THE "EDGE" OF POLITICAL HUMOR: EFFECT OF DARK HUMOR ON AFFECTIVE POLARIZATION

By
Mukhtar Tashkenbayev

Submitted to Central European University - Private University
Department of Political Science

In partial fulfillment of the requirements for the degree of Master of Arts in Political Science

Supervisor: Professor Mariyana Angelova

Vienna, Austria
2022

ABSTRACT

Since the internet came into widespread usage, patterns of exposure to political humor have changed. Due to its relative lack of oversight and censorship, internet-based humor often tends to be more brazen, offensive, and dark compared to humor seen in traditional media. This study attempts to establish the link between dark political humor and affective polarization. It theorizes that dark humor is likely to cause a stronger polarizing reaction compared to casual humor, utilizing a survey experiment to test its hypotheses. The survey focuses on political issues relevant to the United States of America, gathering participants from the country, and presents the jokes in the form of internet "memes" collected on social media platforms. Results vary depending on the group of respondents and the type of measurement, but for most measurements, no statistically significant difference between dark and casual humor in generating affective polarization was found. However, significant support for several supporting hypotheses was established.

Keywords: political humor, dark humor, offensive humor, affective polarization, political polarization

ii

Dedicated to my grandfather, who has always supported me, and my supervisor, who has shown me nothing but kindness.

TABLE OF CONTENTS

1. Introduction

Political actors, must these be parties or individual candidates, typically seek to promote their interests, no matter what these are. One of the main ways to do so is to increase their popularity relative to other actors, which can lead to changes in voting behavior or mobilization of their supporters. Presumably, to achieve that, they must be strategic and make conscious choices about the ways how they approach things like campaigning, how they present themselves, and their opponents. One of the ways to do that is humor.

Political humor is an integral part of popular political discussion. It is engaged by a variety of actors, such as citizens, media personalities, and politicians themselves. Its importance can be emphasized by its enduring wide appeal under a whole range of circumstances. Political humor was disseminated by citizens even under the most oppressive regimes, in which engaging in it could lead to serious sanctions, such as Franco's Spain (Pi-Sunyer 1977) or the Stalinist USSR (Brandenberger 2009). Popular comedy talk shows in contexts such as the US can, while still primarily being focused on comedy, carry as much substantive information about politics as designated news programs do (Fox, Koloen, and Sahin 2007). Humor can be effective in the hands of politicians themselves. For example, during the 2008 presidential election in the US, Obama's campaign's messages on social media that carried an enthusiastic and humorous content typically yielded the most likes and shares (Borah 2016).

While it is clear that humor is a tool that's used for political reasons, it is not a uniform concept. Jokes can be very different in their tone, concept, form, direction, and various other aspects. While there is a lot of literature that tries to dissect the effects of political humor, it typically focuses on a few select dimensions. Some research focuses on the differences in attitude changes caused by "hostile" or "attack" humor, targeted at disparaging opponents, and more self-conscious, self-deprecating humor. Another major theme is the dyad of initial attitudes and

1

perceptions of humor directed at different topics (e.g. whether Republicans find jokes about Democrats or Republicans more funny or convincing).

Naturally, the breadth of research on political humor has historically focused on traditional media, such as TV shows (Becker and Waisanen 2013). While they still give useful insight into the ways people perceive humor, they are not always capable to answer the questions arising in the current context of the newly established importance of internet-based content. The emergence of social media in particular allows political actors to have high levels of freedom in creating and disseminating media content. While some of them had access to traditional forms of media such as print and TV, they were still constrained by editorial standards as well as the formats of these outlets. In contrast, social media can enable a variety of actors to engage in almost uncensored content creation, which opens up their capability to utilize a whole variety of previously unavailable types of humor in their favor.

Moreover, the internet does not only allow a greater degree of freedom to political and non-political actors but also establishes new standards for what is seen as fun or acceptable. Many formats of internet humor are radically different from those exercised in traditional media, while at the same time having high appeal. This can manifest in many forms, including it being cruder and presenting in "do-it-yourself" form, involving references to previous internet jokes and memes, or being highly ironic.

At the same time, politicians often have trouble navigating the social media landscape, not engaging with it interactively, possibly, due to understanding problems about appropriate presentation and boundary setting (Hoffmann and Suphan 2017). Explicitly engaging in "edgy" or "unprofessional" humor can be seen as off-brand for them, and it may actively dissuade certain voters from supporting such politicians, especially given the presumable youth-based appeal of such jokes. At the same time, this type of humor has not been fully ignored.

2

Particularly on the fringes of the political spectrum, such as extreme right-wing in Finland, such humor has been adopted to propagate messaging related to recent news, mottos, or even historical arguments for causes such as anti-migration (Hakoköngäs et al. 2020).

Of course, anti-migrant and other types of exclusionary humorous messaging do not necessarily have to originate from a centralized or organized effort. In regions with high anti-immigrant sentiment, users are happy to utilize humor to ostracize outgroups on their own. For example, it is a common occurrence for Northern Chileans to make and share "memes" about Bolivians, a poorer group that typically resides in the region for migrant work. Those "memes" often do carry extreme messaging, including that of racial inferiority in traits like intellect, but the humorous wrapping often makes the messaging more palatable (Haynes 2019).

While there are studies that aim at establishing the effect of certain types of humor, they typically do not focus on things such as form or tone. Jokes may be relative softballs that poke fun at inoffensive topics about candidates, but they can also be examples of black humor touching on heavy or even taboo themes. Similarly, and particularly within the internet discourse, jokes may take the form of different viral templates ("memes") or more traditionally conventional forms such as anecdotes. They can also differ in the way content is delivered; some of the jokes are heavily visual, while other ones function via text, and that may affect the perception of the readers. Targets of the content may also differ; some jokes may focus on the elite level, such as politicians, some on the voter base of said politicians, and some on general concepts and ideas espoused by political ideologies.

If such factors affect the appeal of humor to people who see it, it can lead to numerous important effects. Since political humor is often tied to certain messaging, it can affect the degree of persuasiveness that these messages hold, or at least the length of time these messages appear to be salient in the reader's minds. Similarly, "funnier" jokes might have a higher chance

3

to be shared by the readers; at the same time, if a joke is too dark or inappropriate, it might stop that process even if the joke is perceived as relatively funnier than a milder one. This dynamic can play differently within different modes of sharing (e.g. through public pages, private messaging, or word of mouth).

Political actors, must these be politicians themselves, groups related to them, or even regular supporters can make conscious choices about what types of political humor to generate or engage with. If political humor can have effects in forms of things like affective polarization, political mobilization, and changes in voting behavior, and if different types of humor generate different effects, then it is worth exploring.

Therefore, tentatively, the research question that I present for this research is: ***How does dark political humor affect affective polarization?***

That is a broad question that I will attempt to turn into workable hypotheses that are possible to operationalize in the following section.

2. Literature review

Several findings suggest that political humor can, directly or indirectly, affect the political behavior of its recipients. Becker finds, for example, that exposure to political humor can increase recipients' political trust, which can, in turn, lead to more engagement in politics in general (Becker 2011). According to Verhulsdonk et al., humor used for attack messaging can help reduce the chance of backlash against the attack (Verhulsdonk et al. 2021). Coronel et al. claim that humor in political messaging not only leads to a higher chance of the information being shared but also increases recipients' ability to remember political information (Coronel et al. 2021). All that suggests that humor can be used for a variety of reasons by both politicians and the general public.

At the same time, even if humor is possible to use effectively, it does not mean that it is always utilized like that. Humor often has to be designed to deliver certain messages in a specific way due to the context in which it is presented and the medium it utilizes to do so. This is evident, for example, in a dissection of a speech by a suffragist Anna Howard Shaw; she specifically framed her messaging following the norms acceptable to the 1910s United States while simultaneously ridiculing anti-suffrage arguments in a way that would call for pre-existing values of consistency in her audience (Innocenti and Miller 2016).

An influential article by Holbert et al. establishes a significant difference between the effects of different types of humor on different groups, specifically focusing on Juvenalian (more acidic and aggressive) and Horatian (more self-referential, palatable) satire. They find that individuals with a higher degree of "ability" to access political humor find Juvenalian satire more persuasive, while individuals with lower "ability" are more persuaded by Horatian humor, with reverse effects if groups are switched. Horatian comedy is also perceived to be significantly funnier by low-ability individuals, however, little difference exists in perception

between groups regarding Juvenalian satire (Holbert 2011). This suggests that different types of humor have significantly different effects on different groups, both as far as persuasive messaging and as far as its ability to appear funny. While it might appear that only the first metric matters, it can be argued that more "funny" humor is likely to be distributed more often.

Just as jokingly advocating for suffrage in the early 20th century, effectively using humor in social media can be a non-trivial task. On the example of journalists, Holton and Lewis find that they tend to use humor and generally act "informally" in social media not outright, but after getting sufficiently accustomed to their norms of behavior. Even then, however, it is mostly journalists from non-elite publications that engage in humor (Holton and Lewis 2011). Similar findings exist on politicians, who largely use social media as purely one-way information-providing "electronic brochures" due to their inexperience in handling more complex types of communications on the platforms (Hoffmann and Suphan 2017).

This suggests that politicians' use of humor in social media can be hit or miss, and depends on their degree of comfortability with the norms and informal rules of the conversation in online spaces. Additionally, it appears that some types of humor must be more suitable for some contexts than others, and variation might exist even between different social media websites. One of the most important differences between social media and conventional media is the relative lack of editorial oversight of social media. It means that social media users, must they be politicians or citizens, can engage in types of humor not typically present or even allowed in traditional media. One of these differences is the ability to use extremely dark, cynical, crude, off-color, "edgy" humor that traditional media typically tends to avoid or censors outright.

However, little to no research exists into the effect of this type of political humor. While, as noted above, there is research on "attack" humor, which can be relatively aggressive at times, it rarely looks into the degree of offense that is put in this humor. I believe that this degree of

offense is important since it can generate different degrees of emotional feedback in the recipients. This lack of research creates a large gap in knowledge in current academia on political humor, which I am aiming to cover by establishing the effect of some of the most aggressive, dark, and "edgy" humor on audiences of different political views and age groups.

3. *Theory*

3.1. Humor as a delivery service

Bail et al. find that exposure to opposing viewpoints can increase the degree of polarization. In particular, according to them, conservative voters became even more conservative after exposure to liberal viewpoints; while the same happened with liberal voters, the effect was smaller and not statistically significant (Bail et al. 2018). To some extent, the treatment used in the study by Bail et al. is dissimilar to mine, since it focused on tweets by opinion leaders, media, and elected officials. At the same time, I believe that the treatment of offensive humor I propose is likely to generate even stronger polarization since it is likely to cause a higher emotional response than relatively tamer and more politically correct messaging by the aforementioned groups.

It is also very difficult to achieve opinion change. According to a survey of almost 40 academic works and an independent experiment, even during political campaigns, the time of most heightened degree of argumentative and emotional appeals, the average opinion change in society is negligible and close to zero (Kalla and Broockman 2018). This, together with Bail et al., suggests to me that most people are relatively stubborn in their beliefs, and are not only unlikely to change them even under prolonged exposure to alternative viewpoints, but even have tendencies to entrench their beliefs further.

The effects of offensive humor are likely closely related to its call-out nature. Attack messaging is likely to generate somewhat stronger engagement in general; for example, Stromer-Galley et al. find that attack messages by political candidates are retweeted more often than advocacy messages (Stromer-Galley et al. 2018). I believe that similarly, dark humor is likely to generate a stronger emotional response from recipients than a more casual type of humor, would. When people see an undeniably hostile message, which most dark humor by nature is, they are likely

to associate it with being an attack on something. At the same time, there are different ways to present attack messages. It is unclear if dark jokes can be very effective in generating buy-in into opposing ideas given the shocking way in which they are presented – instead, they might offend the recipient of the joke, causing an opposite reaction.

Because of that, I predict that ideas behind dark humor are more likely to strongly resonate well with people already in agreement with it. While shocking, dark humor can often have a fairly strong underlying message that points out certain things rarely brought up. Typically, these messages also point at incidents that people have strong opinions on; examples of that could be war crimes, sex crimes, torture, death, and other unsettling topics. By presenting such a message, it is likely not to seem simply offensive for the sake of being offensive, but a sharp and witty critique. Moreover, given their likely nature of creating a big emotional response, they will be able to resonate stronger than casual humor, generating a higher degree of polarization against the group targeted by the humor.

At the same time, likely, if a person disagrees with the messaging, it will create a stronger polarization in favor of the group targeted by humor, and by extension, dislike towards the opposing group. Very few people want to feel like they are on the wrong side of incidents they have strong opinions on, and given that many people deeply relate to the ideas they hold or politicians they support, they may feel attacked personally. That, similarly, is likely to lead to a stronger reaction than a call-out on something milder and more "debatable" in nature.

It is important to note that casual humor can also touch on relatively taboo subjects, must it be war, abortion, or religion, so the topic alone cannot differentiate between casual and dark humor. However, dark humor is also often described as "immoral", and as going beyond the tactful ways to talk about those topics. I believe this combination to be key in establishing what makes dark humor special. It violates the moral norms that people hold as important, and it

violates them in a way that does not make them feel safe and secure. I will support this description in the sections below, by combining two theoretical frameworks of Moral Foundation Theory and Benign Violation Theory.

3.2. Finding the root of polarization

As noted above, political humor needs to carry a certain message to elicit a reaction from the recipients. There is a variety of topics humorous messages can focus on, but not all of them are similarly effective in generating polarization. Therefore, it is important to outline the causal mechanisms that underly polarization that could potentially be used in humorous messaging.

Empathy is often cited to be an important factor in the spread of political polarization. A natural assumption may tell that higher empathy can help to reduce polarization by allowing people to understand the other side better. However, much of the research claims the opposite, saying that empathy works as one of the mechanisms of polarization. Simas et al. claim that empathy does not necessarily reduce polarization, but can exacerbate it on partisan grounds. They focus on empathic concern, which they define as the "tendency to experience other-oriented emotions, such as sympathy or compassion, for another person who is in distress". According to the authors, people tend to feel the highest degree of empathy towards others from their in-group and to be particularly sensitive to offenses committed by the out-groups (Simas et al. 2020). Allamong and Peterson similarly, but more broadly, research the effect of the *emphatic ability*, defining it as a "general ability to determine the emotions of others", on affective polarization and policy support. They ran a survey experiment including a story about an individual who would be affected by the repeal of the Affordable Care Act in the US. Their findings are that partisanship plays a big role; while higher empathic ability strengthened the support towards the individual and against the policy for Democrat recipients, it produced a

10

reverse effect for the Republicans. Moreover, findings were stronger for opinions on individuals compared to policy support change (Allamong and Peterson 2021).

This suggests that the perception of relations and differences between in-group and out-group modulates the degree of affective polarization. Huddy and Yair ran an experimental study, presenting respondents with a mock story about a policy meeting on migration between a Democrat and a Republican leader. The variations included the meeting proceeding warm or hostile, and leaders managing or failing to come to a compromise. They found that the warmth of the meeting managed to decrease polarization, but a policy compromise did not (Huddy and Yair 2021). This finding seems to suggest that the main driver behind polarization is not policy, but other factors, which seems to be consistent with the findings by Allamong and Peterson. An interesting suggestion coming out of that is the empathetic concern that partisans show towards their fellow in-group members is not based on purely analytical grounds and calculation of a particular policy's impacts. Instead, they may rely on emotional cues corresponding to their pre-existing beliefs, such as opinions of their leadership, as well as perceived harm to in-group or out-group individuals. Moreover, if harm is likely to be done to an out-group individual, it might drive higher polarization, even if there is no evident direct benefit to the polarized person.

Huddy and Yair's study seems to suggest that voters care about the demeanor and behavior of their leaders, and might change their preferences accordingly. This finding is supported by Armstrong and Wronski, who tried to determine whether moral foundations and party cues can change people's views on a controversial topic, namely, restrictions on free speech. To do so, they have invoked different moral frames (harm reduction and loyalty) about hate speech about Muslims and flag burning. What they have found is that partisans are broadly likely to reduce their tolerance toward free speech when moral frames are invoked by in-party actors, even if those moral frames are incongruent with the general stances of the party. Interestingly, the

harm reduction moral frame reduced tolerance to free speech against Muslims among Republicans when invoked by in-party cues but decreased it when used by the out-party (Armstrong and Wronski 2019). This, again, appears to coincide with findings by Allamong and Peterson and suggests that at least Republicans are likely to backlash against moral suggestions when they come from the other side of the political spectrum. It is yet unclear whether this effect can be reproduced with Democrats, or more broadly, with more left or right-leaning partisans in other countries.

It's important to note that Armstrong and Wronski's study does not study polarization specifically, but more so the extent to which parties are capable of manipulating their supporters' opinions on issues using moral cues, which appears to be significant. This is important for studying polarization since numerous other studies discuss the evaluations of the moral character of opposing party supporters as an important part of the polarization process.

Tappin and McKay attempt to determine whether partisan actors view each other's moral character negatively and whether that leads to discriminatory behavior on a basis of an experimental study, assuming "moral polarization" to be an important part of affective polarization. The experimental setup used by Tappin and McKay was a variation on a prisoner's dilemma, where subjects had three options: allocating a larger sum of money to themselves, a smaller sum to everyone in their (Republican or Democrat) in-group, or the same small sum to every participant. What they found is that partisans indeed view each other's moral character heavily negatively; however, that almost does not translate into action (Tappin and McKay 2019). At the same time, it is unclear to what extent such a blatantly experimental setup does reflect real-life behavior, and if there can be far-reaching conclusions on the low correlation between polarization and action.

The relationship between partisanship and polarization can be quite complex overall, for example, West and Iyengar find that the salience of partisan identity, and its effects on polarization, can shift depending on the circumstances surrounding the individual. According to them, individuals internalize partisan identity way more during times of heightened political salience, such as right before elections. Additionally, partisanship internalization can be decreased by invoking ideas of self-affirmation. Researchers asked some of the respondents to rate themselves on a list of traits (such as "wisdom") before surveying them and found a significant shift in partisanship. Interestingly, however, shifts in partisanship did not necessarily decrease the out-party hostility and general group bias (West and Iyengar 2020). Seemingly, on the contrary, Levendusky finds that decreasing partisanship does decrease affective polarization. However, the method Levendusky did choose is invoking a common identity, instead of invoking an individualistic one. He managed that by gathering a sample of American subjects, and invoking their American national identity as a treatment. (Levendusky 2018). Therefore, it appears that different sources of partisanship generate different results at different times, making it necessary to find a deeper mechanism that invokes mutual dislike between people with opposing political opinions.

Grubbs et al. suggest that "moral grandstanding", the portrayal of one's moral beliefs as superior to others, may explain at least some political polarization. According to them, moral grandstanding used for prestige-seeking, in an attempt to gain respect and admiration from the individual's peers, did significantly correlate to more extreme ideological viewpoints, as well as affective polarization. Surprisingly, dominance-seeking grandstanding that seeks to humiliate or embarrass opponents, however, did not correlate either with extreme ideology, or higher affective polarization. One of the explanations provided by Grubbs et al. was that prestige-seeking grandstanding was comparative, and therefore, made respondents attempt to stand out by adopting more extreme views (Grubbs et al. 2020).

13

There might be an increase in the prevalence of moral issues and viewpoints in political discourse in general. Kozlowski noted a significant decline in trust in science among US conservatives over the forty years before 2022. What is interesting is that in 1980, moral conservatives were skeptical of science already, but economic conservatives expressed disproportionately high trust in it compared even to liberals. According to Kozlowski, moral conservativism has since consolidated under political conservativism, making more moral conservatives identify as political conservatives (Kozlowski 2022). While that does not suggest that moral issues became the forefront for all conservatives, it at least means that their representation in the overall discourse must have increased.

3.3. Moral Foundation Theory

Haidt claims that the moral predispositions that people have are not sourced exclusively through logical reasoning, but also social convention, and general feelings such as disgust. To test that claim, he conducted an experiment, where he surveyed people from Brazil and the US about their opinions on the morality of different scenarios with no harm. One of the examples was a family dog getting hit by a passing car, and the family eating the dog after its death; another was a woman finding an old unused flag in her closet, and cutting it up for rags to wash the floor with. What he found is that people, even when they recognized that there is no harm, would still consider such scenarios immoral. While Brazilians would broadly be more condemning of the action, the immorality claim was highly correlated with the socioeconomic status as well, with well-off residents in both countries being closer to each other than to their fellow poor countrymen, who would generally also be more judgemental. Lastly, he reports that many of the respondents would be quick to proclaim a certain action to be immoral, but when confronted with the need to explain their opinion, would struggle to provide moral reasoning, resorting to haphazardly inventing victims, and often failing to find reasons when

victims would be ruled out. However, that would typically not change their opinions on the morality of the action (Haidt 2012).

What can be drawn from Haidt is that there are significant differences between moral perceptions in various cultures, as well as across classes. Moreover, there is a perception of certain actions' morality that does not necessarily come from reasoning, but more so from feelings – certain actions just appear to be disgusting or shameful to some people. Varying preferences of liberal and conservative voters must originate somewhere. Graham et al. find that there may exist a difference in importance that liberals and conservatives assign to different moral values. They presuppose the existence of five foundational moral intuitions, "harm/care", "fairness/reciprocity", "ingroup/loyalty", "authority/respect", and "purity/sanctity". Based on four independent experiments, liberals would consistently put more importance on the first two intuitions, while conservatives did tend to use all five more evenly (Graham et al. 2009). This suggests an underlying reason for the "moral polarization" between left-leaning and right-leaning partisans. If they operate based on different moral strata, then it must mean that each side, according to another, violates certain important principles that underly their respective worldviews. If we believe Haidt, it would also suggest that those violations can invoke purely emotional responses, which do not necessarily appeal to reason. These findings are key for establishing why dark humor can be unique in its ability to elicit reactions from recipients. By disregarding the moral norms of certain societal groups, it manages to override the logical decision-making process in people that receive it, and generate polarization to the extent that more morally compliant types of humor cannot.

Koleva et al., using the same framework of five moral intuitions, find that they predict support for a wide range of political issues better, than ideology, age, gender, religious attendance, and general interest in politics. Additionally, the authors found "purity/sanctity" to be the strongest predictor of approval or disapproval for the biggest number of topics, from sexuality to

gambling. It was also very close to ideology in being the strongest predictor of abortion approval or disapproval (Koleva et al. 2012). This sounds logically correspondent to the issues that are often most divisive and polarizing in modern political discourse, which is often concerned with topics such as sexual freedoms and the sanctity of life. At the same time, the finding on abortion appears interesting, since it does directly correlate with the "sanctity of life", however, despite the high predictive power of "purity/sanctity" intuition, still loses out to ideological predisposition. I assume that abortion is one of the, if not the most controversial issue in the United States, where the research was conducted, and it is a major contention point for both the Democrats and the Republicans. As such, support or disapproval for it might have become a staple issue closely related to partisan identity in the country overall. No matter the mechanism, however, it appears that ideology can overtake pure morals in predictive power for at least very heavily discussed topics.

3.4. Benign Violation Theory

McGraw and Warren, drawing from previous literature, propose the idea that amusement or laughter originates from violations that are seen as benign or harmless. They claim that humor is often linked with two seemingly opposing conditions: aggression, hostility, or disparagement, but at the same time, also safety, playfulness, and non-seriousness. An idea that reconciles those factors is that humor also requires simultaneity, or the ability to hold two contradictory ideas at the same time. Authors claim that humorous situations arise when all three conditions are present: there is a violation, but it's perceived to be benign. An experiment study run by them tested the hypothesis based on moral violations, which found that respondents found scenarios that involved a moral violation funnier than those without one. Moreover, participants that found described situations simultaneously wrong and not wrong were significantly more likely to also find the situations funny. They were also more likely to find the situation funny if breaking the moral norm did not lead to harmful consequences, or

when their commitment to the moral norm was weak. What was additionally important is that disgust did not necessarily make people find jokes less funny (McGraw and Warren 2010). Benign violation theory's relation to people's perception of humor was additionally tested by Yam et al. What they found is that people with strong or activated moral predispositions both found humor less funny, and were less likely to produce humor others found funny, especially when humor involved benign violations. Additionally, people with strong moral identities did not tell more jokes without involving violations, ultimately simply producing less humor (Yam et al. 2019).

3.5. Linking humor and polarization

Ferguson and Ford define disparagement humor as humor that seeks to denigrate, derogate, or belittle a target, and attempt to explain the reason why it is funny. They provide a review of three theories: psychoanalytic, superiority, and social identity theory, lending support to the latter. It claims that social identities often become important in the comparisons between different groups, and personal identities, drawing from those, may feel threatened when faced with unpleasant comparisons. Disparagement humor may be a way to fight back, by providing a cultural tool to assert the positive distinctiveness of one's group. One possible example the authors bring up is a majority group feeling threatened by the advances of minority groups. To maintain their social position, they may resort to disparaging humor (Ferguson and Ford 2008).

Koszałkowska and Wróbel have tested the effect of disparagement humor, displayed in the form of sexist, racist, homophobic, or blasphemous jokes on people with five different moral foundations used in research by Graham and Koleva. They found that sexist, racist, and homophobic jokes were found less funny by people with high "harm/care" and "reciprocity/fairness" scores, while blasphemous ones were not amusing for people scoring high in "ingroup/loyalty", "authority/respect", and "purity/sanctity". What was also important

17

is that neutral jokes were found more funny, moral, and less disgusting than disparaging ones on average, with disparaging jokes creating mixed feelings. Authors suggest that it could distinguish them into a separate class of humor. One important consideration was the severity of the disparaging jokes. While authors did not measure the perception of the recipients on it, they speculated that excessively high severity could make the joke not sufficiently benign to be funny. At the same time, excessively low severity has a chance of not invoking the moral foundations in recipients (Koszałkowska and Wróbel 2019).

All of that suggests that moral foundations appear to significantly correlate with the polarization process, as they, at least to an extent, dictate the values that people hold as important, therefore being a possible root of differences in political preferences. At the same time, humor often has to rely on challenging moral foundations to be amusing, making seemingly offensive statements one of its staples. Finally, it appears that the difference in moral foundations also dictates what types of humor respondents find funny or disgusting (which might correlate with offensiveness). Given that morality appears to be fairly often invoked in political discourse, as well as to be an underlying factor of many preferences, it appears that dark humor must be capable of striking at the core of political beliefs that individuals hold, separating it from other types of humor through this mechanism.

That leads me to the following set of hypotheses:

H1: Exposure to dark humor will generate a stronger degree of affective polarization than exposure to casual humor.

H2: Joke recipients agreeing with the jokes' messaging will show more affective polarization.

H3: Joke recipients finding the jokes funny will show more affective polarization.

H4.1: Jokes recipients agree with will be shared more

H4.2: Jokes recipients find funny will be shared more

In addition, since the humor employed in the research will be presented in the form of internet memes, and social media is broadly used more by younger people who are likely to be better-versed in this type of humor, there is an additional set of hypotheses that follows:

H5.1: Younger joke recipients will find the jokes funnier.

H5.2: Younger joke recipients will agree with the jokes more.

H5.3: Younger joke recipients will share the jokes more.

It is not always trivial to figure out the "opposing group", especially if the message is not presented by any discernible member of the group. However, given that much of modern politics is presented on a left-right spectrum scale, as well as that most countries do have a limited number of predominant political actors with opposing viewpoints, it must be relatively easy for recipients to identify it. Still, to ease the task, my research design will specifically focus on a context in which determining the "opposing group" is the easiest.

4. Research design

While exposure to certain content might have a variety of effects, not all of them can be measured easily. For example, it is largely unfeasible to attempt measuring direct changes in voting behavior within the scope of this research, since it is difficult to track what things other than the treatment have affected the participants over a long period before voting. However, certain measures can be good predictors of voting behavior. One of them is affective polarization. Affective polarization can most easily be described as the degree of dislike and distrust towards the opposing political group, which can extend both to its elites and voters (Druckman and Levendusky 2019). Presumably, a voter polarized against a certain party would be unlikely to support it, and therefore, would most likely cast a vote for the other side.

The research was carried out utilizing an internet-based survey experiment with 238 participants (after ruling out unreliable results). The reasoning behind that choice was to have a low price per participant, which would allow establishing a decent sample size while simultaneously providing sufficient variation between different treatment groups and the control group. Despite that, due to budgetary limitations, the study focused exclusively on exposure to hostile humor, no matter if it was dark or casual. At the beginning of the survey, participants were asked about their party identity, being presented with four options, "Democrat", "Republican", "Independent", and "Non-major party supporter". In case the participant chose one of the latter two options, they were asked whether they would lean Democrat or Republican if they had to choose, and grouped with participants who chose Democrat or Republican options initially. In case participants chose the option "Neither" at that point, they were identified as Democrats. The self-selection produced a skew to the Democrat side during the sampling process, with 146 participants being identified as Democrats, and 92 as Republicans. The "Independent"/"Non-major party supporter" group was

20

29 people, 13 of which identified as closer to Democrats, 12 as closer to Republicans, and 4 as neither.

The survey was initially planned to be advertised via Facebook, offering participants a chance to join a lottery with a 100 USD cash prize with other surveyed people. However, assumedly due to its attempts to limit political content on the platform, and the nature of the survey, which included dark humor, the advertisements effectively were not allowed to run. That led to the second choice of platform in Amazon's Mechanical Turk. The platform allows workers to participate in surveys for a cash payment of the size selected by the "requester". Workers were recruited in two waves. During the first wave, they were offered a rate of 1 USD per response. Unfortunately, due to a technical error in the survey that occurred during the adaptation from Facebook to Mechanical Turk, approximately 100 of 250 responses in the first wave were left unusable. Because of that, the second wave of 100 responses was initiated, offering a rate of 0.5 USD per response. This decrease in rate happened because the first wave has fully depleted the budget allocated to the project, which forced making up for unusable responses with limited out-of-pocket funds. However, the quality of the data between both waves was relatively consistent. Both waves had similarly small rates of responses rejected as unreliable, and participants, on average, spent a similar amount of time on the survey, with approximately 24 minutes for the first wave, and 20 minutes for the second one.

After self-selection, respondents were randomly divided into two treatment groups, as well as a control group. The first treatment group received "dark" humor with more offensive content, poking jokes at sensitive topics such as abortion, war, and religion. The second treatment group received "casual" humor with less hostile messaging. Treatment with both "dark" and "casual" humor differed for Democrats and Republicans, with Democrats receiving jokes attacking liberal and left-wing ideas and politicians, and Republicans, respectively, jokes attacking conservative and right-wing ones. Each treatment consisted of 7 images followed by three

21

questions, asking participants if they found the joke funny, whether they agreed with its messaging, and whether they would share the joke with others (publicly or in private). The first two questions used a five-point Likert scale, while the latter had three options. Two images in each of the sets were accompanied by a manipulation check. One manipulation check asked participants to identify some kind of a name that was written on the picture (for example, one of the pictures featured a school shooter, with his name explicitly spelled out, and another one featured the name of the Twitter account that posted the joke). The second manipulation checked for the understanding of the content of the picture (for example, one of the images discussed a ban on guns, and the manipulation check asked whether guns, abortion, or migration were banned in the image). However, two out of those manipulation checks proved unreliable, with highly inconsistent answering patterns. That might be due to ambiguity in the formulation of the questions themselves, or due to the unclear nature of the images presented to the participants. In either case, those two manipulation checks were disregarded during the process of removal of unreliable answers.

Since the goal of the study is to test affective polarization, participants were asked to answer several questions about their political attitudes after receiving the treatment, or, in the case of the control group, instantly. The questions were divided into three groups, based on the methodology used by Druckman and Levendusky, who compared the ways to measure affective polarization commonly applied in the scholarship. The first method was measuring "social distance" between groups, which came in the form of questions such as "How comfortable are you having close personal friends who are of the opposing party?", and focused on measuring respondents' opinions on the other party's voters. According to Druckman and Levendusky, this type of measurement is the most unique, since it tends not to correlate with the other two methods used in the study, and generally shows a lower level of polarization (Druckman and Levendusky 2019).

The second method was a "feeling thermometer". It invited participants to rate how they felt towards different groups on a scale of 0 to 100. On this feeling thermometer scale, ratings between 0 and 49 degrees mean that you feel unfavorable and cold (with 0 being the most unfavorable/coldest). Ratings between 51 and 100 degrees mean that you feel favorable and warm (with 100 being the most favorable/warmest). A rating of 50 means you have no feelings one way or the other (Druckman and Levendusky 2019). Participants were separately asked their opinions on voters, elected officials, and parties of Democrats and Republicans respectively.

The third method required participants to attribute certain traits to voters of both the Republican and Democratic parties. They were presented with 8 traits, five of which were positive, and three were negative, with each being possible to rate on a 5-point Likert scale ranging from describing respective voters "not at all well" to "very well".

The Likert scales used throughout the survey always had every choice presented on them labeled, and attempted to stick to a similar format both in the size of the scale (which ranged between 3 and 5) and in the formulation of the questions and answers.

Jokes spreading on the internet come in a multiplicity of forms, for example, visual and text-based ones. Some jokes utilize templates that allow conveying similar messages via contextual perception. Therefore, participants a mix of text-based and visual jokes, with visual jokes taking up the majority, and some of the jokes utilized popular templates widely spread on the internet as "memes". Most of the jokes were found on the internet and were posted by real users on various social media platforms such as Reddit, Facebook, and Twitter. Some of these jokes were edited for clarity or remade in a higher resolution to be consistent with the rest of the set in terms of visual quality. A select number of jokes were produced by the researcher.

However, all of the jokes were presented as being taken from social media to reduce the unnatural effect of the treatment.

For this research, humor was heavily focused on the context of the United States, albeit alternative options such as the United Kingdom, Australia, or Canada were considered. The main reason for it is the language and cultural barrier. The study involved creating or choosing a lot of content, which limits the possible languages to English, Russian, and Kazakh. Due to largely undemocratic contexts in the countries in which the latter two languages are spoken, it is unclear what kind of political action can jokes lead to in these states. The United States appeared to be the best option out of the Anglophone countries due to my personally higher degree of familiarity with the local politics. Since much of the context will have to reference politics due to the nature of being political humor, having a good understanding of those is crucial. Additionally, securing survey respondents from the US was easier and cheaper in general due to the abundance of survey companies operating specifically in this context. Lastly, a lot of the literature on social media engages specifically with the US, which will allow me to both integrate their findings into the end theory, framework, and design, and to interact with it through end findings.

It creates a natural limitation on the generalizability of the findings since different countries might have different perceptions of humor. On the other hand, the media presence of US politics is unrivaled and is anecdotally theorized to bleed over into the political culture of other countries as well. Similarly, US popular media such as movies, books, and movies is generally consumed all around the world. That might mean a degree of connection between the "American" perception of humor and what appears to be humorous in other contexts due to sheer exposure to American humor there.

To test the hypotheses presented above, it is important to ensure that the humor that will be used in the research will achieve its goal of being shocking. This, however, is not as trivial of a task as simply making jokes on "controversial" topics. Gubanov et al. claim black humor became more popular over the years, which implies a higher degree of exposure to it, and possibly, higher palatability of such messaging. Authors link the change to an increase in reporting of morbid incidents such as deaths or disasters, while at the same time those events continue to be relatively remote to most citizens in the West, creating a desensitizing effect. (Gubanov 2018). At the same time, humor must be palatable enough so that the recipients can read into the underlying political messaging in it, and not just be grossed out by the overtly offensive nature of it.

What exacerbates the problem is the context-based differences in understanding of what is a "good" or an "acceptable" joke. Hirsch, for example, finds differences between popular jokes in English, Spanish, and Hebrew both in the ways they were constructed but also thematically. Jokes in English, for example, would more often target men, while jokes in Hebrew would often target women and minorities (Hirsch 2017). Given this research is planning to focus on the American context, it is important to explore the American approach to political humor.

Political humor, including attack humor, is a staple of American discourse. The degree of fervor in media has been specifically high during the period of Donald Trump's election campaign, as well as post the fact of his election. A wide host of actors, from regular comedians, to journalists, to talk show hosts have attempted to chip in on jokes ranging from criticizing Trump's moral character to his appearance (Kersten 2019). Figures like Biden and Obama also attracted some degree of mockery in the media. At the same time, not all of the political humor is focused on specific politicians. This leads me to believe that jokes in the research must be mixed, focusing both on individual politicians, and on the broader concepts that opposing parties disagree on, or on the followers of those parties. While jokes on those topics are also

not uncommon, they are less likely to be overused to the degree of desensitization of the recipients. There is a variety of topics that cause a lot of debate in society, but not all of them have the same degree of potential for morbid humor. As discussed in the theory section, the topics with the highest potential touch on perceived core values that people hold, such as identity and morals, as well as on violence, due to its extremity.

The jokes were selected from American websites in contexts with predominantly American userbases, which was supposed to make the likelihood of their relevance to the participants higher. While the selection process is inevitably subjective, two main criteria were employed. Firstly, jokes were selected on the variability of topics. While some topics might repeat in the sets, they are presented in reasonably different contexts with different messaging. This was done to cover most of the issues that could concern every individual respondent. Secondly, jokes were selected for ease of understanding. Most of the jokes made straightforward points, criticizing widely discussed societal issues. Even in cases when they referenced relatively lesser-known issues (for example, one of the jokes referenced Ilhan Omar, who, despite being a relatively well-known figure, is not necessarily known to everyone), they were supplemented with visuals that must have helped the participants to digest the messaging. I will elaborate on the success of those measures in the discussion section.

5. *Results*

As mentioned previously, the research focused on exposure to partisan humor, where respondents were provided with jokes that attacked the party they felt closest to. Let's start with examining the first hypothesis, which focused on the difference between exposure to casual and dark humor.

To examine this effect, firstly, a Welch t-test was used on the dyadic pairs of dark and casual humor focused on Democrats and Republicans respectively. Values for all the measurement categories were aggregated. For social distance, the sum of answers was used, with each additional point on the Likert scale adding one point to the total of each respondent's score. For the feeling thermometer, the average value of scores given to voters, elected officials, and parties for both Democrats and Republicans were taken. For traits, a sum of the respondents' answers was used, with positive traits adding their value, and negative traits subtracting their value from the total. All of the values were subsequently normalized to a range from 0 to 1.

In the tables below, SD represents the social distance measurement, FT represents the feeling thermometer, and TR represents traits measurement. Suffixes of _D and _R indicate values about the Democrats and Republicans respectively.

Republican dark and casual humor					
	estimate	statistic	p.value	conf.low	conf.high
SD	-0.026819923	-0.5491788	0.58548720	-0.12507067	0.071430827
FT_R	0.007204433	0.1651896	0.86940103	-0.08020014	0.094609003
FT_D	-0.112762726	-1.9459021	0.05783386	-0.22942955	0.003904103
TR_D	-0.051377771	-1.5325561	0.13326749	-0.11913557	0.016380030
TR_R	-0.016548645	-0.7385313	0.46335386	-0.06146147	0.028364182

Democrat dark and casual humor					
	estimate	statistic	p.value	conf.low	conf.high
SD	-0.044444444	-1.18661937	0.2383873	-0.11881778	0.02992889
FT_R	0.002773913	0.05202046	0.9586261	-0.10314305	0.10869088
FT_D	0.016066667	0.40222349	0.6884390	-0.06325146	0.09538479
TR_D	-0.011820652	-0.55573124	0.5797817	-0.05408156	0.03044026
TR_R	-0.025652174	-0.81339602	0.4185601	-0.08847504	0.0371706

The only statistically significant difference in direct comparisons of feelings about Democrats and Republicans was found between the feeling thermometer towards Democrats for Republican respondents, at $p<0.1$. The rest of the values show very high p-values, except for the trait measurement of Democrats by Republicans. To examine this effect further, let's look at the means of feeling thermometer and trait measurements for Republicans.

Republicans	Dark humor	Casual humor	Control group
SD	0.6666667	0.6934866	0.631746
FT - Rep	0.7289286	0.7217241	0.65
FT - Dem	**0.5432143**	**0.655977**	0.5049524
Traits - Rep	0.5825893	0.5991379	0.5732143
Traits - Dem	0.5100446	0.5614224	0.5125

Interestingly, while as expected, Republicans with exposure to dark humor like Democrats less, both dark and casual humor treatment recipients express a higher opinion of Democrats compared to the control group. In fact, this trend is consistent among all measurements' average values. Let's examine the same for Democrats.

Democrats	Dark humor	Casual humor	Control group
SD	0.6666667	0.7111111	0.7177778
FT - Rep	0.5138406	0.5110667	0.4958
FT - Dem	0.7433333	0.7272667	0.6748667
Traits - Rep	0.5074728	0.533125	0.52875
Traits - Dem	0.5944293	0.60625	0.569375

The picture appears somewhat different for Democrats. Both types of treatment show a higher degree of polarization in terms of social distance compared to control. The same happens to traits attributed to Republicans for dark humor treatment. The change in feeling thermometer towards Republicans is on a way lower scale compared to Republicans' feeling thermometer change about Democrats. At the same time, both Democrats and Republicans consistently show

a higher opinion of their own group after either treatment, although the effect is again muted

for Democrats. Let's test these findings for significance as well.

Democrat control group and dark humor					
	estimate	statistic	p.value	conf.low	conf.high
SD	-0.05111111	-1.4152892	0.1603711	-0.12283890	0.02061668
FT_R	0.01804058	0.3327216	0.7400992	-0.08964333	0.12572448
FT_D	0.06846667	1.6413565	**0.1041462**	-0.01438259	0.15131593
TR_D	0.02505435	1.1316885	0.2606959	-0.01891329	0.06902198
TR_R	-0.02127717	-0.6471352	0.5193535	-0.08668489	0.04413054

Democrat control group and casual humor					
	estimate	statistic	p.value	conf.low	conf.high
SD	-0.006666667	-0.1884261	0.85093471	-0.076884177	0.06355084
FT_R	0.015266667	0.3013790	0.76376495	-0.085260367	0.11579370
FT_D	0.052400000	1.1895725	0.23710442	-0.035020727	0.13982073
TR_D	0.036875000	1.8264862	**0.07084749**	-0.003193773	0.07694377
TR_R	0.004375000	0.1748607	0.86155675	-0.045288056	0.05403806

Republican control group and dark humor					
	estimate	statistic	p.value	conf.low	conf.high
SD	0.034920635	0.71472885	0.47821957	-0.06330222	0.13314349
FT_R	0.078928571	1.75672805	**0.08398319**	-0.01091333	0.16877047
FT_D	0.038261905	0.61600445	0.54048746	-0.08627592	0.16279973
TR_D	-0.002455357	-0.06136379	0.95128129	-0.08255704	0.07764633
TR_R	0.009375000	0.36762837	0.71447042	-0.04165667	0.06040667

Republican control group and casual humor					
	estimate	statistic	p.value	conf.low	conf.high
SD	0.06174056	1.6265649	**0.10894836**	-0.01414950	0.13763062
FT_R	0.07172414	1.5604631	0.12374967	-0.02015852	0.16360679
FT_D	0.15102463	3.0810314	**0.00308885**	0.05301555	0.24903371
TR_D	0.04892241	1.6027650	0.11495243	-0.01230582	0.11015065
TR_R	0.02592365	0.9785807	0.33164034	-0.02704517	0.07889246

As far as control values go, numerous p-values come close to statistical significance at the p<0.1 level, but only a few cross it.

For a measure of polarization, a difference between thermometer score aggregates for Republicans and Democrats was used. It showcases the gap between respondents' feelings about Republicans and Democrats. Interestingly, the lowest polarization was shown not by the control groups, but by the Republicans who received a casual humor treatment.

Mean polarization scores					
Dem. dark	Rep. dark	Dem. casual	Rep. casual	Dem. ctrl.	Rep. ctrl.
0.2294928	0.1857143	0.2162	0.06574713	0.1790667	0.1450476

To examine this effect, similarly, a Welch t-test was used on the dyadic pairs of dark and casual humor focused on Democrats and Republicans respectively, as well as on the control groups.

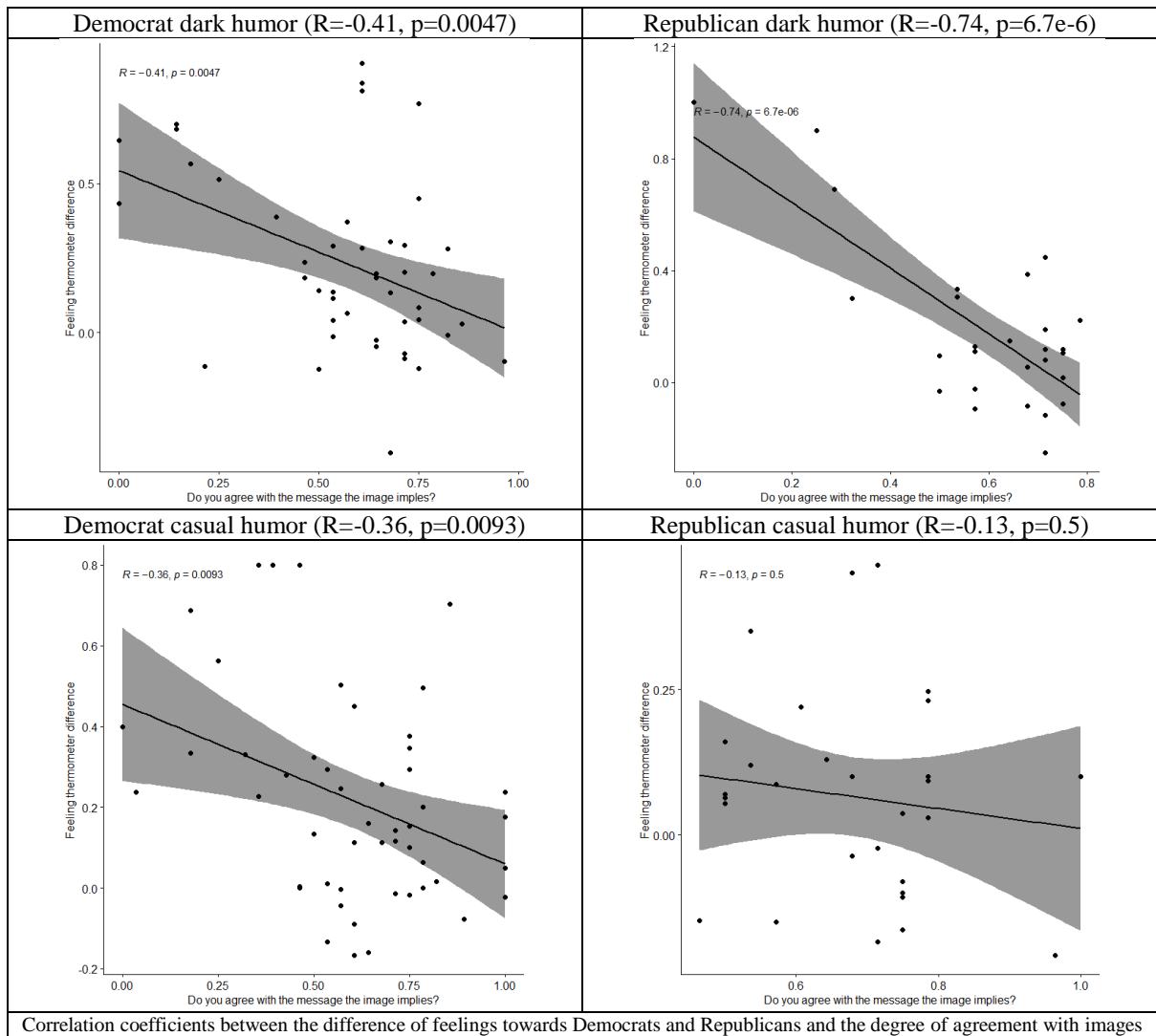|	Estimate	statistic	p.value	conf.low	conf.high
Dem. dark/casual	-0.01329275	-0.2347974	0.81490563	-0.12578382	0.099198308
Rep. dark/casual	-0.11996716	-1.8786438	0.06688940	-0.24864399	0.008709667
Dem. dark/ctrl	0.05042609	0.9129746	0.36379950	-0.05936757	0.160219741
Rep. dark/ctrl	0.04066667	0.5954097	0.55413463	-0.09636561	0.177698944
Dem. casual/ctrl	0.03713333	0.7541670	0.4525646	-0.06058193	0.134848601
Rep. casual/ctrl	-0.07930049	-1.5175780	0.13427510	-0.18378535	0.025184362

The differences in polarization between groups that received dark and casual humor are similarly not statistically significant except for Republicans who received dark and casual

humor, which is significant at p<0.1 level. This appears to reflect the abnormally low degree of polarization among the Republicans who received the casual treatment.

Examining the other measurements, we can see that on average, surprisingly, respondents claimed they are more likely to share dark humor than casual humor, with the difference being particularly prominent for Democrat respondents, holding statistically significant at a p<0.1 level. Democrats also found the dark humor funnier as a whole, in contrast to Republicans, who preferred casual jokes. Republicans agreed with casual humor way more than with dark humor, with statistical significance at p<0.05 level. Democrats showcased a similar pattern, albeit on a smaller scale, and not in a statistically significant way.

Average values	Dem. dark	Dem. casual	Rep. dark	Rep. casual
Fun	0.5357143	0.4521429	0.5076531	0.5972906
Agreement	0.5729814	0.6042857	0.5905612	0.6810345
Share	0.7515528	0.6671429	0.7576531	0.729064
Fun/stat. sig	p-value = 0.0627		p-value = 0.0797	
Agree/stat. sig	p-value = 0.5028		p-value = 0.0400*	
Share/stat. sig	p-value = 0.0929		p-value = 0.6209	

The second hypothesis states that there must be a correlation between the degree of agreement with the jokes recipients show and their degree of polarization. To do that, respondents' answers to the question of whether they agreed with the messaging of the joke were aggregated similarly to the questions on polarization. For a measure of polarization, a difference between thermometer score aggregates for Republicans and Democrats was used. It showcases the gap between respondents' feelings about Republicans and Democrats.

| Democrat dark humor (R=-0.41, p=0.0047) | Republican dark humor (R=-0.74, p=6.7e-6) |
| Democrat casual humor (R=-0.36, p=0.0093) | Republican casual humor (R=-0.13, p=0.5) |

Correlation coefficients between the difference of feelings towards Democrats and Republicans and the degree of agreement with images
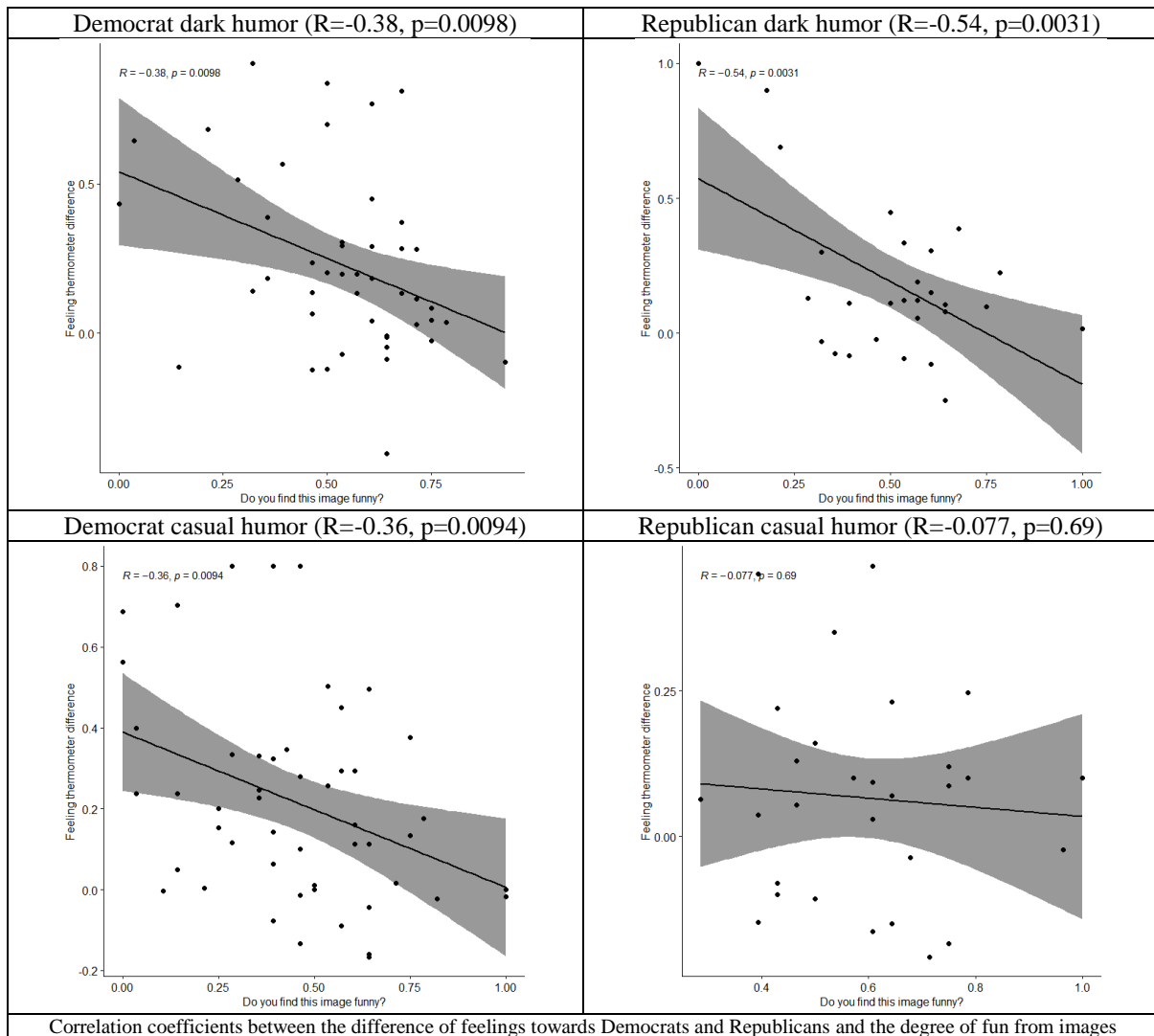
The results show that there is a strong correlation between the degree to which respondents agree with the jokes and the degree of polarization they showcased in three out of four treatment groups. The only group that did not show a statistically significant correlation was the Republican respondents that received casual humor. Interestingly, the correlation for dark humor for Republicans showed the highest explanatory power at $r^2=0.54$. Across all of the graphs, respondents who did not agree with the messaging in the images expectedly showed a higher degree of polarization.
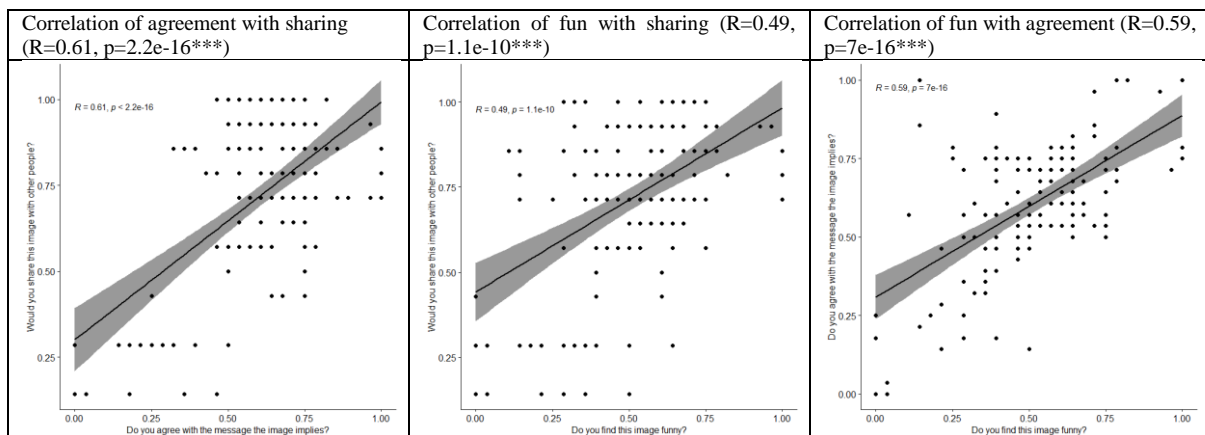
Talking about the third hypothesis, let's see if there is a statistically significant correlation between finding jokes funny, and the degree of polarization shown by respondents. To do that,

again, respondents' answers to the question of whether they found the jokes funny were aggregated similarly to the questions on polarization. In the same fashion, for a measure of polarization, a difference between thermometer score aggregates for Republicans and Democrats was used. It showcases the gap between respondents' feelings about Republicans and Democrats.

The results are strikingly similar to the results on the correlation between agreement and degree of polarization. They show that there is a strong correlation between the degree to which respondents found the jokes humorous, and the degree of polarization they showcased in the same three out of four treatment groups. Again, the only group that did not show a statistically significant correlation was the Republican respondents that received casual humor.

Democrat dark humor (R=-0.38, p=0.0098)	Republican dark humor (R=-0.54, p=0.0031)

Democrat casual humor (R=-0.36, p=0.0094)	Republican casual humor (R=-0.077, p=0.69)

Correlation coefficients between the difference of feelings towards Democrats and Republicans and the degree of fun from images

The fourth set of hypotheses, which is concerned with the likelihood of sharing the jokes, was tested on the aggregated set of data from all treatment recipients.

Correlation of agreement with sharing (R=0.61, p=2.2e-16***)	Correlation of fun with sharing (R=0.49, p=1.1e-10***)	Correlation of fun with agreement (R=0.59, p=7e-16***)

Unsurprisingly, there is a high degree of correlation (statistically significant at $p<0.001$ level) between both the degree of agreement and the degree to which jokes are found funny with the likelihood to share. A similarly high correlation exists between the degree to which jokes are found funny and the likelihood that respondents will agree with them. All three measurements, therefore, appear to closely correlate with each other. The agreement appears to have a higher impact on the likelihood to share than the degree to which the joke is found funny, however, with explanatory power $R^2=0.3721$ against $R^2=0.2401$.

Lastly, I look at the fifth set of hypotheses, concerning the age of the recipients. To test this set of hypotheses, first, responses across all treatment groups were aggregated to test if there are broad trends across the entirety of the set.
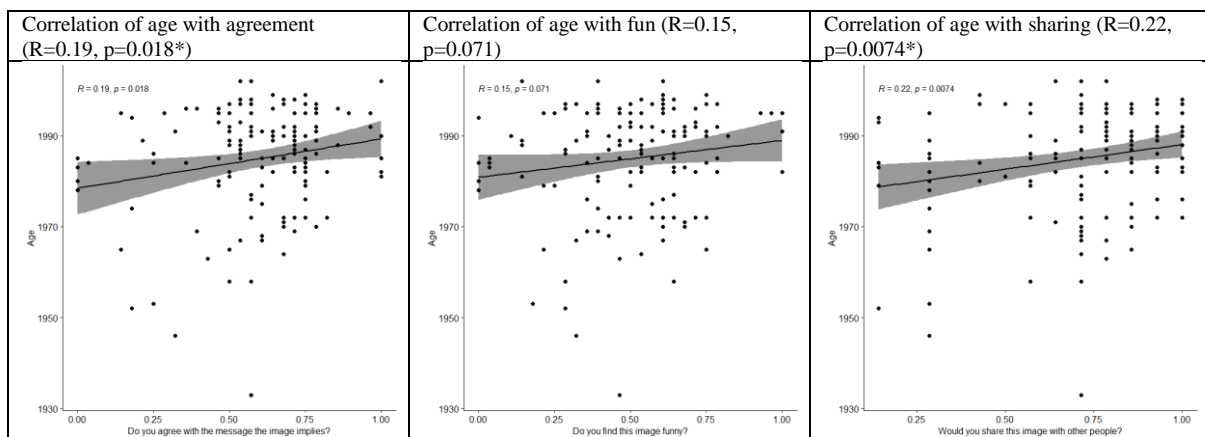
As expected, there is a statistically significant correlation on a $p<0.05$ level on the likelihood to share the jokes, as well as on the degree of agreement with them. Age also correlated with how funny the recipients found the jokes, albeit at a lower significance level of $p<0.1$. However, the correlation proved fairly weak in all of the cases, with the explanatory power ranging between $r^2=0.0225$ for the degree of fun and $r^2=0.0484$ for the likelihood to share. The correlation also did not hold statistical significance for non-aggregate data, which can be attributed to relatively small individual sample sizes of different treatment groups.

6. Discussion and implications

The conducted research did not manage to reliably confirm that the effects of dark humor are significantly different from the effects of casual humor. This, in itself, does not mean that dark humor, therefore, is no different from casual humor. There is a variety of possible explanations for this result that I am going to discuss now.

First of all, the research had several limitations. The sample that was possible to acquire given the budget of the research was relatively limited, with certain treatment groups having less than 30 participants. In particular, that was noticeable with the participants of the conservative slant. That, naturally, limited the ability of the research to provide high explanatory power. A straightforward suggestion for future research could be to explore the question again, using a larger sample of the population. It was difficult to change the research design to include a lesser number of groups since representatives of each ideological slant inherently required at least two treatments alongside a control group. One possible improvement could be to focus only on one side of the political spectrum, namely, Republicans or Democrats. The problem with this is related to the second limitation. The survey was conducted online, which gave little tools for checking participants' party affiliation or ideological slant outside of self-reporting. Even if the call would be explicitly made for exclusively Republican or Democrat supporters, nothing would prevent participants from simply lying, given the monetary incentive behind the survey in both Facebook and Mechanical Turk variations.

A possible solution could be to use another platform, such as Lucid, which collects demographic data from its subjects and allows to construct samples based on it. In the end, the choice to limit the sample to only one partisan group was not made, since it would reduce the findings' applicability to the broader population, as well as disallow potentially interesting comparisons of reactions to various types of humor coming from ideologically distant

respondents. Based on it, the main suggestion remains to conduct a larger study that would involve a higher number of subjects.

Another issue could arise from the nature of the treatment. Humor is inherently subjective and difficult to classify, which makes it difficult to provide appropriate treatment. The sets of treatments were selected entirely by me, and while certain input was received from other students from various social groups and specializations, the final decisions still came from a single person. It would be valuable to track the effectiveness of treatment by asking questions such as "Do you find this image immoral/offensive", however, given the attempt of the research to present the jokes simply as a selection of political humor, it ran a risk of priming the subjects to perceive the humor as specifically aiming to be offensive. Three potential problems with the treatment could be identified based on this. Firstly, the treatments may have simply not differed enough. It is possible that either the "dark" humor was not sufficiently dark to distinguish itself from the "casual" one, or that the casual humor was overtly dark and was not much different from the dark one. This is a complicated problem to solve. Using overtly offensive humor, while it would be easier to distinguish from more casual type, does not necessarily represent the real dark humor that is spread in the media, therefore, rendering the findings less applicable. It is also the case that most casual humor that attacks a partisan group or its beliefs can be perceived as offensive, even if it makes a relatively light-hearted point. Possibly, more experimentation with different variations of jokes could be beneficial to establish a better understanding of what jokes an average person finds offensive or dark.

Secondly, the treatment might have been difficult to understand. At least two manipulation checks produced highly varied answers, which might suggest that the survey respondents did not get some of the jokes immediately. While the problem may lie in the manipulation checks themselves, it is difficult to fully rule out the previous explanation either. A lot of the humor, given its politically-driven nature, required at least some familiarity with the US political

37

landscape. While all of the respondents were from the United States, it does not necessarily mean that they were well-versed in politics. This was foreseen, and an attempt at mitigation of the effect was taken by choosing topics broadly discussed in society, as well as relatively straightforward jokes. However, it is still possible that the perception of what is widely discussed and easily understandable coming from a political scientist differs from the understanding of an average person. The way to improve this would be to run a trial experiment, or, possibly, a series of trial experiments that would determine the degree to which the jokes are digestible. However, due to the same budgetary limitations that constrained the sample size, it was largely unfeasible.

Lastly, it might be that the treatment was simply insufficient. Each respondent, except for the control group, received seven images. While there is no gold standard to the amount of treatment that should be used, and there are reasons why excessive treatment might also be detrimental, it might simply be that this amount of treatment is too small to shift the perceptions of recipients about the other side in a meaningful way. Granted, an average survey response took a little longer than 20 minutes already, which was two times longer than the advertised 10-minute length of the survey. Expanding the survey further would likely mean that each respondent would have to spend even more time on it, which can be problematic for a couple of reasons. Firstly, survey respondents were already somewhat underpaid compared to Amazon's recommended rate per hour for Mechanical Turk workers. Expanding the survey without changing the payment could have damaged the quality of the data. For a test, I have run a small wave of 25 responses with a payment of only 0.25 USD per response that was not used for the analysis. The average response time for that wave was only approximately 10 minutes, two times lower than for 0.5 USD and 1 USD. Therefore, underpaying respondents by using a big survey could have led to less attention being paid to the questions and haphazard responses, among other issues. If the payment was adjusted to be appropriate, it would have

shrunk the already suffering sample size, which would have been highly damaging to the possibility of gaining valid results.

All of these limitations lead me to believe that the lack of the results for the first hypothesis in this research does not mean that the question is not worth exploring, given the theoretical backing that was provided in the earlier sections.

One finding that came close to the statistical significance at $p < 0.05$ level was the difference in the aggregate feeling thermometer about Democrats between Republicans who received dark and casual humor treatments. Another finding that received a relatively low p-value ($p = 0.13$, statistically insignificant) was the difference in trait assessment of Democrats between Republicans who received dark and casual humor treatments. This might appear to suggest that Republicans are more sensitive towards offensive humor than Democrats. This is also supported by the finding that Democrats enjoyed dark humor more than casual humor, while the Republicans did the opposite, with both findings significant at $p < 0.1$ level. Republicans that received dark humor also reported higher scores on the feelings thermometer towards other Republicans compared to the control group on a statistically significant ($p < 0.1$) level, something casual humor came close to, but without crossing the threshold.

At the same time, despite potential sensitivity, Republicans increased their attitude about Democrats when receiving casual humor. This might suggest that there is some degree of persuasiveness behind humor since the casual jokes Republicans received still did disparage their party. Despite that, their reaction was suggestive of agreement with the points made in the jokes, instead of anger on the other side of the partisan aisle.

Coming to the second hypothesis, strong evidence was found that the degree to which respondents agree with the joke correlates with the degree of polarization they report. Overall, if respondents agreed with the jokes, they showcased a lower degree of polarization. What is

important to note is the trite remark that correlation does not mean causation. There is a reasonable chance that the causal mechanism worked in reverse, and the reason why certain respondents disagreed with the jokes was their pre-existing high level of polarization. This explanation sounds no less reasonable than that the treatment had an effect, and therefore, suggests a need for additional research on the topic. It is important to note that again, Republicans who received casual humor stood out since the correlation between their agreement with the jokes and the degree of polarization was very weak, albeit slanted in the same direction, and statistically insignificant. Both partisan groups, on average, agreed with the casual humor more, however, this relationship was significant only for the Republicans.

The results for the third hypothesis tell a similar story. There is a strong correlation between the degree to which the jokes are found fun and the degree of polarization, and this correlation is also negative, with people who reported the jokes to be funny being less likely to exhibit high polarization. Similarly, it might be the case that the causal mechanism is reversed, and more polarized people simply do not find jokes making fun of their party funny. Again, the Republicans receiving casual humor did stand out, not reporting a statistically significant correlation.

Both of these results provide some food for discussion. Much of it is in the possibilities for future research. One possible design for future research could measure the degree of polarization before and after the treatment, which was out of the scope for the purposes of this study. This could help to establish the causal direction better, as well as to separate the effects of the fun and agreeable jokes. Another important suggestion was already discussed in the research design section, and it lies in testing not only jokes that oppose the ideological views of the recipients but also align with them. It might be that those jokes might have a different, less straightforward effect on the polarization process.

The fourth hypothesis gave very expectable results. People are more likely to share things that they agree with and find funny. What is interesting is that agreement with the joke appeared to be more important as an explanatory factor than the degree to which the joke was funny. That implies that people do in fact use jokes as messaging vehicles, and not only as entertainment, and that this metric might be even more important to them than bringing enjoyment to others. This, indirectly, suggests a need for additional research in the field of political humor, focusing on the ways jokes and memes are used by the population. As noted in the literature review, it is already known that people can receive information through humor coming from sources such as argumentative late-night shows (Fox, Koloen, and Sahin 2007). What is crucially different here is that while very few individuals have access to TV show production, almost everyone is capable of sharing jokes online. If the public as a whole, to a significant extent, perceives spreading humor as a task of informing others, it might imply both newfound importance of online humor, as well as give hints as to how and why it spreads.

The last hypothesis discussed the age of the participants. While age did statistically significantly correlate with the degree to which the jokes were found fun or were agreed with, as well as the likelihood to share them, all of those correlations were fairly weak, with relatively little difference between older and younger participants. This finding is also important. Social media and the internet as a whole are often characterized as mediums that are primarily used by younger generations. While this survey did not try to collect statistics on older generations' participation in social media, it shows that they are capable of understanding the humor delivered through novel platforms in the relatively novel format of memes. They also are almost as likely to share it as the "perpetually online" younger generations in case if given the opportunity. At the same time, it is important to note that the sample used for this research was not randomly chosen from the broader population, but was a convenience sample coming through an online platform. It is possible that on average, at least older people working on such

an online platform must be more well-versed with the internet compared to their other peers. While there is no way to know that, this may also suggest a need for additional research that will sample people through other methods that are less likely to include people who use the internet as often.

7. Conclusion

Political humor remains to be an important topic of discussion, and one of the ways millions of people engage with the political process on a daily basis. It comes in a variety of forms, transforming together with societal changes and available mediums. This study attempted to establish an answer to the question of whether there is a difference between dark and casual humor in generating affective polarization. While, ultimately, it was not able to establish a clear, statistically significant difference between dark and casual humor for this purpose, it still managed to produce a range of findings that can be valuable and serve as foundations for future research.

The literature often discusses the differences between various political groups and their supporters, often broadly dividing them into various spectrums, such as left-wing and right-wing. One of the foundational puzzles that concern a variety of research topics is the difference in the way how supporters of different political ideologies see the world. This study contributes by establishing certain differences and similarities in humor perception between Republicans and Democrats. It appears that both of the groups, when faced with humor that disparages their side, tend to feel more warmly about it, but this effect holds stronger for the Republicans. Republicans also tend to react better to casual than to dark humor, even going as far as improving their opinion about the other side significantly, which is something that cannot be said about Democrats, who mildly prefer dark humor, but do not tend to change their feelings about the other side as much.

While the United States is not representative of worldwide politics, and while there may be significant differences between its political landscape and the political landscapes of the other states, there is some importance to those findings. First of all, the United States serves as a microcosm of polarized politics due to its strong two-party system. Almost 90% of the

participants readily identified themselves with one of two major parties, and only a bit less than 2% did claim to be fully neutral. It makes the country a perfect example for research on polarization that can be later replicated in other contexts. Secondly, the United States is undeniably a cultural powerhouse, whose political developments are closely monitored around the world. Oftentimes, the sociocultural debates sparking in the US manage to migrate to other contexts, as it happened, for example, with the BLM movement, which gained momentum in Europe, despite starting around an incident that happened in the USA.

This suggests that the Republicans and Democrats might be at least somewhat representative of the political right and left in other states, and their preferences in terms of humor might be applicable there as well. This finding is not only interesting from the academic standpoint but also has practical applications, as politicians increasingly include humor in their campaigning, especially on the internet. The nature of the treatment used in this study closely approximated the type of humor prevalent on social media platforms, making it particularly relevant in the changing media landscape.

This project did not exclusively focus on the differences between dark and casual humor but also produced several other findings. It focused on the other predictors of polarization, such as the degree to which the jokes' messaging finds agreement in recipients, as well as the degree to which recipients see jokes as funny. For these two predictors, it finds a strong correlation between the measurements of polarization and degree of agreement or fun respectively. This finding, while it might seem obvious at the first glance, gives way for further research. The causal link within this correlation is not established, and it is possible both that humor causes polarization, or that polarization changes what humor is perceived as funny or worthy of agreement. It is possible that both of these effects hold, and reinforce each other. Given that massive numbers of people interact with political humor on the daily basis, either of these causal mechanisms could be very important. In particular, a lot of researchers theorize that

people fall into online "echo chambers" that can radicalize their opinions, bringing them to showcase extreme affective polarization. Often, this research does not specify in what forms can the content of those echo chambers come, and how exactly it manages to increase polarization. If the agreement with humor increases polarization, and polarization, in turn, makes people seek out humor they agree with, seemingly innocuous online communities that focus on humor may contribute to the process, bringing people to rabbit holes of echo chambers. This is only one of the possible scenarios and implications, and it requires further research to become more than speculation, but it showcases the contribution of this study to the gaps in the scholarship.

One more finding of this study is that people tend to share humor they find funny or agree with a higher likelihood. While this might seem to be obvious, an interesting part of the finding is that the level of agreement appears to be more important than the degree to which the joke is perceived as funny. This might mean that the patterns through which humor is shared by the population are dictated by its feeling or understanding that jokes carry ideological content, and implies that people, consciously or subconsciously, use humor as a method of information transmission. If that is the case, it invites additional research. Some research has already focused on the informative function of certain humorous mediums, such as TV shows. However, there are two major differences here. Firstly, internet memes typically come in a way shorter form than a TV show or similar media, consisting of one or two pictures, and therefore, their ability to carry information must be different. Secondly, internet memes can be easily generated and shared virtually by anyone, meaning that in case they manage to spread information effectively, it gives a powerful tool to the hands of the population that breaks away from the monopoly of a few media conglomerates.

The last finding of the research is that while there is a correlation between age and the degree to which respondents find internet memes funny or agree with them, as well as between the

45

likelihood to share them, this correlation is quite weak. Older people are not significantly different from younger people in their perception of internet humor, must it be casual or dark. This is important since it clashes with the often-cited idea that social media and other parts of the internet where memes can proliferate are primarily dominated by young people. It appears that older people can understand internet humor, find it funny, and be willing to share it at relatively similar rates. Aside from clashing with the scholarship, this also can have some implications for the social media campaign management of politicians who seek to find the support of various groups.

In the end, while this project had a series of limitations due to budgetary constraints that were outlined in the previous chapters, I believe that it managed to produce useful findings that can be employed in future research, and as such, will serve as a valuable contribution to the body of scholarship on political humor.

References

Allamong, M. B., & Peterson, D. A. (2021). Screw Those Guys: Polarization, Empathy, and Attitudes About Out-Partisans. Political Psychology, 42(3), 349-369.

Armstrong, G. M., & Wronski, J. (2019). Framing hate: Moral foundations, party cues, and (in) tolerance of offensive speech. Journal of Social and Political Psychology, 7(2), 695-725.

Bail, Christopher A., et al. "Exposure to opposing views on social media can increase political polarization." Proceedings of the National Academy of Sciences 115.37 (2018): 9216-9221.

Becker, A. B., & Waisanen, D. J. (2013). From funny features to entertaining effects: Connecting approaches to communication research on political comedy. Review of Communication, 13(3), 161-183.

Becker, Amy B. "Political humor as democratic relief? The effects of exposure to comedy and straight news on trust and efficacy." Atlantic Journal of Communication 19.5 (2011): 235-250.

Borah, P. (2016). Political Facebook use: Campaign strategies used in 2008 and 2012 presidential elections. Journal of Information Technology & Politics, 13(4), 326-338.

Brandenberger, D. (2009). Political humor under Stalin: An anthology of unofficial jokes and Anecdotes. Slavica Publishers.

Coronel, Jason C., et al. "Political Humor, Sharing, and Remembering: Insights from Neuroimaging." Journal of Communication 71.1 (2021): 129-161.

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization?. Public Opinion Quarterly, 83(1), 114-122.

Ferguson, M. A., & Ford, T. E. (2008). Disparagement humor: A theoretical and empirical review of psychoanalytic, superiority, and social identity theories.

Fox, J. R., Koloen, G., & Sahin, V. (2007). No joke: A comparison of substance in The Daily Show with presidential election campaign. Journal of Broadcasting & Electronic Media, 51(2), 213-227.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. Journal of personality and social psychology, 96(5), 1029.

Grubbs, J. B., Warmke, B., Tosi, J., & James, A. S. (2020). Moral grandstanding and political polarization: A multi-study consideration. Journal of Research in Personality, 88, 104009.

Gubanov, N. N., Gubanov, N. I., & Rokotyanskaya, L. O. (2018, November). Factors of Black Humor Popularity. In Proceedings of the International Conference on Contemporary Education, Social Sciences and Ecological Studies (CESSES 2018). Series" Advances in Social Science, Education and Humanities Research (Vol. 283, pp. 379-383).

Haidt, J. (2012). The righteous mind: Why good people are divided by politics and religion. Vintage.

Hakoköngäs, E., Halmesvaara, O., & Sakki, I. (2020). Persuasion through bitter humor: Multimodal discourse analysis of rhetoric in internet memes of two far-right groups in Finland. Social Media+ Society, 6(2), 2056305120921575.

Haynes, N. (2019). Extreme Speech| Writing on the Walls: Discourses on Bolivian Immigrants in Chilean Meme Humor. International Journal of Communication, 13, 21.

Hirsch, G. (2017). Humor appreciation in the digital era. *Lingua*, *197*, 123-140.

Hoffmann, C. P., & Suphan, A. (2017). Stuck with 'electronic brochures'? How boundary management strategies shape politicians' social media use. Information, Communication & Society, 20(4), 551-569.

Holbert, R. Lance, et al. "Adding nuance to the study of political humor effects: Experimental research on juvenalian satire versus horatian satire." American Behavioral Scientist 55.3 (2011): 187-211.

Holton, Avery E., and Seth C. Lewis. "Journalists, social media, and the use of humor on Twitter." Electronic Journal of Communication 21.1/2 (2011): 1-22.

Huddy, L., & Yair, O. (2021). Reducing affective polarization: Warm group relations or policy compromise?. Political Psychology, 42(2), 291-309.

Innocenti, Beth, and Elizabeth Miller. "The persuasive force of political humor." Journal of Communication 66.3 (2016): 366-385.

Kalla, Joshua L., and David E. Broockman. "The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments." American Political Science Review 112.1 (2018): 148-166.

Kersten, H. (2019). America's faith in the laugh resistance–popular beliefs about political humor in the 2016 presidential elections. Humor, 32(2), 299-316.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. Journal of research in personality, 46(2), 184-194.

Koszałkowska, K., & Wróbel, M. (2019). Moral judgment of disparagement humor. Humor, 32(4), 619-641.

Kozlowski, A. C. (2022). How conservatives lost confidence in science: the role of ideological alignment in political polarization. Social Forces, 100(3), 1415-1443.

Levendusky, M. S. (2018). Americans, not partisans: Can priming American national identity reduce affective polarization?. The Journal of Politics, 80(1), 59-70.

McGraw, A. P., & Warren, C. (2010). Benign violations: Making immoral behavior funny. Psychological science, 21(8), 1141-1149.

Pi-Sunyer, O. (1977). Political humor in a dictatorial state: The case of Spain. Ethnohistory, 179-190.

Simas, E. N., Clifford, S., & Kirkland, J. H. (2020). How empathic concern fuels political polarization. American Political Science Review, 114(1), 258-269.

Stromer-Galley, Jennifer, et al. "Tweeting the attack: Predicting gubernatorial candidate attack messaging and its spread." International Journal of Communication 12 (2018): 22.

Tappin, B. M., & McKay, R. T. (2019). Moral polarization and out-party hostility in the US political context. Journal of Social and Political Psychology, 7(1), 213-245

Verhulsdonk, Iris, Alessandro Nai, and Jeffrey A. Karp. "Are Political Attacks a Laughing Matter? Three Experiments on Political Humor and the Effectiveness of Negative Campaigning." Political Research Quarterly (2021): 10659129211023590.

West, E. A., & Iyengar, S. (2020). Partisanship as a social identity: Implications for polarization. Political Behavior, 1-32.

Yam, K. C., Barnes, C. M., Leavitt, K., Wei, W., Lau, J., & Uhlmann, E. L. (2019). Why so serious? A laboratory and field investigation of the link between morality and humor. Journal of Personality and Social Psychology, 117(4), 758.