Populism and polarization in the digital arena: Categorising and measuring political attacks on Twitter

Héctor García Benítez-D'Ávila (2553120)
University of Twente
MSc Communication Science
Main supervisor: Prof. Dr. M.D.T. de Jong
2nd Assessor: Dr. Shenja van der Graaf
Number of words: 10595

Content

Abstract

Purpose

In an increasingly polarised political environment with the rise of populism in Western democracies, some politicians use Twitter to attack their opponents with ad-hominem arguments. Ad-hominem arguments are those that, instead of addressing the point of the discussion, are directed towards personally attacking the other candidate. This study seeks to analyse the types of ad-hominem arguments used by politicians on Twitter while measuring and comparing the engagement of these tweets. The aim is to raise awareness of this issue and to demand counterstrategies and policies from social media platforms and political parties.

Method

To establish a categorisation of political attacks, a content analysis was performed on the 1378 tweets by the main six political candidates that ran for the elections for the Community of Madrid in 2021 that were tweeted during the campaign period. These tweets were coded and a final codebook was established determining 18 types of political personal attacks that were used during this campaign. Additionally, these categorised tweets were tested to look for significant differences in their engagement.

Results

In this study, a categorisation of political attacks was created via coding; 18 types of personal attacks were found to be used by politicians on Twitter. Regarding these attacks, the most used by politicians consists in highlighting their opposition's bad governance by calling them negligent or forcing them to change their point of view. However, when it comes to engagement, addressing others as violent or corrupted in regards to media generates the most responses and attacks directed to resting the credibility of the opponents generates the most RTs and Favs. Addressing others as radicals, anti-democratic, or trying to exclude them from the dialogue (for example, by proposing a cordon sanitaire) generates the highest engagement in comparison to the number of followers, measured through a construct called "interactor ratio". When comparing the engagement among different types of tweets, attacks showed significantly higher engagement than non-attacks for all Twitter metrics, while ad-hominem arguments showed only a significantly higher interactor ratio than non-ad-hominem attacks.

Conclusion

This study shows, by analysing a real case, that populist politicians use up to 18 different types of personal attacks to confront others. In regards to their engagement, the more personal attacks are, the more polarisation and engagement they generate on Twitter. This analysis should serve as a wake-up call for urgent action by political parties and social media platforms to put an end to this populist practice that damages the political system.

1. Introduction

Since the 1980s populist parties have risen in liberal democracies (Mudde, 2004) and have experienced a major boost thanks to the development of the Internet and social media (Carrella, 2020) where politicians are able to directly influence their public (Grosse et al., 2012). When it comes to political communication, one of the most relevant social media platforms is Twitter (Bode & Dalrymple, 2016). Twitter provides a new debate arena for the masses, a digital space that reflects public opinion (Moya Sánchez & Herrera Damas, 2015) which makes it the perfect place for discussing political issues. On this platform, politicians gather more followers than official media accounts (Tong & Zuo, 2020) turning them into one of the main sources of information that replicates throughout the tweetlines.

The nature of Twitter is based on users engaging with each other and broadcasting the information they align with, turning this social media into an "echo chamber" (Bail et al., 2018). The "echo chamber" nature of Twitter is key to understanding the political charge of this social media; Twitter allows for users to interact and respond each other in conversations (Suh et al., 2010), so, in the political scope, users interact mostly with others that they politically align with (Grosse et al., 2012). Populist politicians are part of this digital space, as they use social media to communicate with their followers during campaigns (Lorentzen, 2020) and to address other politicians in their populist strategies (Bruns & Highfield, 2013). Using Twitter as their main asset to replicate information fast through their followership, politicians have influenced the political system with historical happenings such as the Brexit (Flew & Iosifidis, 2020) and the Capitol incident (Twitter Inc., 2021).

All these factors turn Twitter into the selected channel in this study to analyse populist political attacks and the users' reactions to their tweets. Many other studies focusing on political social media communication rely on using Twitter and tweets as the main source of data. Social media use by populist politicians has been addressed by analysing tweet engagement and populist features such as negativity of the message (Carrella, 2020), the use of social media to troll and attack others (Bulut & Yörük, 2017) and analyses of Twitter as a populist tool to target opposition (van Kessel & Castelein, 2016) as well as other studies regarding populist messages, polarisation and populist leaders in social media (Alonso-Muñoz, 2019; Lorentzen, 2020; Primario et al., 2017; Suau-Gomila et al., 2020; Tong & Zuo, 2020; Usher et al., 2019). Other studies have also analysed how users react to polarising or controversial tweets, either from a "cancel culture" perspective (Sailofsky, 2021) or a boycott perspective (Bogen et al., 2021). However, the "cancel" movement relates to messages including social injustices, racism awareness or "wokeness" (Vredenburg et al., 2020) rather than to messages including personal attacks.

It is known that populist politicians have used personal arguments in their campaigns often achieving better results and higher votes (Borovali, 2018; Sobieraj & Berry, 2011) but their typology and their impact on social media engagement remains rather unknown. Therefore, this study has two theoretical goals: to extend the

literature's typology of political personal attacks used in tweets during campaigns and to measure and compare the reactions of users to these categories. To do so, this study seeks to answer the next research questions, in a qualitative and quantitative scope respectively: the first, "what are the types of ad-hominem arguments that political candidates use on Twitter during a campaign?" and the second, "how do Twitter users interact with politicians' attacks and ad-hominem arguments in a campaign environment?".

2. Theoretical framework

2.1. Populism and polarization

According to the philosopher Ernest Laclau (1996), populist leaders utilize a main communicational pillar to captivate the masses: the popular self-identification with the leader or the movement based on the definition of a common antagonist (Cornelissen, 2019). This separates the social system into two different groups, a privileged minority and an oppressed majority, in a simple dichotomy that populist leaders can easily communicate (Kioupkiolis, 2016). Elaborating on Laclau's ideas some authors state that key references to the enemies are not necessary for the populist discourse to take place, as the broad majority, "the people", is the most important element (Engesser et al., 2017). There are some branches of populism that do not even define an antagonist (Jagers & Walgrave, 2007). "The people" is firstly formed by identification without the need of an exterior figure, although it will remain incomplete until the group identity is glued when the leader appears (Kingsbury, 2016). Thus, the populist leader plays an important role as the messenger of the party, usually portraying charisma (Engesser et al., 2017) and the degree of populism in his political party will mostly depend on his discourse (Bernhard, 2020).

Some populist politicians rely on confrontation to construct a group identity. According to Engesser et al. (2017) one of the techniques attributed to populism is polarization (Bucy et al., 2020). Populist politicians usually define two enemies of the people: the controlling elites, and the so-called "others", that is, those who are privileged by the elites and belong to an out-group (Engesser et al., 2017) outside of the common identification (in-group) that the people have with the populist leader. This out-group is normally a part of society, which gives a visible enemy for the people, as the powerful elites may be described too abstractly and broadly (Jagers & Walgrave, 2007). This victimization against the elites is another resource of populist politicians, as it gives provides a moral legitimacy to the political cause (Birks, 2011). The element of confrontation is the key to the self-identification with the leader (Cornelissen, 2019), and as a part of the same team we could expect the followers of the leader (the in-group) to act as the leader justifies, even on radical or violent stances (Golec de Zavala & Keenan, 2021). With the paradigm of social media, polarization and populism can be seen on a mass scale; as anonymous attacks are a possibility, populist politicians use their discourse to entitle trolls to polarize the digital space (Bulut & Yörük, 2017). The political sphere is aware of populist practices and may address others as populists; when

a politician does this, it is often done to attribute insincerity or non-professional behaviour, in order to reduce their level of trust (Hamo et al., 2019).

2.2. Digital populism

Over the last years, populism has raised in many liberal democracies due to the usage of social media in what could be considered populism 2.0 (Tong & Zuo, 2020). Through the Internet, political leaders can communicate more directly with their followers, unlike in traditional media, where the possibilities of communicating freely were much more limited (Alonso-Muñoz & Casero-Ripollés, 2018). Twitter is a very valuable tool for populist politicians (Arroyas Langa & Pérez Díaz, 2016) as Twitter users are more interested in politics and less trusting of mainstream media (Bode & Dalrymple, 2016); it can be seen that popular politicians may even gather more followers in their accounts than people following mainstream media accounts (Tong & Zuo, 2020). The nature of Twitter benefits homogenous political interaction, as due to its system of engagement, messages are replicated among people that follow each other, which turns Twitter into an echo chamber (Bail et al., 2018). The echo chamber theory states that users are most exposed to consonant views and supportive networks, which happens in the case of Twitter (Vaccari et al., 2016). However, some authors avoid reducing Twitter as an echo chamber only, as heterogeneous interactions occur more often than popularly thought (Michailidou, 2017). Nevertheless, this exposure to politically opposed messages does not necessarily bring more understanding; findings show that it increases radicalization (Bail et al., 2018).

Regarding their tweets, most populist leaders stick to a “more is less” approach; they tend to tweet less, as engagement depends on the content of the tweet, and not on the number of tweets (Alonso-Muñoz & Casero-Ripollés, 2018). This content will vary on its aggressiveness depending on the degree of populism, as the more populist a party is, the more possibilities it has of adopting a discourse to attack the elites or out-groups (Jagers & Walgrave, 2007). According to findings, this polarization should generate a reaction from users; tweets including polarising content are linked to higher engagement in social media (Bulut & Yörük, 2017) which may explain that when politicians address others in Twitter it is often done to increase engagement (Lorentzen, 2020); just mentioning another candidate by their name and not using the symbol “@” increases the engagement of the tweet (Pancer & Poole, 2016). However, the content of the tweets is not qualitatively categorised in these studies, and other findings show that politicians may also mention others just speak about them (Bruns & Highfield, 2013) as not all mentions to other politicians have to be attacks nor have a populist nature (Conover et al., 2011).

Hence, these results could indicate that the nature of the content (attacking) is the differential factor increasing engagement. This would match with results showing that polarised attacks increase the engagement of tweets and that addressing other politicians is usually done to increase engagement (Bulut & Yörük, 2017; Conover et al., 2011; Lorentzen, 2020). Additionally, messages with criticism have been shown to generate positive effects and increase votes for politicians (Sobieraj & Berry, 2011).

When linking these results to Laclauist's theories of identification based on a common adversary, an engagement is expected by the users that follow the political leader. These reasonings lead to the proposal of the first hypothesis of this study, H1:

H1 - Attacking tweets generate more engagement than non-attacking tweets.

In Twitter, the acceptability of messages can be measured by a social voting system like Twitter metrics (retweets, favourites and responses) (Alsinet et al., 2017). Interactions have been used in other studies to design models to predict political behaviour on Twitter (Aragõn et al., 2013; Baviera et al., 2019; Cardaioli et al., 2020; Keller & Kleinen-von Königslöw, 2018; Khatua et al., 2020; Makazhanov et al., 2014; Tumasjan et al., 2010), predicting political reactions to advertising campaigns (Altoaimy, 2018; Bogen et al., 2021), and measuring polarisation (Lorentzen, 2020; Primario et al., 2017).

2.3. Ad-hominem arguments

The framework of abstract argumentation from Dung (1993) distinguishes that discussion proposal argumentations can be either justified or opposed, making it possible for other actors to be persuaded by identifying new information (Dragoni et al., 2018). The interactive process of argumentation is usually divided into the first step of evaluating the strength of arguments and then selecting the most acceptable arguments (Cayrol & Lagasquie-Schiex, 2005). The first step, evaluability of arguments, is produced by presenting concise, valuable information although it can also be promoted by persuasive techniques such as proposing dichotomic terms like "good" or "bad" (Carenini & Moore, 2006) as some ad-hominem arguments portray (Harris et al., 2012) following a dichotomic nature like the one that populist leaders communicate (Kioupkiolis, 2016). The second step, which is selecting the most acceptable arguments, may also be influenced by the use of ad-hominem, as it is easier to choose an ad-hominem argument to another of similar strength if there are time or informational constraints (Finn, 2019). That might be why, in politics, ad-hominem arguments are often used as a resource to discredit the opposition (Borovali, 2018), them being commonly used by populist politicians (van Kessel & Castelein, 2016).

For an argument to be considered ad-hominem, there has to be an attack on somebody (Kotzee, 2010) that aims to discredit the argument of the other party (Walton, 2000). Although ad-hominem arguments can be considered a fallacy, due to referring to the identity of the actor and not validating an argument (Orkibi, 2018), they can still provide new information that may be perceived as important (Harris et al., 2012) and they can influence decisions and even force one of the interlocutors to withdraw from his arguments (Macagno, 2013). According to Walton (2000), they should be only considered fallacious when the goal is to block the discussion between the two parts while Mizrahi (2010) considers the fallacy if the figure's behaviour does not align with the message he is criticising about the other party.

There are different types of ad-hominem argumentation according to literature; Borovali (2018) categorises ad-hominem arguments in 4 categories; the "you too"

argument, the "whataboutery" ("where were you when this happened?"), pointing out the bias of the opponent and the direct personal attack while Macagno (2013) creates another categorisation based on the roles of the discussion; Ad-hominem 1 (excluding the interlocutor from the dialogue), ad-hominem 2 (Forced persuasion to force the interlocutor to support a viewpoint), ad-hominem 3 (Argument to trigger negative judgement from others) or ad-hominem 4 (Diminish credibility of the source). Other authors have established new categories such as the Ad-Hitlerum by (Harris et al., 2012) "Hitler also did it, therefore this person should not be trusted". The lack of categorisations in literature in regards to a social media campaign environment motivated the qualitative goals of this study, that is, to identify existing and new categories of ad-hominem arguments used on political campaigns to then be inventoried.

Previous findings show that politicians use these arguments on Twitter, but their usage varies based on their gender and race, as well as the campaign environment; female candidates tend to attack more on tweets than male candidates, especially if there is a higher percentage of other female candidates in an election (Evans & Clark, 2016). The main reason why populist political leaders generate contexts of confrontation and outrage with personal attacks (van Eemeren et al., 2012) is to provoke an emotional response on citizens (Sobieraj & Berry, 2011). In polarised campaigns, ad-hominem argumentation increases polarisation and makes it easier for citizens to pick a side and support one of the parties (Borovali, 2018). Ad-hominem arguments can be considered as a more polarising element in social media (Macagno, 2013) than other types of attacks, and a stronger tool than "normal attacks" to point at adversaries (Sobieraj & Berry, 2011). Following Laclau's theories, more engagement is expected than in "normal attacking" tweets due to the increased polarisation element. Thus, the second hypothesis of this study, H2, is proposed:

H2 – Attacking tweets with ad-hominem arguments generate more engagement than attacks with non-ad-hominem arguments.

3. Method

A content analysis was performed to categorise tweets, and then the metrics of the tweets were collected. These metrics were introduced in the statistical program SPSS and a statistical test was performed on the categorised tweets to assess the difference of the engagement among the different types of tweets that emerged during the coding. The study was approved by the Ethics Committee of the University of Twente.

3.1. Corpus

The case study of the 2021 elections for the Community of Madrid is taken to analyse the metrics of ad-hominem political attacks on Twitter. This campaign has a clear date of start as it was provoked due to a failed impeachment on the 10th of March and a clear end on election day on the 4th of May, which will mark the period of tweet collection. Almost all of the parties with representation in 2019 managed representation

in elections of the case study of 2021, easing the choice of the candidates whose tweets will be analysed (Ángel Gabilondo for PSOE, Isabel Díaz Ayuso for PP, Mónica García for Más Madrid, Rocío Monasterio for Vox, Pablo Iglesias for Unidas Podemos and Edmundo Bal for Ciudadanos). These candidates were the main actors of the campaign, they were invited to debate on media on multiple occasions and all were generally considered suitable candidates for the creation of a government. For the tweet categorisation, the number of 2438 tweets were extracted from these accounts for this period. For the sake of more precise coding and re-coding, only the text contained within the tweets was analysed. This means that possible arguments on graphic material such as videos or pictures were not coded. Retweets, mentions and tweets containing only media were discarded, leading to 1378 tweets to analyse.

3.2. Codebook

Tweets were categorised from a starting set of types based on literature; the definition for attack and non-attack was taken from del Valle et al. (2020) while the definition of ad-hominem arguments was taken from Kotzee (2010). Four types of ad-hominem arguments were collected from Borovali (2018), Macagno (2013) and Harris et al. (2012). Although these categories served to code some tweets, there was a clear necessity of creating new ad-hominem categories with the hopes of extending the ad-hominem typology existing in current literature. During the coding, some tweets showed attacks that were not directed to other politicians but to other political parties. This was not contemplated in the original research, as the study is oriented towards personal attacks, but these tweets cannot be considered non-attacks, and therefore needed to be categorised and measured as a different type of attacks.

Following the general rule of thumb (Lavrakas, 2013), a random selection of 10% of tweets (more than 50 units) was re-coded by a second independent coder to assess for reliability in the coding. Therefore, 140 tweets were randomly selected to be re-coded. The intercoder, who was familiar with qualitative research, was given a brief instruction on the use of Atlas.ti and the codebook as a guide. Disagreements were revised after each round of coding and guidelines were set to re-code following an agreed criterion. The guidelines were: a party attack must include a party's name or very clear identifiable information over a specific party. The same goes for ad-hominem. Every attack must have an obvious negative connotation. General confrontational statements that are not directed to a clear individual or party, but to a group that implies the political opposition, are included as Attacks. Additionally, the categories Ad-hominem and Direct personal attack merged and the category of general Party Attack is removed. The final Codebook can be found in Appendix A.

After the first round of coding, Cohen's kappa for intercoder reliability was 0.56 for identifying attacks, 0.64 for identifying ad-hominem arguments, 0.35 for identifying families of ad-hominem arguments, 0.35 for identifying party attacks and 0.3 for identifying the arguments within messages. After the final round of recoding, the kappa for intercoder reliability was 0.8 for identifying attacks, 0.96 for identifying ad-

hominem arguments, 0.77 for identifying party attacks, 0.86 for identifying types of party attacks and 0.82 for identifying the arguments within messages.

3.3. Engagement measurements

To measure engagement the Twitter metrics (retweets, favourites and responses) were assessed separately and compared for differences among groups. These metrics represent the social voting system of Twitter (Alsinet et al., 2017), and each one has its nuances: retweets are used to disseminate information, although they can also be used to add or respond to a tweet. Favourites show support and also help to disseminate information (Yaqub et al., 2017), while responses can vary in their degree of support and can also portray a negative tone, especially against polarising statements (Bogen et al., 2021).

These metrics can be used to build other constructs such as the interactor ratio. The interactor ratio determines the reach of engagement of a tweet outside of the follower base (Anger & Kittl, 2011). This ratio is constructed with the formula:

$$\frac{Retweets + Favourites + Responses}{Followers *}$$

*The number of followers is determined at the end of the campaign date, May 4th 2021.

Tweets scoring for a high interactor ratio would therefore be engaged by users that are not following the politician's account, which is not always done positively. Users normally follow the politicians that they agree with, while they do not follow those with who they disagree, due to Twitter's "echo chamber" nature (Bail et al., 2018; Michailidou, 2017; Vaccari et al., 2016). Additionally, findings show that negative political messages trigger emotional responses from opposed citizens (Sobieraj & Berry, 2011). Based on the echo chamber theory, the interactor ratio, normally used to measure reach, is presumably the metric that carries the highest possibilities of portraying controversy or polarisation in the political context.

4. Results

4.1. Categories of arguments

The final codebook can be seen in Table 1. This categorisation collects all the different types of tweets and attacks that were found during the study. Descriptive data of these tweets is displayed in this table. As it can be expected, most of the tweets emitted by politicians are non-attacks, containing types of information regarding political rallies, television appearances or the call for political measures without a confrontational nature. Some examples of non-attacks are:

> "In some minutes I give an interview in @ondamadrid. 📻 You can listen to it here: https://telemadrid.es/emision-en-directo-ondamadrid" (Iglesias P., 2021, April 30).

> "Thankful for the neighbours of the Chamberí neighbourhood and our militants and sympathizers for their effort and dedication. Today and always." (Gabilondo A., 2021, April 12).

> "Spain is not a banana republic. #Freedom" (Díaz Ayuso I., 2021, April 26)

However, almost 40% of tweets contain at least one attack, and one out of four attacks consists of an ad-hominem argument. An example for a non-ad-hominem argument that still represents an attack is:

> "Only bad people single out, attack and focus their hate on vulnerable children" (García M., 2021, April 20).

From ad-hominem attacks, the most used type is the Negligent, consisting in 3% of the sample, more than double of the second most used type of ad-hominem, that is, Ad-hominem 2. Examples of Negligent ad-hominem are:

> "Living in Madrid alone is not difficult, it is difficult to do so while Ayuso and the PP rule" (García M., 2021, April 19).

> "Ayuso's inefficiency is a danger, let's not let her keep us away from hugs without a mask" (García M., 2021, April 2).

Although the majority of tweets represent a non-attack, 40% of the tweets include at least one attack. This is a striking percentage, as it means almost half of the sample is dedicated to confronting other politicians. Approximately one out of four attacks being an ad-hominem argument, conforming around 11% of the total sample, which leaves non-ad-hominem attacks at 29%, almost a third of the sample. Party attacks were coded for around 8% of the sample, making them the less common type of attacks. Attacks that were not either party attack or ad-hominem arguments shape 21% of the sample. These attacks do not address personal characteristics, but rather political criticisms.

The newly coded ad-hominem arguments result in much more specific content-wise than those described in previous literature. There are clear descriptions of the aim of the attack, while literature types portray much more ambiguity, as in the case of Macagno's (2013) categorisation. This specificity in the arguments may be induced by the short-message nature of Twitter. As there is not much space to elaborate, politicians focus on very specific attributes in their tweets. As it can be seen, most of the newly coded ad-hominem arguments focus mainly on undesirable political attributes of the opposition. They can address others by their incapacity of governance (e.g. Negligence, Corrupt, Reactionary) or described as having bad intentions when it comes to governing (e.g. Polarising, Radical, Violent, Liar), but there is a lack of usage of ad-hominem based on physical characteristics. It is possible that those kinds of attacks would be too polarising and possibly damage the image of the attacking politician and are therefore not used at all. The category of Direct personal attack, which is one of the poorest when it comes to its "argument power" as it discredits based on personal attributes only, is one of the lowest used arguments, which may be for that reason.

Table 1:

Codebook.

Categories	Name	Description	N	%
Non-attack		The tweet is a mention but it does not include criticism directed to another politician's actions (del Valle et al., 2020).	829	60
Attack		The tweet is a mention and includes some criticism directed at another politician's actions (del Valle et al., 2020). However, it is not considered an ad-hominem.	549	40
Party		The tweet is an attack directed at a political party or political organisation.	111	8
Ad-hominem		The tweet is an Attack on somebody's personal character (Kotzee, 2010).	150	11
	"You too"	The person is discredited due to violating the principles he is advocating (Borovali, 2018).	6	.4
	Direct personal attack	The person is discredited based on his personal attributes (Borovali, 2018).	4	.3
	Ad-hominem 1"Excluding from dialogue"	The attack tries to exclude the other person from the dialogue (Macagno, 2013).	9	.7
	Ad-hominem 2"Forcing viewpoint"	The other person is attacked for not supporting a viewpoint (Macagno, 2013).	18	1.3
	Ad-hominem 3"Negative judgment"	The attack tries to trigger negative judgement from other users (Macagno, 2013).	16	1.2
	Ad-hominem 4"Diminishing credibility"	The attack tries to diminish the credibility of the other person (Macagno, 2013)	3	.2
	Antidemocracy	The person is addressed as being against democratic practices and principles.	6	.4
	Copycat	The person is addressed as being a copy or ally of another political opponent.	17	1.2
	Inverted whataboutery	The person is discredited based on the convenience of not speaking up on an instance that is not aligned with their interests. This is the inverted argument of "whataboutery" by Borovali (2018)	5	.4
	Liar	The person is accused of lying or hiding relevant information.	10	.7
	Negligent	The person is addressed as a negligent politician, ignoring groups of people or applying incoherent or unnecessary policies.	43	3.1
	Corrupt	The person is addressed as being corrupt or taking part in corrupt activities.	8	.6
	Polarising	The person is addressed as a polarising politician or trying to polarise society with their messages.	5	.4
	Radical	The person is addressed as a radical or extremist.	11	.8
	Reactionary	The person is addressed as opposed to political, scientific or societal progress.	4	.3
	Violent	The person is addressed as violent or taking part in violent activities.	3	.2

Categories	Name	Description	N	%
	Populist	The person is addressed as a populist.	2	.2
	Media-corrupt	The person is accused of having relations with media sources or using media for their political benefit.	5	.4
Total			1378	100

4.2. Engagement

Table 2 shows the information related to the engagement of the coded sample. The upper bound of the 95% confidence interval for the means of Favs [3323.5, 6591] RTs [1027.3, 2131.5], responses [375.6, 689.4] and IR [.03, .04] is taken as a threshold to identify the highest scores. It can be seen that in all forms of engagement but interactor ratio, the types of Ad-hominem 4 "Diminishing credibility", Liar and Media-corrupt show the most scores over the upper bound of the CI. These three types of arguments have similar characteristics as they directly attack the credibility of the political opponent.

Additionally for Favs, Ad-hominem 1 "Excluding from dialogue", Anti-democracy and Violent also provide the highest scores. For responses, the higher scores are provided by radical. For the interactor ratio, Ad-hominem 1 "Excluding from dialogue", Antidemocracy and Radical provide the highest scores.

Table 2:

Information of categorised tweets.

Category	Name	Interaction type							
		Favs		*RTs*		*Responses*		*IR**	
		M	SD	M	SD	M	SD	M	SD
Non-attack		2990	181	868	50	280	20	.02	.00
Attack		4267	306	1429	93	518	48	.03	.00
Party		2829	361	1031	117	339	77	.02	.00
Ad-hominem		4408	469	1453	156	516	70	.03	.00
	"You too"	2676	1412	866	347	193	103	.03	.00
	Direct personal attack	4032	1378	1017	222	361	163	.04	.02
	1"Excluding from dialogue"	7289**	1981	1892	460	1149**	481	.05**	.02
	2"Forcing viewpoint"	3139	1079	956	339	306	131	.03	.01
	3"Negative judgment"	2628	757	1055	328	219	78	.02	.00
	4"Diminishing credibility"	13502**	3632	4396**	1021	698**	207	.04	.00
	Anti-democracy	8011**	4141	2664**	1380	437	210	.05**	.02
	Copycat	2324	470	791	161	341	65	.02	.00
	Inverted whataboutery	4044	1517	1134	371	341	126	.03	.01
	Liar	13505**	3336	4665**	1086	883**	282	.04	.01
	Negligent	3518	703	1124	228	534	155	.03	.01
	Corrupt	2193	708	956	325	299	149	.02	.01
	Polarising	1356	552	566	248	566	248	.02	.00

Category	Name	Interaction type							
		Favs		*RTs*		*Responses*		*IR**	
		M	SD	M	SD	M	SD	M	SD
	Radical	6253	1730	2042	618	704**	183	.05**	.02
	Reactionary	1778	316	576	94	340	143	.03	.00
	Violent	7217**	4053	1088	48	1273**	540	.04	.02
	Populist	438	79	208	43	42	25	.01	.00
	Media-corrupt	10660**	3582	3970**	1358	1375**	798	.04	.03

*Interactor ratio ($\frac{Retweets+Favourites+Mentions}{Followers}$)

**Scores higher than the upper bound of the 95% CI of favs [3323.5, 6591], RTs [1027.3, 2131.5], responses [375.6, 689.4] and IR [.03, .04]

Two tests were performed to test the two proposed hypotheses: the first one was used to measure the significant difference of engagement on attacking and non-attacking tweets (H1) and the second one was used on ad-hominem and non-ad-hominem attacks (H2). The dataset was tested for normality with Kolmogorov-Smirnov and homogeneity of variance was tested with Levene's test, failing to fulfil the assumptions of normality ($p > 0.05$) and homogeneity of variance ($p < 0.01$) required to perform a t-test. Therefore, the non-parametric Mann Whitney U test is chosen to perform this analysis.

The first Mann Whitney U test showed significant lower scores for non-attacks on FAVs, RTs, responses and interactor ratio than for attacks on the same metrics of FAVs, $U = 180957.5$, $z = -6.444$, $p < .001$, RTs, $U = 163607.5$, $z = -8.843$, $p < .001$, responses, $U = 168566$, $z = -8.158$, $p < .001$ and interactor ratio, $U = 168588.5$, $z = -8.154$, $p < .001$. These results, shown in Tables 3 and 4 (Appendix B), support the hypothesis H1 in regards to attacking tweets generating more engagement than non-attacking tweets. There is a clear inclination of attacking tweets to increase engagement and therefore the interactor ratio in a significant manner.

Tables 6 and 7 (Appendix B) display the results of the test for H2, where the same effect is contemplated, with a significant difference between ad-hominem and non-ad-hominem attacks ($p < 0.01$) but only for the interactor ratio. The second Mann Whitney U test did not show a significant difference between scores for non-ad-hominem attacks on FAVs, RTs, responses and ad-hominem. However, a significant difference was found between the non-ad-hominem score of interactor ratio and the ad-hominem score for interactor ratio $U = 22053$, $z = -4.753$, $p < .001$. Despite the lack of significance, the overall mean ranks of Ad-hominem attacks are higher than non-ad-hominem attacks, following the expected effect of H2. This result partially supports H2, in regards to ad-hominem attacks generating a higher interactor ratio than non-ad-hominem attacks.

5. Discussion

5.1. Main findings

In this study, a categorisation of 18 different types of personal attacks was established in a campaign environment. Of these 18 categories, 13 were also used to attack political parties. Although 60% of political communication on Twitter does not involve attacks, there is still a significant number of tweets that do contain attacking statements (around 40% of tweets).

Looking at the specific types of arguments, it can be seen that Negligence is the most used type of ad-hominem argument. Negligence incises in the opposition's incompetence for governance, which seems fundamental in a campaign environment; performance is linked to political trust (van der Meer, 2017). This result may also be related to the case study taking place within the COVID-19 context, as governments and parties have been repeatedly judged by their actions towards the control of the pandemic. Ad-hominem 2 is the second most used ad-hominem argument, probably due to highlighting the mistakes of the opposition as well as Negligence. Resigning implies a loss of trust (Isotalus & Almonkari, 2014) and therefore could help a politician to transfer votes originally going towards the opposition. It seems like politicians are likely to use attacks that seek to dimmish the trust of their opposition by attacking their performance or forcing them to amend their errors.

Although these two categories are the most used by politicians, these are not the types of attacks with the most engagement. When looking at engagement, Ad-hominem 4 "Diminishing credibility", Liar and Media-corrupt show the highest scores for RTs and Favs, which are the metrics that show overall positive support by Twitter users. Users are more likely to disseminate information if it emerges from an ideologically similar source (Barberá et al., 2015) which is the main use of RTs and Favs; to broadcast information (Yaqub et al., 2017). These three types of attacks aim to diminish the credibility of the opposition, which may give self-confirmation to users that agree with the politician that they are correct. This correlates with the perception of Twitter as an echo chamber, where users broadcast information that supports their views (Michailidou, 2017).

On the other hand, Media-corrupt and Violent show the highest scores for responses. Responses may show either support or backlash by Twitter users, as they can be triggered by a disagreement with the tweet. A previous study analysing political controversy showed that most responses to a controversial tweet were negative, even though most of the tweets referring to the issue without responding showed a positive response (Bogen et al., 2021). Users are triggered to respond to opposed messages, as a previous study by Bail et al. (2018) shows that exposure to Tweets from an opposed political source significantly increase radicalization. Media-corrupt and Violent types of arguments are part of the populist language where victimization is of key importance to give the populist leader moral legitimacy (Birks, 2011). This tone of a victim may be

the key element for these two arguments to trigger responses by either supporters or ideologically opposed users.

The ad-hominem types that provide the highest scores in the interactor ratio are Anti-democracy, Radical and Ad-hominem 1 “Excluding from dialogue”. These three types are linked, as Ad-hominem 1 “Excluding from dialogue” was mostly used from candidates that suggested a “cordon sanitaire” on others, being the usual argument that they are too extremist or harmful for democracy. This is a clear use of polarising language, as the opposition is completely labelled and censored, as populists do to bring the group together (Engesser et al., 2017). Hence, it would appear that populist statements would increase interactions in comparison to the follower base, that is, the interactor ratio. Although these populist attacks succeed in engagement, addressing others as populists provide low scores for every engagement metric, as can be seen in the stats from the Populist ad-hominem type.

5.2. Theoretical contributions

This study contributes to the existing literature by disclosing a categorisation of 18 types of ad-hominem arguments that were used by candidates on Twitter during a campaign. Although there were some types already identified by previous authors (Borovali, 2018; Harris et al., 2012; Macagno, 2013) more ad-hominem attacks were found to be used by the candidates of this case study. This new categorisation includes mainly personal attacks on the political attributes of the other candidates. The different types of attacks were compared in their metrics (Favs, RTs and responses) as well as the interactor ratio to compare the engagement outside the follower-base of a candidate.

This study finds that politicians mostly attack their competitors’ capacity for governance, by calling them negligent and highlighting their errors. The attacks that generate the most supportive engagement (RTs and Favs) are those that diminish the credibility of the opposition which supports the perception of Twitter as an echo chamber (Vaccari et al., 2016). The attacks that generate the most responses are those that either call the opposition violent or corrupt in regards to the media, which are attacks commonly used by populist politicians who victimize themselves. Victimization is a classic populist strategy (Birks, 2011), and this result supports other findings showing that polarising messages trigger engagement (Bail et al., 2018; Sobieraj & Berry, 2011). Additionally, this study uses for the first time the interactor ratio as a metric with polarising connotations in the political environment.

Findings also showed that attacks generate more engagement than non-attacks and that ad-hominem arguments generate more interactor ratio than non-ad-hominem attacks. These results contribute to the existing literature by suggesting a positive correlation between polarisation and engagement (Conover et al., 2011), which would explain the success of populist strategies in social media and the urgency to address this political issue.

5.3. Practical implications

The inventory of personal attacks that are shown in this case study could be used by social media companies to identify personal attacks. This can be done without interfering with the free speech of politicians; for example, by learning how personal attacks can be identified, Twitter can create digital debate spaces where attacks are not allowed and only policymaking can be discussed while still maintaining freedom of speech on their current platform. Creating a clean political arena would ultimately benefit citizens, as politicians would have to strictly defend and strategize their political program without using attacks.

On the other side of the screen, these results should serve for users to identify the populist practices that some politicians use by dedicating a gross part of their communicative strategy to polarising attacks instead of relevant information for voters (such as discussing their political program). Users could see this data as a wake-up call to stop interacting with controversial tweets, fighting the emotional trigger that populists seek for their tweets to succeed. Politicians serve the citizens as the Twitter algorithm serves their users, but this power needs to be claimed back. By not falling into triggers and polarization, and by demanding real information and change, the future of digital political communication could aim towards discussing information instead of creating polarisation.

5.4. Limitations

These results are limited from a location standpoint; the Spanish demography and the tense political environment that revolves in Spain and the Community of Madrid may differ from other campaign scenarios, where engagement on attacks might be different. The Community of Madrid elections from 2021 was a particular and tense campaign; the vice-president of Spain surprisingly resigned from his position to run as a candidate for the Community (Jones, 2021), candidates for this campaign received bullets and threats among other members of the Government (Europa Press, 2021) and political rallies were marked by explicit violence (el País, 2021). All these factors increased the participation to over 75%, a historic peak (Viúdez & Gálvez, 2021). Although the particularities of Spanish politics might be a limitation in the sense of global representativeness, this tension was a key factor that provided enough examples of attacks and ad-hominem arguments for the qualitative part of the study. Another particularity in the case study that may have influenced the results is that the campaign and the elections could have been conditioned by the COVID-19 pandemic, and results may differ in another time frame, once the COVID-19 pandemic is over.

Using interactions as a measurement may also be a limitation due to the crudeness of these metrics when it comes to representing the intentions of the users. These metrics could be substituted for constructs that may portray engagement more accurately, especially if there is access to other types of hidden metrics such as proposed by Muñoz-Expósito et al. (2017). However, hidden metrics were not available for this study. Additionally, metrics on this study are considered positive when coming

from the follower-base and possibly controversial when coming from outside of the follower-base, such as the ones measured by the interactor ratio. This is based on the echo chamber theory, as users are supposed to follow those politicians that they agree with. But still, there might be exceptions for this case. For example, it might be possible that some "trolls" also follow politicians they disagree with to interact with them faster. As trolling is a common practice in politics (Bulut & Yörük, 2017) the validity of the interactor ratio based on the echo chamber theory is worthy of revision in future studies.

The number of tweets per politician was not taken into account when selecting the sample of tweets. However, some politicians may tweet more than others, or in a more polarising way than others, inflating the number of attacks in the final results. This limitation could be addressed in a future study by equivalating the sample size of the tweets for each politician.

5.5. Suggestions for future research

Future research needs to be performed to analyse these populist strategies that are rapidly spreading in western democracies using social media as their main tool. The aim of these potential future studies should be to help develop counterstrategies to these polarising techniques, for example by providing the algorithm of social media companies with more information to identify these attacks. Hence, this same study could be extended in the future by expanding the coding to new arguments and new typologies such as party attacks.

Additionally, the usage of attacks could be compared between different political inclinations or political parties, which could serve voters to judge the role of different parties in creating polarisation. These results could also be linked to voting intention polls or even election results to study the relationship between digital attacks and electoral success. Additionally, a future study could be extended to include other political actors such as activists, bloggers, congressmen or senators to identify the toxicity of other political figures. And naturally, the study could be extended to other social media platforms or politician's interviews or articles on digital newspapers with measurable metrics

This study could be replicated for an in-depth analysis of engagement by measuring the subjectivity of Twitter metrics. For example, performing a sentiment analysis in the responses with a natural language processor such as TextBlob could provide a polarity and subjectivity score that could be compared and tested for a significant difference among the different types of tweets. Other potential future studies on the topic could use new methods to understand how triggering works outside of the digital space. With neuro-marketing techniques, the emotions of users can be measured when reading these attacking tweets. Other techniques such as surveys or interviews can also be used to identify the feelings of users when reading these tweets. An experimental take on studying users' reactions would provide more insights into how triggering works, which is key to understanding the populist phenomenon.

5.6. Conclusion

Populism has risen in Western democracies partly due to the development of the Internet and social media such as Twitter, where politicians can easily connect with their followers and attack their opposition. This research was focused on discovering the types of attacks that politicians use and how users engage with these types of attacks. By analysing a real case study, it is shown that politicians use up to 18 different types of attacks, mostly focusing on undesirable political attributes. Additionally, attacks generate more engagement than non-attacks while ad-hominem arguments generate only more engagement out of the follower base, which has a polarising nuance. These findings should serve for society to reflect on their role in the development of political polarisation and populism. Users should disengage from polarising statements and claim the accountability of political parties and social media platforms to develop a clean political debate space.

References

Alonso-Muñoz, L. (2019). The 'More is more' effect: a comparative analysis of the political agenda and the strategy on Twitter of the European populist parties. *European Politics and Societ*, *21*(5), 505–519.

Alonso-Muñoz, L., & Casero-Ripollés, A. (2018). Communication of European populist leaders on Twitter: Agenda setting and the 'more is less' effect. *El Profesional de La Información*, *27*(6), 1193. https://doi.org/10.3145/epi.2018.nov.03

Alsinet, T., Argelich, J., Béjar, R., Fernández, C., Mateu, C., & Planes, J. (2017). Weighted argumentation for analysis of discussions in Twitter. *International Journal of Approximate Reasoning*, *85*, 21–35. https://doi.org/10.1016/j.ijar.2017.02.004

Altoaimy, L. (2018). Driving change on Twitter: A corpus-assisted discourse analysis of the Twitter debates on the Saudi Ban on women driving. *Social Sciences*, *7*(5). https://doi.org/10.3390/SOCSCI7050081

Anger, I., & Kittl, C. (2011). Measuring influence on Twitter. *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies - i-KNOW '11*, 1. https://doi.org/10.1145/2024288.2024326

Aragõn, P., Kappler, K. E., Kaltenbrunner, A., Laniado, D., & Volkovich, Y. (2013). Communication dynamics in Twitter during political campaigns: The case of the 2011 Spanish national election. *Policy and Internet*, *5*(2), 183–206. https://doi.org/10.1002/1944-2866.POI327

Arroyas Langa, E., & Pérez Díaz, P. L. (2016). La nueva narrativa identitaria del populismo: un análisis del discurso de Pablo Iglesias (Podemos) en Twitter. *Cultura, Lenguaje y Representación - CLR*, *15*, 51–63. https://doi.org/10.6035/clr.2016.15.4

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Fallin Hunzaker, M. B., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *PNAS*. https://doi.org/10.1073/pnas.1804840115

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right. *Psychological Science*, *26*(10), 1531–1542. https://doi.org/10.1177/0956797615594620

Baviera, T., Calvo, D., & Llorca-Abad, G. (2019). Mediatisation in Twitter: an exploratory analysis of the 2015 Spanish general election. *The Journal of International Communication*, *25*(2), 275–300.

Bernhard, L. (2020). Revisiting the inclusion-moderation thesis on radical right populism: Does party leadership matter? *Politics and Governance*, *8*(1), 206–216. https://doi.org/10.17645/pag.v8i1.2515

Birks, J. (2011). The politics of protest in newspaper campaigns: Dissent, populism and the rhetoric of authenticity. *British Politics*, *6*(2), 128–154. https://doi.org/10.1057/bp.2011.5

Bode, L., & Dalrymple, K. E. (2016). Politics in 140 characters or less: Campaign communication, network interaction, and political participation on Twitter. *Journal of Political Marketing*, *15*(4), 311–332. https://doi.org/10.1080/15377857.2014.959686

Bogen, K. W., Williams, S. L., Reidy, D. E., & Orchowski, L. M. (2021). We (want to) believe in the best of men: A qualitative analysis of reactions to #Gillette on Twitter. *Psychology of Men and Masculinity*, *22*(1), 101–112. https://doi.org/10.1037/men0000308

Borovali, M. (2018). Ad hominem argumentation in politics. *Philosophy and Social Criticism*, *44*(4), 426–436. https://doi.org/10.1177/0191453718755206

Bruns, A., & Highfield, T. (2013). Political networks on Twitter. *Information, Communication & Society*, *16*(5), 667–691. https://doi.org/10.1080/1369118X.2013.782328

Bucy, E. P., Foley, J. M., Lukito, J., Doroshenko, L., Shah, D. v., Pevehouse, J. C. W., & Wells, C. (2020). Performing populism: Trump's transgressive debate style and the dynamics of Twitter response. *New Media and Society*, *22*(4), 634–658. https://doi.org/10.1177/1461444819893984

Bulut, E., & Yörük, E. (2017). Digital populism: Trolls and political polarization of Twitter in Turkey. *International Journal of Communication*, *11*, 4093–44117.

Cardaioli, M., Kaliyar, P., Capuozzo, P., Conti, M., Sartori, G., & Monaro, M. (2020). Predicting Twitter users' political orientation: An application to the Italian political scenario. *Proceedings of the 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2020*, 159–165. https://doi.org/10.1109/ASONAM49781.2020.9381470

Carenini, G., & Moore, J. D. (2006). Generating and evaluating evaluative arguments. *Artificial Intelligence*, *170*(11), 925–952. https://doi.org/10.1016/j.artint.2006.05.003

Carrella, F. (2020). #Populism on Twitter : statistical analysis of the correlation between tweet popularity and "populist" discursive features. *Brno Studies in English*, *2*, 5–23. https://doi.org/10.5817/BSE2020-2-1

Cayrol, C., & Lagasquie-Schiex, M. C. (2005). Graduality in argumentation. *Journal of Artificial Intelligence Research*, *23*, 245–297. https://doi.org/10.1613/jair.1411

Conover, M. D., Ratkiewicz, J., Francisco, M., Goncalves, B., Flammini, A., & Menczer, F. (2011). Political polarization on Twitter. *Fifth International AAAI*

Conference on Weblogs and Social Media. http://truthy.indiana.edu/site_media/pdfs/conover_icwsm2011_polarization.pdf.

Cornelissen, L. (2019). We, the peoples: Populist leadership, neoliberalism and decoloniality. *Araucaria*, *21*(42), 529–551. https://doi.org/10.12795/araucaria.2019.i42.23

del Valle, M. E., Sijtsma, R., Stegeman, H., & Borge, R. (2020). Online deliberation and the public sphere: Developing a coding manual to assess deliberation in Twitter political networks. *Javnost*, *27*(3), 211–229. https://doi.org/10.1080/13183222.2020.1794408

Dragoni, M., da Costa Pereira, C., Tettamanzi, A. G. B., & Villata, S. (2018). Combining argumentation and aspect-based opinion mining: The SMACk system. *AI Communications*, *31*(1), 75–95. https://doi.org/10.3233/AIC-180752

Dung, P. M. (1993). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning and logic programming. *Ijcai-93, Vols 1 and 2: Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, 852–857.

el País. (2021, April 9). Vox en Vallecas: violencia inaceptable. *El País*. https://elpais.com/opinion/2021-04-09/vox-en-vallecas-violencia-inaceptable.html

Engesser, S., Fawzi, N., & Larsson, A. O. (2017). Populist online communication: introduction to the special issue. *Information, Communication & Society*, *20*(9), 1279–1292. https://doi.org/10.1080/1369118X.2017.1328525

Europa Press. (2021, May 16). Se atasca la investigación de las cartas con balas de la campaña de Madrid porque no hay huellas de los autores. *El Mundo*. https://www.elmundo.es/elecciones/elecciones-madrid/2021/05/16/60a0f00afc6c8321278b45b6.html

Evans, H. K., & Clark, J. H. (2016). "You tweet like a girl!": How female candidates campaign on Twitter. *American Politics Research*, *44*(2), 326–352. https://doi.org/10.1177/1532673X15597747

Finn, H. C. (2019). The defamatory potential of ad hominem criticism: Guidance for advocacy in public forums. *Pacific Conservation Biology*, *25*(1), 92–104. https://doi.org/10.1071/PC17022

Flew, T., & Iosifidis, P. (2020). Populism, globalisation and social media. *International Communication Gazette*, *82*(1), 7–25. https://doi.org/10.1177/1748048519880721

Golec de Zavala, A., & Keenan, O. (2021). Collective narcissism as a framework for understanding populism. *Journal of Theoretical Social Psychology*, *5*(2), 54–64. https://doi.org/10.1002/jts5.69

Grosse, K., Chesñevar, C. I., & Maguitman, A. G. (2012). An argument-based approach to mining opinions from twitter. *CEUR Workshop Proceedings*, *918*, 408–422.

Hamo, M., Kampf, Z., & Weiss-Yaniv, N. (2019). Populism as a keyword and as a meta-discursive resource for positioning in mediated political discourse. *Discourse, Context & Media*, *29*, 100283. https://doi.org/10.1016/j.dcm.2018.11.005

Harris, A. J. L., Hsu, A. S., & Madsen, J. K. (2012). Because Hitler did it! Quantitative tests of Bayesian argumentation using ad hominem. *Thinking and Reasoning*, *18*(3), 311–343. https://doi.org/10.1080/13546783.2012.670753

Isotalus, P., & Almonkari, M. (2014). Political scandal tests trust in politicians. *Nordicom Review*, *35*(2), 3–16. https://doi.org/10.2478/nor-2014-0011

Jagers, J., & Walgrave, S. (2007). Populism as political communication style: An empirical study of political parties' discourse in Belgium. *European Journal of Political Research*, *46*(3), 319–345. https://doi.org/10.1111/j.1475-6765.2006.00690.x

Jones, S. (2021). *Pablo Iglesias leaves Spanish politics, 'very proud' of Podemos legacy.* The Guardian. https://www.theguardian.com/world/2021/may/05/pablo-iglesias-leaves-politics-podemos-spain

Keller, T. R., & Kleinen-von Königslöw, K. (2018). Followers, spread the message! Predicting the success of Swiss politicians on Facebook and Twitter. *Social Media + Society*, *4*(1), 205630511876573. https://doi.org/10.1177/2056305118765733

Khatua, A., Khatua, A., & Cambria, E. (2020). Predicting political sentiments of voters from Twitter in multi-party contexts. *Applied Soft Computing*, *97*, 106743. https://doi.org/10.1016/j.asoc.2020.106743

Kingsbury, D. V. (2016). From populism to protagonism (and back?) in Bolivarian Venezuela: Rethinking Ernesto Laclau's on populist reason. *Journal of Latin American Cultural Studies*, *25*(4), 495–514. https://doi.org/10.1080/13569325.2016.1230939

Kioupkiolis, A. (2016). Podemos: the ambiguous promises of left-wing populism in contemporary Spain. *Journal of Political Ideologies*, *21*(2), 99–120. https://doi.org/10.1080/13569317.2016.1150136

Kotzee, B. (2010). Poisoning the well and epistemic privilege. *Argumentation*, *24*(3), 265–281. https://doi.org/10.1007/s10503-010-9181-8

Laclau, E. (1996). *Emancipación y diferencia*. Ariel.

Lavrakas, P. (2013). Intercoder reliability. *Encyclopedia of Survey Research Methods*. https://doi.org/10.4135/9781412963947.n228

Lorentzen, D. G. (2020). Bridging polarised Twitter discussions: the interactions of the users in the middle. *Aslib Journal of Information Management*, *73*(2), 129–143. https://doi.org/10.1108/AJIM-05-2020-0154

Macagno, F. (2013). Strategies of character attack. *Argumentation*, *27*(4), 369–401. https://doi.org/10.1007/s10503-013-9291-1

Makazhanov, A., Rafiei, D., & Waqar, M. (2014). Predicting political preference of Twitter users. *Social Network Analysis and Mining*, *4*(1), 193. https://doi.org/10.1007/s13278-014-0193-5

Michailidou, A. (2017). Twitter, public engagement and the Eurocrisis: More than an echo chamber? In *Social Media and European Politics* (pp. 241–266). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-59890-5_11

Mizrahi, M. (2010). Take my advice-I am not following it: Ad hominem arguments as legitimate rebuttals to appeals to authority. *Informal Logic*, *30*(4), 435–456. https://doi.org/10.22329/il.v30i4.2990

Moya Sánchez, M., & Herrera Damas, S. (2015). How can Twitter contribute to more advanced political communication? | Cómo puede contribuir twitter a una comunicación política más avanzada. *Arbor*, *191*(774). https://doi.org/10.3989/arbor.2015.774n4012

Mudde, C. (2004). The Populist Zeitgeist. *Government and Opposition*, *39*(4), 541–563. https://doi.org/10.1111/j.1477-7053.2004.00135.x

Muñoz-Expósito, M., Oviedo-García, M. Á., & Castellanos-Verdugo, M. (2017). How to measure engagement in Twitter: advancing a metric. *Internet Research*, *27*(5), 1122–1148. https://doi.org/10.1108/IntR-06-2016-0170

Orkibi, E. (2018). Precedential ad hominem in polemical exchange: Examples from the Israeli political debate. *Argumentation*, *32*(4), 485–499. https://doi.org/10.1007/s10503-018-9453-2

Pancer, E., & Poole, M. (2016). The popularity and virality of political social media: hashtags, mentions, and links predict likes and retweets of 2016 U.S. presidential nominees' tweets. *Social Influence*, *11*(4), 259–270. https://doi.org/10.1080/15534510.2016.1265582

Primario, S., Borrelli, D., Iandoli, L., Zollo, G., & Lipizzi, C. (2017). Measuring polarization in Twitter enabled in online political conversation: The case of 2016 US presidential election. *2017 IEEE International Conference on Information Reuse and Integration (IRI)*, 607–613. https://doi.org/10.1109/IRI.2017.73

Sailofsky, D. (2021). Masculinity, cancel culture and woke capitalism: Exploring Twitter response to Brendan Leipsic's leaked conversation. *International Review for the Sociology of Sport*. https://doi.org/10.1177/10126902211039768

Sobieraj, S., & Berry, J. M. (2011). From incivility to outrage: Political discourse in blogs, talk radio, and cable news. *Political Communication*, *28*(1), 19–41. https://doi.org/10.1080/10584609.2010.542360

Suau-Gomila, G., Pont-Sorribes, C., & Pedraza-Jiménez, R. (2020). Politicians or influencers? Twitter profiles of Pablo Iglesias and Albert Rivera in the Spanish general elections of 20-d and 26-j. *Communication and Society*, *33*(2), 209–225. https://doi.org/10.15581/003.33.2.209-225

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. *2010 IEEE Second International Conference on Social Computing*, 177–184. https://doi.org/10.1109/SocialCom.2010.33

Tong, J., & Zuo, L. (2020). Mainstreaming populism through the Twitter practices of politicians and the news media: A case study of the 2016 Brexit referendum debates. *Information Polity*, *25*(3), 361–375. https://doi.org/10.3233/IP-190197

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. *ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 178–185.

Twitter Inc. (2021). *Permanent suspension of @realDonaldTrump*. Twitter Blog. https://blog.twitter.com/en_us/topics/company/2020/suspension

Usher, J., Morales, L., & Dondio, P. (2019). BREXIT: A granger causality of twitter political polarisation on the FTSE 100 Index and the Pound. *Proceedings - IEEE 2nd International Conference on Artificial Intelligence and Knowledge Engineering, AIKE 2019*, 51–54. https://doi.org/10.1109/AIKE.2019.00017

Vaccari, C., Valeriani, A., Barberá, P., Jost, J. T., Nagler, J., & Tucker, J. A. (2016). Of echo chambers and contrarian clubs: Exposure to political disagreement among German and Italian users of Twitter. *Social Media + Society*, *2*(3), 205630511666422. https://doi.org/10.1177/2056305116664221

van der Meer, T. W. G. (2017). *Economic Performance and Political Trust* (E. M. Uslaner, Ed.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190274801.013.16

van Eemeren, F. H., Garssen, B., & Meuffels, B. (2012). The disguised abusive ad hominem empirically investigated: Strategic manoeuvring with direct personal attacks. *Thinking and Reasoning*, *18*(3), 344–364. https://doi.org/10.1080/13546783.2012.678666

van Kessel, S., & Castelein, R. (2016). Shifting the blame. Populist politicians' use of Twitter as a tool of opposition. *Journal of Contemporary European Research*, *12*(2), 594–614.

Viúdez, J., & Gálvez, J. J. (2021, May 4). La participación en las elecciones de Madrid se dispara por encima del 75%. *El País*. https://elpais.com/espana/elecciones-madrid/2021-05-04/datos-de-participacion-en-las-elecciones-de-la-comunidad-de-madrid.html

Vredenburg, J., Kapitan, S., Spry, A., & Kemper, J. (2020). *Police brutality and running shoes: Authentic brand activism or woke washing: An abstract*. 623–624. https://doi.org/10.1007/978-3-030-39165-2_255

Walton, D. N. (2000). Case study of the use of a circumstantial ad hominem in political argumentation. *Philosophy and Rhetoric*, *33*(2), 101–115. https://doi.org/10.1353/par.2000.0015

Yaqub, U., Chun, S. A., Atluri, V., & Vaidya, J. (2017). Analysis of political discourse on Twitter in the context of the 2016 US presidential elections. *Government Information Quarterly*, *34*(4), 613–626. https://doi.org/10.1016/j.giq.2017.11.001

Appendix

Appendix A

Category	Description
Attack	The tweet is a mention and includes some criticism directed at another politician's actions (del Valle et al., 2020). However, it is not considered an ad-hominem.
Non-attack	The tweet does not include criticism directed to another politician's actions (del Valle et al., 2020).
Ad-hominem	The tweet is an Attack on somebody's personal character (Kotzee, 2010).
Ad-hominem – "you too"	The person is discredited due to violating the principles he is advocating (Borovali, 2018).
Ad hominem – "whataboutery"	The person is discredited based on the convenience of speaking up on an instance that is aligned with their interests (Borovali, 2018).
Ad-hominem – Bias in the opponent	The person is discredited based on their argument being part of their agenda of interests (Borovali, 2018).
Ad-hominem – Direct personal attack	The person is discredited based on his personal attributes (Borovali, 2018).
Ad-hominem 1 "Excluding from dialogue"	The attack tries to exclude the interlocutor from the dialogue (Macagno, 2013).
Ad-hominem 2 "Forcing viewpoint"	The attack tries forced persuasion to force the interlocutor to support a viewpoint (Macagno, 2013).
Ad-hominem 3 "Negative judgement"	The attack tries to trigger negative judgement from other users (Macagno, 2013).
Ad-hominem 4 "Diminishing credibility"	The attack tries to diminish the credibility of the source (Macagno, 2013)
Ad-Hitlerum	"Hitler also did it, therefore this person should not be trusted" (Harris et al., 2012)
Ad-hominem - Antidemocracy	The person is addressed as being against democratic practices and principles.
Ad-hominem - Copycat	The person is addressed as being a copy or ally of another political opponent.
Ad-hominem - Inverted whataboutery	The person is discredited based on the convenience of not speaking up on an instance that is not aligned with their interests. This is the inverted argument of "whataboutery" by Borovali (2018)
Ad-hominem - Liar	The person is accused of lying or hiding relevant information.
Ad-hominem - Negligent	The person is addressed as a negligent politician, ignoring groups of people or applying incoherent or unnecessary policies.
Ad-hominem - Corrupt	The person is addressed as being corrupt or taking part in corrupt activities.

Ad-hominem - Polarising	The person is addressed as a polarising politician or trying to polarise society with their messages.
Ad-hominem - Radical	The person is addressed as a radical or extremist.
Ad-hominem - Reactionary	The person is addressed as opposed to political, scientific or societal progress.
Ad-hominem - Violent	The person is addressed as violent or taking part in violent activities.
Ad-hominem - Populist	The person is addressed as populist.
Ad-hominem - Elitist	The person is addressed as living as an elitist.
Ad-hominem - Use media for own benefit	The person is accused of having relation with media sources or using media for their political benefit.
Party – Negative judgement	The attack tries to trigger negative judgement from other users. Based on Ad-hominem 3 by Macagno (2013).
Party - Antidemocratic	The party is addressed as being against democratic practices and principles.
Party - Attack	The tweet mentions another party and includes some criticism directed at another party's actions
Party - Copycat	The party is addressed as being a copy or ally of another political opponent.
Party - Corrupt	The party is addressed as being corrupt or taking part in corrupt activities.
Party - Inverted whataboutery	The party is discredited based on the convenience of not speaking up on an instance that is not aligned with their interests. This is the inverted argument of "whataboutery" by Borovali (2018)
Party - Liars	The party is accused of lying or hiding relevant information.
Party - Negligent	The party is addressed as a negligent party, ignoring groups of people or applying incoherent or unnecessary policies.
Party - Polarising	The person is addressed as a polarising party or trying to polarise society with their messages.
Party - Populists	The party is addressed as populist.
Party - Radical	The party is addressed as a radical or extremist.
Party - Reactionary	The party is addressed as opposed to political, scientific or societal progress.
Party – Media-corrupt	The party is accused of having relation with media sources or using media for their political benefit.
Party - Violent	The party is addressed as violent or taking part in violent activities.
Party - You too	The person is discredited due to violating the principles they are advocating. Based on ad-hominem "you too" by Borovali (2018).

Appendix B

Table 3:
Ranks of H1 test.

	Attack	N	Mean Rank	Sum of Ranks
FAVs	*No*	829	633.28	524992.5
	Yes	549	774.39	425138.5
	Total	1378		
RTs	*No*	829	612.36	507642.5
	Yes	549	805.99	442488.5
	Total	1378		
Responses	*No*	829	618.34	512601
	Yes	549	796.96	437530
	Total	1378		
Interactor Ratio	*No*	829	618.36	512623.5
	Yes	549	796.92	437507.5
	Total	1378		

Table 4:
H1 Test Statistics[a]

	FAVs	RTs	Responses	Interactor Ratio
Mann-Whitney U	180957.5	163607.5	168566	168588.5
Z	-6.444	-8.843	-8.158	-8.154
Sig. (2-tailed)	.000	.000	.000	.000

a. Grouping Variable: Attack

Table 5:
Medians of H1

Attack	FAVs	RTs	Responses	Interactor Ratio
No	1054	334	88	.009
Yes	1995	695	198	.015
Total	1322	445	125	.011

Table 6:
Ranks of H2 test

	Attack	N	Mean Rank	Sum of Ranks
FAVs	Non ad-hominem	399	271.06	108153.00
	Ad-hominem	150	285.48	42822.00
	Total	549		
RTs	Non ad-hominem	399	273.45	109108.50

	Ad-hominem	150	279.11	41866.50
	Total	549		
Responses	Non ad-hominem	399	273.92	109294.00
	Ad-hominem	150	277.87	41681.00
	Total	549		
Interactor Ratio	Non ad-hominem	399	255.27	101853.00
	Ad-hominem	150	327.48	49122.00
	Total	549		

Table 7:

H2 Test Statistics[a]

	FAVs	RTs	Responses	Interactor Ratio
Mann-Whitney U	28353	29308.5	29494	22053
Z	-.949	-.372	-.260	-4.753
Asymp. Sig. (2-tailed)	.343	.710	.795	.000

a. Grouping Variable: Type of attack

Table 8:

Medians of H2

Attack	FAVs	RTs	Responses	Interactor Ratio
Non ad-hominem	2020	709	204	.013
Ad-hominem	1968	636.5	184.5	.02
Total	1995	695	198	.015