Memes to an End: A Look into what makes a Meme Offensive[*]

Yehia Elkhatib[1,2][0000−0003−4639−436X] and Kieran Hill[1]

[1] School of Computing and Communications, Lancaster University, UK
`mail@kieranhill.xyz`
[2] School of Computing Science, University of Glasgow, UK
{`firstname.lastname`}`@glasgow.ac.uk`

Abstract. Using memes is a popular method of online communication, including both lighthearted and dark humor as well as polarizing subjects. But what makes a meme offensive? Does this depend on a meme's content or format? Does a meme offensiveness vary across audiences? In order to answer these questions, we recruit 316 participants to score two sets of meme variants in terms of their offensiveness. We augment this with interviews to extract additional insights.

Keywords: Memes · Internet meme · Offending content · Online abuse · Media consumption · Social medium

1 Introduction

An *Internet meme* (or just *meme*) is a multimodal medium of communicating a certain message using the juxtaposition of textual and visual formats. This commonly takes the form of a brief comment over an image referencing pop culture, politics, or social issues. An example of a meme is shown in Figure 1.

Memes have been very widely used in social media of various types and in innumerable languages [1]. The use of memes has helped accelerate the formation and evolution of Internet subcultures of different types; *cf.* [3].

Anecdotally, memes are quite efficient at delivering a complex message in a succinct way. Generally, this is due to the mix of text and visual content (in image or video format), which allows for almost immediate portrayal of context and illustration of contrast. This also allows low-effort embedding into trending phenomena, as well as crossing cultural and language barriers.

The significant role of memes in modern online culture and society in general is well documented [3,4,21]. We are interested in the common use of memes by Internet subcultures that exhibit counter-cultural, abusive or hate attitudes, such as far-right subreddits and politically incorrect 4chan boards. Given such prominent use in social media where it is easily exploited to spread misinformation [24]
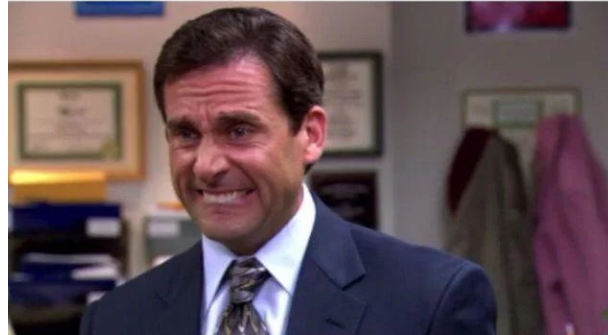

Fig. 1. An example of a meme.

even mobilizing voters [11,20], we are motivated to identify the characteristics of a meme that enables it to deliver a potentially inflammatory or offensive point of view. We are also motivated to investigate how this varies across audiences of varied demographics.

We carry out a randomized controlled trial to answer the above questions. We recruit a total of 316 participants to review two different groups of potentially offensive memes, and ask them to quantitatively score their reaction towards them. We also invite them to provide qualitative feedback. Consequently, we are able to draw some fundamental insights into the nature of memes. This is the first study of its kind to try to isolate, using mixed methods analysis, the characteristics of a meme that makes it palatable or offensive to different audiences.

In the remainder of this paper we will be using politically incorrect and potentially inflammatory language as and when context dictates. Section 2 briefly surveys related work. Section 3 describes our methodology for collecting memes, recruiting participants, and designing the survey. Section 4 presents the results of our study. Section 5 reflects on the results and discusses their wider implications. Section 6 concludes and outlines future work.

2   Related Work

Memes play a significant role in modern online culture [3,4,21] and indeed in society at large. For instance, several studies have looked into the effect of memes on the rhetoric and voter mobilization of alt-right political groups during the 2016 US Presidential Election [10,20,11,14,18].

Various works have been published on the spread of memes in online media. Shifman [22] studied features that help a creative derivative culture around YouTube memes. Guadagno et al. [9] analyzed emotional response to memes and what contributes to one *going viral*, while Huntington [12] looked into a meme's *persuasiveness* and effect on online political engagement. Neither study investigated a meme's potential to cause offense. Zannettou et al. [27] followed the spread of memes across online platforms, and observed the popularity of memes in racist online forums and fringe Internet subcultures. Johann and Bülow [13] investigated reasons for uptake in meme use, and attributed success to adoption by journalists. They also highlighted how novelty of meme derivatives through image editing increases diffusion. Lonnberg et al. [15] built models of meme spread, incorporating the users who help it spread and their willingness to do so. Other studies focused on developing methods to automatically recognize memes [2,19,23].

Yoon [26] found deep rooted racist tendencies in using memes, using color-blindness as pretext without acknowledging the adverse effects of the implied racist message. Williams and Dupuis [24] used memes to study the correlation between political affiliation and levels of personal openness and dispositional affects.

Despite all of this interest, however, our work is the first (trial or observational) to dive into why some memes could be more offensive than others.

3   Methodology

We utilized a set of 'real' memes collected from Reddit, and complemented them with synthesized ones in order to vary a certain feature of the meme whilst controlling the others. We designed a survey to collect data on perceived emotions from a large number of user of different backgrounds.

3.1   Meme collection and creation

We collected a set of memes from the /r/ImGoingToHellForThis subreddit where offensive content is typically posted. The memes, collected in November-December 2019, covered a variety of sensitive subjects such as politics, religion, and race. We attempted to select a variety of meme formats to appeal to different audiences. However, we aimed *not* to choose memes that are highly topical in order to minimize the effect of context or *trendiness* on meme perception.

We focus in this study on the two main features that constitute a meme: **image format** and **textual message**. In order to identify the degree to which each can play on perception of offense, we vary one while keeping the other the same as in the original meme collected from Reddit. We used MemeGenerator.com to alter either the format or text to include in the survey variants. When text is changed, a message from another /r/ImGoingToHellForThis meme is used. Examples are shown in Figures 2–3, with the full list included in Table 1. All memes could be accessed at https://github.com/kieranhill97/Memeology/.

Table 1. A summary of the memes used in each of the survey sets as well as the themes they discuss.

Q#	Survey A Meme	Theme	Survey B Meme	Theme
1	I'm kind of retarded	Social niceties	I'm kind of retarded	Web culture
2	Pro gamer move	Atomic bombings	Pro gamer move	Murder, Gaming
3	Keep your secrets	Infidelity, Domestic abuse	Keep your secrets	Abandonment
4	Starter packs	Sexuality	Starter packs	Race
5	Daily struggle	Death, Parenthood	Star Wars good vs evil	Death, Parenthood
6	Xbox Series X Parodies	Religion	Xbox Series X Parodies	9/11
7	Crying woman	Disability, Poverty	African kid crying with a knife	Disability, Poverty
8	Troll quotes	Assassination (Lincoln)	Troll quotes	Assassination (Kennedy)
9	Batman slapping Robin	Poverty	I worry about you sometimes, Candice	Poverty
10	I am speed	Suicide	I am speed	Race
11	I don't want any damn vegetables	Disability (child)	I don't want any damn vegetables	Disability (sibling)
12	Plants vs Zombies	Atomic bombings	Rice Krispies	Atomic bombings, Race
13	She can't do that	Gender, Religion	She can't do that	Gender, Trump
14	Roll safe	Death	When Your Parents Ask Where All Your Money Went	Death
15	What's funnier than 24	Nazism	What's funnier than 24	9/11
16	Careful, he's a hero	Gender	Careful, he's a hero	Race, Murder
17	Daily struggle	Slavery	Daily struggle	Murder, Cancer
18	Woman yelling at a cat	Gender	Who would win	Gender
19	Nut button	Terrorism	Highway exit turn	Terrorism
20	Surprised Pikachu	Terrorism	Surprised Pikachu	Abuse

Fig. 2. The meme in Q2 where the text is modified, but the format remains the same.

Fig. 3. The memes in Q7 where the format is modified, but the text remains unchanged.

3.2 Survey design

We used a survey questionnaire as the tool to collect input specifically due to its inherent nature of disconnecting the participant from the researcher and also from other participants (as in a focus group). The aim was to help the participants answer freely without feeling implicit pressure to select a politically correct answer when faced with a polarizing topic.

The survey had 2 variants, each with 20 memes. The size was calibrated to sustain participant interest throughout the survey and elicit earnest responses. Upon start, each participant was assigned one of the 2 sets based on a simple round robin protocol. The corresponding questions in each survey matched each other in either visual format or textual content. Variation was made between sets to identify the characteristic of the meme that has potential to cause more

offense. Each question asked the participants how offended the meme made them feel on a 5-point Likert scale between 'Not offensive at all' and 'Extremely offensive'. Subjects were informed at the beginning of the survey that the study was intended to offend and that the memes do not represent the views of the researchers. The survey was IRB approved.

At the end of the survey, participants were asked to indicate if they are willing to answer follow-up open-ended questions.

3.3   Participant recruitment

Between February 23rd and 26th 2020, different social media (namely Reddit, Facebook, and Twitter) were used to spread the survey in an attempt to make the group of participants as diverse as possible in terms of online culture. We also recruited locally on our university campus. Demographic information was collected for summary statistics but all responses were anonymized.

Participation was limited to a single response linked to a Google account in order to prevent multiple responses from skewing the data. This restriction limited participation to those with a Google account and were willing to use it to sign in. There were 550 visits to the survey landing page, 320 of which completed the survey (a response rate of 57.5%). 4 results had to be disregarded due to incomplete data, resulting in 316 valid responses. Of these, 37 participants provided further input in the form of answers to our open-ended questions.

4   Result Analysis

In this section we present and examine the results of the study.

4.1   Participant demographics

First, we inspect how our round robin assignment of survey affected the participant demographics of each survey set. This resulted in fairly equivalent sets, as is depicted in Figures 4–5. The only observable disparity is in the frequency of meme usage, but the difference is not statistically significant.
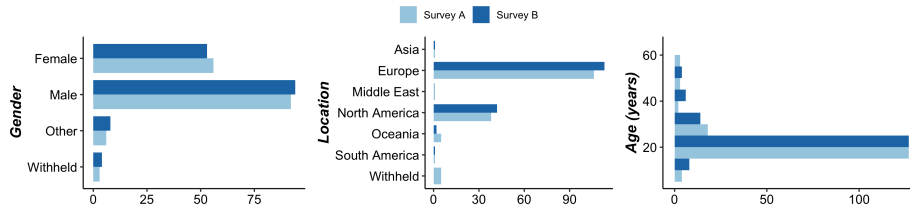

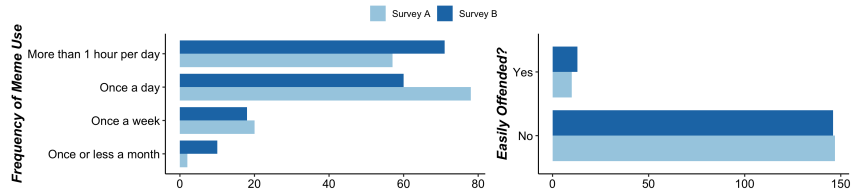**Fig. 4.** Demographics of the participants over the two randomly assigned surveys.

Fig. 5. Behavior and attitudes towards memes as self-reported by the participants.
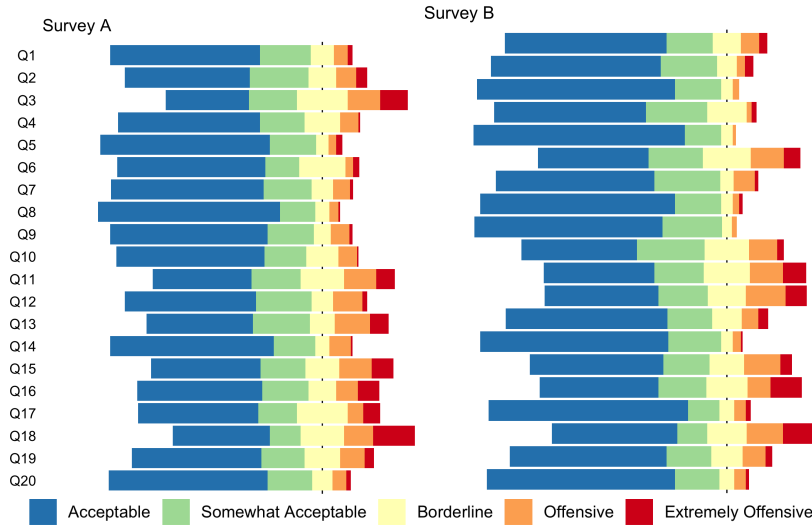

Fig. 6. A comparison of the aggregated response scores of the two survey sets.

4.2 Overall comparison

To systematically identify the difference in scores between corresponding questions of the two surveys (shown in Figure 6), we used the t-test and Mann-Whitney-Wilcoxon methods. Both are non-parametric and produce statistically significant results [25]. We found the results of the t-test to be a little more sensitive to differences in survey scores (Figure 7). This method identified the following questions as those with highly significant differences across surveys ($p < 0.01$): Q3, Q6, Q9, Q12, Q13, and Q17. Furthermore, the ones with significant differences ($0.01 < p < 0.05$) were identified as: Q2, Q5, and Q10. We now discuss our observations that account for these differences.

4.3 Form vs. substance

Participant response seems to be associated with the message of a meme more than with its visual aspects. This can be particularly observed in Q7, Q8, Q11,
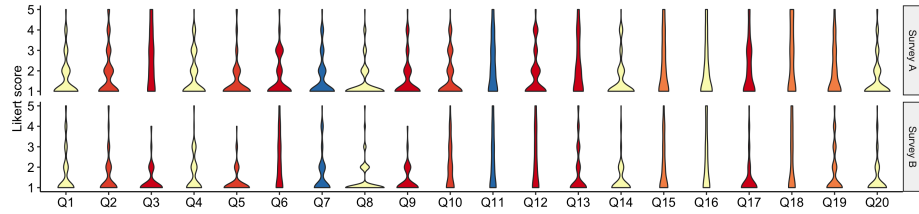
Fig. 7. Violin plots of the collected responses, color-coded in how different they are between surveys. Red indicates significant divergence between the results of the two surveys, followed in order by orange, yellow, and blue which marks insignificant divergence.

Q18 and Q19 where the message was constant while the meme changed. Exceptions are Q5, Q9, and Q14, which can be attributed to slightly awkward meme use when the same message was retained without adjustment while the meme changed. In effect, the framing in most cases seems less effective than the content.

Similarly, this is also observed in instances where the meme is reused with a different message (such as Q3, Q13, and Q17) where it can be significantly more or less offensive depending on the content. Interestingly, though, 86% of those interviewed after the survey provided input that agreed with the statement "the format of a meme can change its meaning", which is contrary to what the data shows us.

4.4   Theme matters

The theme of tragic events, especially large-scale ones (*e.g.,* the atomic bombings of Hiroshima and Nagasaki) are consistently received as being more offensive. This is evident from the significant changes in Q2, Q6, Q12, and from the lack thereof in Q15 where both themes were large-scale tragic events (The Holocaust and 9/11, respectively). Moreover, combining more than one offensive theme tends to garner higher scores signaling higher level of perceived offense. Examples include Q3 and Q12. An exception is Q17 where the theme of slavery was more offensive than murder and cancer combined.

Some themes were deemed to be less offensive than expected, such as religion. However, we avoid to draw any conclusion on this as we did not collect nor control for participants' religious and spiritual beliefs.

4.5   The demographic effect

We observed some associations between participant demographics and response to memes. Although no extremely strong correlations were detected, a few deserve remark (Figure 8). Most obvious, perhaps, is that age is a fairly good indicator of the frequency of meme use. A stronger correlation, however, exists
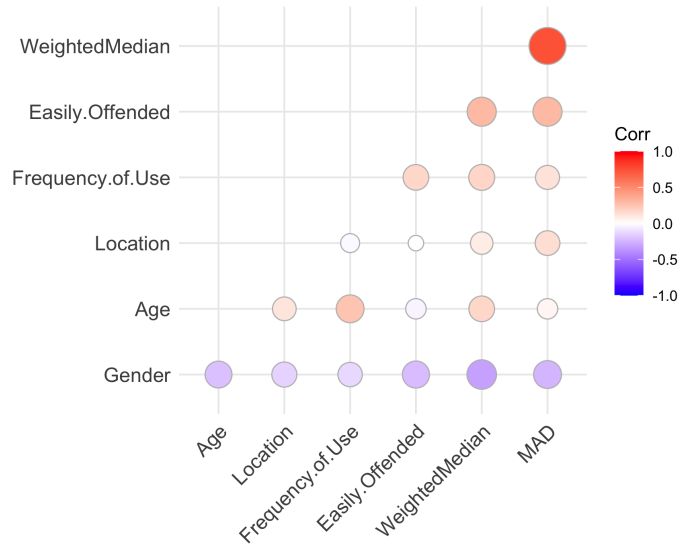
Fig. 8. Pearson correlation coefficient (r) between participant demographics and the magnitude and variability of their responses.

between a participant's gender[3] and how easily offended they are by the memes. Moreover, gender is also moderately correlated to varying Likert scores using *median absolute deviation*, a robust measure of variability. Essentially, female participants were more likely to use the full range of the Likert scale over the various survey questions, while many male participants tended to score using only part of the scale, *e.g.,* between 1 and 3. Non-binary gender participants showed no distinguishable tendency. Elsewhere, participants who signified that they are easily offended did indeed score higher and with more variability. Finally, participant location showed no significant correlation with scores.

5 Discussion

We now discuss the implications and limitations of the above results, and comment on threats to validity.

5.1 Automated analysis of meme effect

Based on our findings in Section 4.3, the automated detection of potential offense and, similarly, misinformation in memes might be easier than anticipated as it is more dependent on the textual rather than visual elements of a meme. Fortunately, textual analysis methods are more advanced and readily available

---

[3] After removing those who withheld their gender (7), but keeping non-binary answers.

than visual ones. Thus, we propose the development of systems to predict the effect of a meme in terms of its offense as an important future direction. This would be useful for detecting misinformation as well as commercial uses such as advertising campaigns. Most efforts to date on automated analysis of memes have focused on distribution, identification and clustering [8,27,7].

5.2   A meme as a humorous Trojan Horse

Some attribute the success of memes to its reliance on humor as a tool, which allows its creator to 'soften the blow' of the contained message [17]. This is certainly true for political use; Miltner comments that "the humorous nature of memes [. . .] makes them an ideal venue for political critique and commentary" [17]. This is one reason why memes tend to be quite common in online political culture [16,20,5,18,6].

   This inherent property has the potential to be exploited to mask the offensive disposition of the message, or equally political critique or misrepresented information within. In other words, using memes makes it slightly more easier to 'swallow the pill' of a message that is otherwise less palatable due to its offensive or misleading character.

   However, there seems to be a limit to this capacity to subvert. Our results are encouraging as they indicate that it is more difficult to dress a overtly offensive message with the humorous visual elements of a meme. More work is needed to investigate if this also applies to other emotions that a meme can trigger, and the potential to carry misinformation.

5.3   Threats to validity

In our study, we do not aim to establish any causal relationships hence threats to internal validity are minimal. Survey results are susceptible to be swayed by any biases that the participants already have. This aspect is difficult to control in an anonymous online survey without an additional step to calibrate perspectives, and by significantly increasing sample size. Although we comment on the correlations we observe, we are cautious not to generalize in order to mitigate threats to external validity.

   There are some threats to construct validity due to the nature of how memes are created, and we did notice this effect with a few memes (see Section 4). Furthermore, memes seldom live in isolation and are typically part of a wider discussion that is context- and community-specific. As such, there will always be a degree of threat to construct validity with synthesizing memes purely for research purposes.

6   Conclusion and Future Work

Through a randomized controlled trial, we studied how different memes are received as being offensive or otherwise. We established an understanding of the

importance of a meme's content as opposed to just its visual and cultural aspects. We also uncovered some relationships between user demographics and how likely they are to being offended by a meme. As such, our work helps address a gap in terms of how sensitive or polarizing messages could be interpreted by different audiences. This knowledge is important not just in understanding how the potential of propagating certain messages through the medium of memes, but also how controversial opinions (due to political, racial, religious, etc. perspectives) can have varied effects. We also discussed the degree to which a meme could be used as a delivery vehicle to normalize what is arguably controversial, extreme, or misleading. However, this area requires more research which is a future target of ours.

There are numerous other directions for future work. Further investigation could be made into the relationship between the timeliness of a meme and its offense effect. How this may vary across topics, online communities, and social platforms are all interesting question. Another interesting direction is to investigate the possible dependency of the 'memorability' of a meme and its inherent properties.

References

1. Agarwal, P., Garimella, K., Joglekar, S., Sastry, N., Tyson, G.: Characterising user content on a multi-lingual social network (2020)
2. Beskow, D.M., Kumar, S., Carley, K.M.: The evolution of political memes: Detecting and characterizing internet memes with multi-modal deep learning. Information Processing & Management **57**(2), 102170 (2020). https://doi.org/10.1016/j.ipm.2019.102170
3. Börzsei, L.K.: Makes a meme instead: A concise history of internet memes. New Media Studies Magazine **7** (Mar 2013)
4. Breheny, C.: "By any memes necessary": Exploring the intersectional politics of feminist memes on Instagram. Master's thesis, Uppsala University, Media and Communication Studies (Jun 2017)
5. Chagas, V., Freire, F., Rios, D., Magalhães, D.: Political memes and the politics of memes: A methodological proposal for content analysis of online political memes. First Monday **24**(2) (Feb 2019). https://doi.org/10.5210/fm.v24i2.7264
6. Dean, J.: Sorted for memes and gifs: Visual media and everyday digital politics. Political Studies Review **17**(3), 255–266 (2019). https://doi.org/10.1177/1478929918807483
7. Du, Y., Masood, M.A., Joseph, K.: Understanding visual memes: An empirical analysis of text superimposed on memes shared on twitter. International AAAI Conference on Web and Social Media **14**(1), 153–164 (May 2020)
8. Ferrara, E., JafariAsbagh, M., Varol, O., Qazvinian, V., Menczer, F., Flammini, A.: Clustering memes in social media. In: International Conference on Advances in Social Networks Analysis and Mining (ASONAM). pp. 548–555 (2013). https://doi.org/10.1145/2492517.2492530
9. Guadagno, R.E., Rempala, D.M., Murphy, S., Okdie, B.M.: What makes a video go viral? an analysis of emotional contagion and internet memes. Computers in Human Behavior **29**(6), 2312–2319 (2013). https://doi.org/10.1016/j.chb.2013.04.016

10. Heikkilä, N.: Online antagonism of the alt-right in the 2016 election. European Journal of American Studies **12**(12-2) (2017)
11. Heiskanen, B.: Meme-ing electoral participation. European Journal of American Studies **12**(12-2) (2017)
12. Huntington, H.E.: The Affect and Effect of Internet Memes: Assessing perceptions and influence of online user-generated political discourse as media. Ph.D. thesis, Colorado State University (Sep 2017), https://hdl.handle.net/10217/183936
13. Johann, M., Bülow, L.: One does not simply create a meme: conditions for the diffusion of internet memes. International Journal of Communication **13**, 23 (2019)
14. Lincoln, T.C.: A(meme)rican Politics: Gender Representation in Political Memes of the 2016 Election. Master's thesis, Rochester Institute of Technology (Aug 2019), https://scholarworks.rit.edu/theses/10116/
15. Lonnberg, A., Xiao, P., Wolfinger, K.: The growth, spread, and mutation of internet phenomena: A study of memes. Results in Applied Mathematics **6**, 100092 (2020). https://doi.org/10.1016/j.rinam.2020.100092
16. Martínez-Rolán, X., Piñeiro-Otero, T., et al.: The use of memes in the discourse of political parties on twitter: analysing the 2015 state of the nation debate. Communication & Society **29**(1), 145–160 (2016)
17. Miltner, K.M.: Internet memes. The SAGE handbook of social media pp. 412–428 (2018)
18. Moody-Ramirez, M., Church, A.B.: Analysis of facebook meme groups used during the 2016 US presidential election. Social Media + Society **5**(1) (2019). https://doi.org/10.1177/2056305118808799
19. Perez-Martin, J., Bustos, B., Saldana, M.: Semantic search of memes on twitter (2020)
20. Ross, A.S., Rivers, D.J.: Digital cultures of political participation: Internet memes and the discursive delegitimization of the 2016 u.s presidential candidates. Discourse, Context & Media **16** (2017). https://doi.org/10.1016/j.dcm.2017.01.001
21. Rowett, G.: The strategic need to understand online memes and modern information warfare theory. In: International Conference on Big Data. pp. 4437–4442. IEEE (2018). https://doi.org/10.1109/BigData.2018.8622051
22. Shifman, L.: An anatomy of a youtube meme. New Media & Society **14**(2), 187–203 (2012). https://doi.org/10.1177/1461444811412160
23. Theisen, W., Brogan, J., Thomas, P.B., Moreira, D., Phoa, P., Weninger, T., Scheirer, W.: Automatic discovery of political meme genres with diverse appearances (2020)
24. Williams, A., Dupuis, M.: I don't always spread disinformation on the web, but when i do i like to use memes: An examination of memes in the spread of disinformation. In: International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC) (Mar 2020)
25. de Winter, J.F., Dodou, D.: Five-point likert items: t test versus mann-whitney-wilcoxon. Practical Assessment, Research, and Evaluation **15**, 11 (2010). https://doi.org/10.7275/bj1p-ts64
26. Yoon, I.: Why is it not just a joke? analysis of internet memes associated with racism and hidden ideology of colorblindness. Journal of Cultural Research in Art Education **33** (2016)
27. Zannettou, S., Caulfield, T., Blackburn, J., De Cristofaro, E., Sirivianos, M., Stringhini, G., Suarez-Tangil, G.: On the origins of memes by means of fringe web communities. In: Internet Measurement Conference. p. 188–202. IMC '18, ACM (2018). https://doi.org/10.1145/3278532.3278550