%
Advantageous Comparison	0.8%	12.9%
Diffusion of Responsibility	4.6%	2.9%
Displacement of Responsibility	10.7%	6.4%
Consequences Distortion	9.2%	7.6%
Euphemistic Labelling	2.7%	16.4%
Moral Justification	0.0%	1.2%
Prosocial Processes	<i>Pro</i>	<i>Against</i>
Principle Recalling	6.5%	5.8%
Structured Argumentation	4.2%	9.9%
Assumption of Responsibility	6.9%	0.0%
Value Attribution	0.6%	0.6%
Disconfirm of Negative Consequences	0.8%	0.0%
Exaggerate Negative Consequences	1.1%	0.0%
Realistic Labelling	1.7%	0.0%
Agentic Trigger	1.7%	0.0%
Humanization	1.0%	0.6%

When we consider the expressed emotions with respect to the moral stances we can report several significant differences [$\chi^2(1296)= 189$; $p < .001$] in the sense that the aggressive one as we could expect is featured by anger (59,8%) similarly to the proself one (49,1%) and differently from part of prosocial stance that is featured mainly by sadness (63,2%) but also anger (29,2%). But a tendency that characterized the aggressive and the proself is the presence of good percentages of joy (18,8% and 12,3% respectively) that signal how the aggression and not supportive toward women (against the aggression) is expressed by derision,

irony and ridicule that can be considered an indirect act of aggression [29].

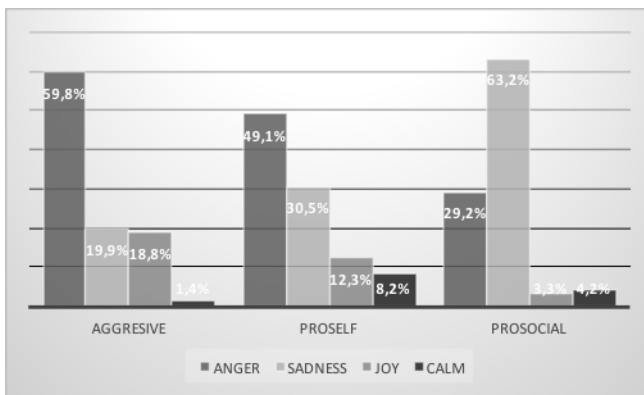


Fig. 1. Moral stance*Emotions

IV. DISCUSSION

The present study confirmed how online communication can be largely denoted by negative tone and aggressive forms in most of the case, even when a commenter would support the target of aggressive communication. The denounce of an aggressive behavior towards a victim only partly stimulates users' sorrow or other deactivated emotional states that could lead to a supporting scenario for the victim of the aggression (Carola Rackete). Instead both supporters and unsupporters of the journalist's denounce commented being mainly angry. Anger, that, as in other studies, was associated with a series of aggressive cognitions that feed the vicious cycle of aggression. In particular, we saw that users who support the sexist denounce on their side attribute the blame and dehumanize, and they are more aggressive and less prosocial than expectations. Pro sexists denounce people even get more enraged than against ones, in which, on the contrary, positive emotions emerged accompanied by acts of ridicule or derision, presumably deriving from the attempt to use humor to diminish and defend one's own social image [29; 30; 31]. The proself unexpectedly dehumanize less and mostly focus on behavior with 'euphemistic labeling' and 'advantageous comparison', dampening the force of the public denounce through lower emotional activation. On the social network the gender aggression is carried out by the same women who became aggressive and much less prosocial than what one would expect. The high occurrence of dehumanization mechanisms could be explained by the fact that the Lucarelli's post focused not only on sexist attack against Carola Rackete, but also on woman who made that attack. Thus, it is possible that dehumanization is also related to the fact that this specific woman represents a member of a "political" outgroup [32], that is the citizens sustaining the political forces against the hosting toward immigrants. The study confirmed that hate scenario can be spread on the web highlighting the affective and cognitive processes that can feed it. The reference to the real world is filtered by mechanisms of non-responsibility that

increasingly leave people "protected" behind their screen and in peace with their conscience despite the words written (aggressive, sarcastic, violent). It is conceivable that the modeling processes are working and the presence of pervasive aggressive cues support these dynamics that people might not necessarily be aware of. Future studies should focus on the possible moderators of these dynamics and the role of awareness in a process that seems above all supported by unruly emotions and processes that denote a loss of moral control.

REFERENCES

- [1] Navarrete, C. D., McDonald, M. M., Molina, L. E., & Sidanius, J. (2010). Prejudice at the nexus of race and gender: an outgroup male target hypothesis. *Journal of personality and social psychology*, 98(6), 933.
- [2] Faulkner, N., & Bliuc, A. M. (2016). 'It's okay to be racist': moral disengagement in online discussions of racist incidents in Australia. *Ethnic and Racial Studies*, 39(14), 2545-2563.
- [3] Pacilli, M. G., & Mannarini, T. (2019). Are women welcome on facebook? a study of facebook profiles of italian female and male public figures. *Tpm: Testing, Psychometrics, Methodology in Applied Psychology*, 26(2).
- [4] D'Errico, F., & Paciello, M. (2019). Online moral struggles in hosting immigrant's discourses: The underlying role of expressed anger and socio-cognitive processes. *Journal of Language Aggression and Conflict*, 7(2); 10.1075/jlac.00024.dar
- [5] D'Errico, F., & Paciello, M. (2018). Online moral disengagement and hostile emotions in discussions on hosting immigrants. *Internet Research*, 28(5), 1313-1335.
- [6] Bandura, A. (1978). Social learning theory of aggression. *Journal of communication*, 28(3), 12-29.
- [7] Engelhardt, C. R., & Bartholow, B. D. (2013). Effects of situational cues on aggressive behavior. *Social and Personality Psychology Compass*, 7(10), 762-774.
- [8] Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology*, 7(2p1), 202.
- [9] Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature. *Psychological science*, 12(5), 353-359.
- [10] Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in human behavior*, 28(2), 434-443.
- [11] Runions, K. C., & Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychology, Behavior, and Social Networking*, 18(7), 400-405.
- [12] Zimmerman, A. G., & Ybarra, G. J. (2016). Online aggression: The influences of anonymity and social modeling. *Psychology of Popular Media Culture*, 5(2), 181.
- [13] Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., & Sears, R. R. (1939). Frustration and aggression.
- [14] Zillmann, D. (1979). Hostility and aggression.
- [15] Konecni, V. J. (1975). The mediation of aggressive behavior: Arousal level versus anger and cognitive labeling. *Journal of Personality and Social Psychology*, 32(4), 706.
- [16] Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological bulletin*, 106(1), 59.
- [17] Zillmann, D. (1988). Cognition-excitation interdependences in aggressive behavior. *Aggressive Behavior*, 14(1), 51-64.
- [18] Averill, J. R. (1983). Studies on anger and aggression: implications for theories of emotion. *American psychologist*, 38(11), 1145.

- [19] Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American psychologist*, 46(8), 819.
- [20] Rudolph, U., Roesch, S., Greitemeyer, T., & Weiner, B. (2004). A meta-analytic review of help giving and aggression from an attributional perspective: Contributions to a general theory of motivation. *Cognition and emotion*, 18(6), 815-848.
- [21] Van Doorn, J., Zeelenberg, M., & Breugelmans, S. M. (2014). Anger and prosocial behavior. *Emotion Review*, 6(3), 261-268.
- [22] Vitaglione, G. D., & Barnett, M. A. (2003). Assessing a new dimension of empathy: Empathic anger as a predictor of helping and punishing desires. *Motivation and Emotion*, 27(4), 301-325.
- [23] Batson, C. D., Kennedy, C. L., Nord, L. A., Stocks, E. L., Fleming, D. Y. A., Marzette, C. M., ... & Zenger, T. (2007). Anger at unfairness: Is it moral outrage?. *European Journal of Social Psychology*, 37(6), 1272-1285.
- [24] Craig, K. M. (2002). Examining hate-motivated aggression: A review of the social psychological literature on hate crimes as a distinct form of aggression. *Aggression and Violent Behavior*, 7(1), 85-101.
- [25] Navarrete, C. D., McDonald, M. M., Molina, L. E., & Sidanius, J. (2010). Prejudice at the nexus of race and gender: an outgroup male target hypothesis. *Journal of personality and social psychology*, 98(6), 933.
- [26] Richardson-Self, L. (2018). Woman-Hating: On Misogyny, Sexism, and Hate Speech. *Hypatia*, 33(2), 256-272.
- [27] Bandura, Albert. 2015. *Moral Disengagement: How People Do Harm and Live with Themselves*. New York: Macmillan Higher Education.
- [28] Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of personality and social psychology*, 8(4p1), 377.
- [29] D'Errico, F., & Poggi, I. (2016). "The Bitter Laughter". When Parody Is a Moral and Affective Priming in Political Persuasion. *Frontiers in psychology*, 7, 1144.
- [30] D'Errico, F., & Poggi, I. (2012). Blame the opponent! Effects of multimodal discrediting moves in public debates. *Cognitive Computation*, 4(4), 460-476.
- [31] Hodson, G., & MacInnis, C. C. (2016). Derogating humor as a delegitimization strategy in intergroup contexts. *Translational Issues in Psychological Science*, 2(1), 63.
- [32] Pacilli, M. G., Roccato, M., Pagliaro, S., & Russo, S. (2016). From political opponents to enemies? The role of perceived moral distance in the animalistic dehumanization of the political outgroup. *Group Processes & Intergroup Relations*, 19(3), 360-373.