The Influence of Moral Grandstanding on Perceived Consensus

Antonia H. Jager

S5241804

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Group number: 27

Supervisor: dr. Wim Meerholz

Second evaluator: dr. Toon Kuppens

In collaboration with: Berk Muratoglu, Daan van der Veen, Eline Jansen, Lilly Schönleiter, Linde van Loon

June 26, 2025

A thesis is an aptitude test for students. The approval of the thesis is proof that the student has sufficient research and reporting skills to graduate, but does not guarantee the quality of the research and the results of the research as such, and the thesis is therefore not necessarily suitable to be used as an academic source to refer to. If you would like to know more about the research discussed in this thesis and any publications based on it, to which you could refer, please contact the supervisor mentioned.

Abstract

Previous literature has suggested multiple negative effects of moral grandstanding, the use of moral talk to increase one's moral status, on public moral discourse as well as interpersonal relationships. However, its impact on the perception of consensus has not yet been subject to research. This study explored the effects of moral grandstanding on perceived consensus regarding the subject discussed by the grandstander. Additionally, previous agreement with the topic was included as a moderator variable. It was hypothesized that moral grandstanding would increase the perception of consensus, and that this effect would be stronger for those who agreed with the grandstander's views than for those who did not. To research these hypotheses, an experiment was conducted using a sample of 146 participants. Participants were exposed to a fictional online discussion and placed either in the experimental condition, which included a grandstanding comment, or in the control condition, which included a non-grandstanding comment. Contrary to the hypotheses, results showed a negative effect of moral grandstanding on perceived consensus, and no moderating effect of agreement with the topic. The findings presented in this paper support the notion of potentially negative effects of moral grandstanding on public moral discussions, as it might skew perceptions of consensus among the general population. This has implications both for individual communication as well as larger contexts, such as public discourse in real life and on social media.

Keywords: moral grandstanding, perceived consensus, agreement, moral discourse

The Influence of Moral Grandstanding on Perceived Consensus

When engaging in moral discussion, we might hope that our counterparts are fair and sincere in their arguments, so that a productive discourse can take place. However, Tosi and Warmke (2016) suggest that moral talk is often not what we would like it to be. Not all moral talk is the result of genuine interest in the topic, and, instead, some people engaging in moral talk are driven by ulterior motives for status and recognition, a behavior Tosi and Warmke (2016) termed moral grandstanding. In trying to gain moral standing, grandstanders often make use of especially dramatic and emotional rhetoric aimed at convincing their audience of their moral superiority, which might cause several problems, such as a sense of distrust in moral talk (Tosi & Warmke, 2016). A less obvious consequence might be a skewed image of moral opinions or, in other words, a distorted perception of consensus regarding the subject discussed by the grandstander. People draw information about norms from the behavior of those around them to inform their own attitudes and behaviors (e.g., Ross et al., 1977; Sherif, 1936), a process that might be disrupted through the exaggerated presentation of moral views typical of moral grandstanding. Such presentations might give undue weight to these views in people's minds, leading them to overestimate the consensus on the grandstander's opinion. Additionally, people's perceptions are influenced by beliefs they already hold (Nickerson, 1998). Because of this, the way in which people estimate consensus based on grandstanding cues might differ between people who already agree with the grandstander's views and those who do not. In this study, we explored the effects of moral grandstanding on the perceived consensus about a given topic and how this relationship might be influenced by previous agreement with this topic.

Perceived Consensus

Consensus can be seen as the degree to which people agree on a topic or believe the same thing (Akiyama et al., 2016). In this study, perceived consensus will be defined as the

estimate of this agreement. People often estimate consensus on a topic based on what behavior they observe in others (Krueger, 1998). It lies in the nature of estimation that it is not always possible to perceive consensus accurately based on the often-limited information we receive (e.g., Ross et al., 1977), which is why perceived consensus is not necessarily a reflection of the actual consensus (Zou et al., 2009). Perceived consensus has been shown to influence how people see the world and how they behave in it (e.g., Prentice & Miller, 1993; Ross et al., 1977; Zou et al., 2009), making it relevant in multiple ways, three of which will be briefly discussed. Firstly, the perception of consensus for one's own opinion might lead to differences in the strategy of resolving conflicts, which might cause problems beyond the initial issue, as groups may not be able to find agreement on conflict resolution (Whitney & Miller, 2002). Secondly, people might not act in their best interest if their perception of consensus is distorted, since being in the majority versus minority calls for different strategies of action, which cannot be chosen in one's best interest if one's position is falsely estimated (Whitney & Miller, 2002). Lastly, perception of consensus about one's own stance was also shown to lead people to engage in selective reinforcement, where they praised those who shared their views while not engaging in praise with people holding different views (Wallace et al., 2001). It seems likely that in the context of morality, this could lead people who perceive high consensus for their own moral stances to reinforce those who also share those views, while not rewarding those who disagree, leading to one-sided discourse. Because of the notable impact perceived consensus might have, it is important to consider which factors might influence it.

Past research has often explored factors leading to biased consensus perception (e.g., Prentice & Miller, 1993; Ross et al., 1977). Different factors might lead people to believe that others generally agree with their views (Ross et al, 1977). This has been summarized as false consensus "[resulting] from nonrandom sampling and retrieval of evidence and from

idiosyncratic interpretation of situational factors and forces." (Ross et al., 1977, p. 21). This indicates that people tend to draw information from those close to them, who are naturally similar, and infer the commonality of views and behaviors among the general society based on the frequency with which they observe them in people around them and themselves. Additionally, the authors suggest that such behaviors will come to mind more easily because of this, which might make them seem disproportionately common. For example, a person who often discusses social issues in their friend group, where everyone shares similar attitudes, might come to conclude that most people think the same way. This is because they use the observed frequency of opinions in themselves and among their friends as an estimate and apply it to those outside their friend group, failing to account for the fact that their social circle might be selected based on similarities, and that others might hold vastly different views. In turn, this could lead to an overestimation of consensus for one's own beliefs. Another factor that might influence the amount of perceived consensus in a similar way is moral grandstanding.

Moral Grandstanding

If a speaker makes a moral statement not as much for the sake of conveying their true opinion as for the purpose of gaining admiration from their audience for their moral views, this person might be engaging in moral grandstanding (Tosi & Warmke, 2016). The main distinction between genuine engagement in moral discussion and moral grandstanding is the attempt to be seen as moral and the disappointment that would set in if this goal were not met (Tosi & Warmke, 2016). Naturally, it is difficult to determine whether someone wants to gain recognition and make themselves out to be morally superior or if they truly care (Grubbs et al., 2019; Savejnarong et al., 2021). However, Tosi and Warmke (2016) proposed five manifestations that are typical rhetoric of moral grandstanding, which might help identify manipulative moral talk. Firstly, grandstanders might engage in piling on, meaning they

repeat speaking points to integrate themselves into the discussion and be seen as expressing the right opinion. Secondly, when ramping up, the speaker will state a more extreme moral view than the person before them to stand out as even more moral in comparison. Thirdly, another strategy used to show one's outstanding morality is to claim self-evidence, making opponents look morally inferior by making it seem as though the grandstander's point is naturally true to a moral person. A fourth manifestation is to create moral problems that do not appear to be problematic at all, described as trumping up. Lastly, the grandstander might also bring their views across in an overly emotional way, implying they care more than those who do not feel so strongly regarding the moral issue.

Moral grandstanding might cause a myriad of unfavorable outcomes (Grubbs et al., 2019; Tosi & Warmke, 2016), four of which are mentioned here. For one, Tosi and Warmke (2016) suggested people might become suspicious of all speakers and moral discourse as a whole when they sense that a speaker is taking part in moral discussions mainly for their own sake. Secondly, they also brought forward the concern that moral grandstanding might make discourse less efficient, for example, through the repetition of arguments for one's own recognition. A third unfavorable outcome might be caused by ramping up, which might lead to extremism and hate toward the outgroup as ingroup members push each other toward increasingly more extreme claims (Grubbs et al., 2020; Tosi & Warmke, 2016). Lastly, findings also showed interpersonal problems, as people with the motivation to partake in moral talk for their own benefit were found to report more conflict with others (Grubbs et al., 2019). Given the impact that moral grandstanding can have on public moral discourse and beyond, it is relevant to explore additional areas and processes it might impact.

Of particular relevance for the current study, moral grandstanding might increase the level of perceived consensus. We suggest that this occurs through two complementary mechanisms. Firstly, previous research suggests that people infer consensus based on the

information they receive from those around them and that perceptions of consensus are, at least in part, shaped by observing other people's behaviors and making inferences for the broader population based on such observations (Miller, 2023; Ross et al., 1977; Sherif, 1936). Secondly, in what they called the availability heuristic, Tversky and Kahneman (1973) explained that information that is more readily accessible serves as a cue to infer the frequency of some event. While the relationship between moral grandstanding and the availability heuristic has not been directly established, if this theory is extended to the often dramatic and highly emotional rhetoric typical of moral grandstanding, it suggests that such rhetoric could become more accessible when estimating consensus. In this way, moral grandstanding might serve as a faulty cue and lead receivers to overestimate the consensus of certain views. Thus, people might, in the case of moral grandstanding, infer a view as common not by how frequently they observed it in themselves and those around them, as was previously suggested (Ross et al., 1977). Rather, the grandstander's view might become more salient due to the exaggerated grandstanding rhetoric, causing it to be more accessible when estimating the consensus on a topic, thereby gaining heuristic quality. To our knowledge, the relationship between moral grandstanding and perceived consensus has not been previously researched.

Based on these earlier findings, we hypothesized that people exposed to moral grandstanding would perceive higher consensus for the subject discussed by the grandstander. However, we believe this relationship might be moderated by a person's level of agreement with the topic.

Agreement With the Topic

Agreement with the topic is a relatively straightforward variable, as it concerns the extent to which people endorse a certain view. People's own views have been shown to impact how much they perceive others to feel or act similarly (e.g., Ross et al., 1977). In fact,

it has been shown in past research that people who agree with a subject or behavior tend to assume that other people will also agree, which is known as the false consensus effect (Mullen et al., 1985; Ross et al., 1977).

Agreement with the topic might influence how moral grandstanding affects the perception of consensus. When people already agree with a topic, they are more likely to pay attention to such information that aligns with these beliefs or rate it as more important, a phenomenon known as the confirmation bias (Nickerson, 1998). This biased attention toward information in line with one's own opinion can show adverse effects in many areas, from poor policy decisions to errors in medicine, and can impact one's ability to reason soundly (Nickerson, 1998). If people, influenced by the false consensus effect, assume that others tend to share their views, they might be inclined to pay increased attention to the grandstander if he confirms this view by grandstanding about these beliefs. This might, in turn, further increase their perception of consensus on the topic, as they pay more attention and give more weight to such congruent information. Conversely, people tend to give less weight to information or even ignore it altogether if it does not align with their existing views (Nickerson, 1998). Thus, those who disagree with the grandstander's attitudes might downplay the grandstanding or regard the grandstander's views as unimportant altogether. This might lead to a reduced perception of consensus compared to those who agree, as the disagreeing person might not incorporate the grandstanding message into their evaluation of consensus to the same extent as a person agreeing with the grandstander would.

Given what was discussed above, we predict that moral grandstanding will increase participants' perception of consensus for people with high levels of agreement on the topic. For people low in agreement, we predict that this effect still exists, but to a smaller degree.

The Present Study

The present study aimed to contribute to the topic of moral grandstanding in three central ways. Firstly, it adds to an experimental body of research about moral grandstanding, which, at this point, is still somewhat scarce due to the relative novelty of the concept of moral grandstanding. Secondly, although the consequences of moral grandstanding have started to be explored in research (e.g., Grubbs et al., 2019), additional work on them is still needed. This study aims to provide new insights into one such aspect, namely perceived consensus, that might be influenced by moral grandstanding to better understand why grandstanding might be problematic. Lastly, although research on perceived consensus and its possible underlying mechanisms has been established (e.g., Ross et al., 1977), moral grandstanding as a factor influencing this perception has not yet been explored. By examining this relationship, this study aims to fill a gap in research on a new mechanism that might influence perceived consensus.

Methods

Participants and Design

The initial dataset consisted of 254 cases. Subsequently, 106 of those cases were removed for three reasons. Firstly, 98 cases were removed from the final data set due to incompleteness. The seriousness check at the end of the survey was set as the criterion to decide whether a case was considered incomplete: If a participant did not fill out the seriousness check, they were excluded due to incompletion of the study. However, it should be mentioned that most of these removed cases were likely participants who only briefly viewed the survey and decided not to participate either immediately or after answering only a small portion of the questionnaire. Yet others might have viewed the survey and then participated via another device, which might have led the system to count them as two separate cases, only one of whom then finished the questionnaire. Only 23 of the excluded cases reached the manipulation, and 11 of these answered any questions on the dependent

variables. Thus, the majority of the cases that were removed due to incompleteness decided to discontinue the survey early on, as opposed to quitting after having completed most of it. This is important to clarify, as the latter might have suggested a methodological issue in our study. Secondly, despite having completed the seriousness check, two cases were removed after they indicated they had not participated seriously. Lastly, six participants were removed after they no longer wanted their data to be used after the debrief about the study's true purpose. Participants ranged in age from 18 to 63 (M = 27.24, SD = 10.41). Of these participants, 84 indicated they were female, 61 indicated they were male, 1 person indicated they were non-binary, and 2 people preferred not to indicate their gender at all.

Participants were recruited using convenience sampling. Initially, the researchers reached out to their personal networks, mainly via WhatsApp. The link to the survey was additionally shared on online platforms such as Instagram and Reddit to reach a broader participant pool.

The study employed a between-subjects design in which participants were randomly assigned to one of two levels of the independent moral grandstanding variable: the experimental condition (exposure to moral grandstanding) or the control condition (exposure to a non-grandstanding statement). Agreement with the topic was used as the moderator, and perceived consensus was assessed as the dependent variable. This study was part of a bigger research project, and additional variables were assessed in the survey. The findings on these models are beyond the scope of this paper, however, a table with all variables that were included in the research can be found in Appendix A.

Materials and Procedure

The ethics committee of the University of Groningen approved the study after reviewing it. Students shared a link with their social circle and online, inviting potential participants to take part in a study about how online discussions are experienced. This

description served as our cover story, as we did not want participants to be aware of the grandstanding manipulation. The survey used for the study was designed as an online questionnaire using Qualtrics. Participation in the study took about fifteen minutes and was voluntary, meaning that participants could withdraw from the study whenever they chose.

Agreement With the Topic

After filling out the informed consent form, participants were exposed to the questions regarding the moderator, agreement with the topic. Agreement with the topic was assessed using a self-developed three-item 7-point Likert scale ranging from *strongly disagree* to *strongly agree* ($\alpha = 0.85$). As an example, one item read, "I support the use of gene modification in unborn children to reduce the risk of disorders or disabilities."

Moral Grandstanding

Following this, participants were randomly assigned to either of the conditions. In both conditions, they were presented with a screenshot from a social media discussion about gene modification and asked to read it. This was not a real discussion, but rather one created by the researchers and made to look like a real conversation on X. The topic of this discussion was gene manipulation, which was chosen based on two main considerations. Firstly, we needed a subject that would not lead to overwhelmingly one-sided opinions, i.e., a subject we could measure nuanced opinions on. This was to make sure we could measure the variability in opinions necessary to detect any effects of differences in agreement in a meaningful way. Secondly, it was important that participants did not have too many or overly strong preconceived notions about societal consensus regarding the chosen topic. This was important to ensure that the estimation they would later make about the general consensus was primarily influenced by the experimental manipulation, rather than by information they had perceived prior to the study that could have served as heuristic cues. Both conditions contained a discussion consisting of four comments, the last of which differed per condition. In the

experimental condition, we had manipulated the last comment to include typical moral grandstanding expressions as defined by Tosi and Warmke (2016), while the rest of the discussion, as well as the arguments being used, remained the same as in the control condition. The last comment in the control condition contained the same arguments as the grandstanding comment, however, it did not make use of any of the grandstanding expressions. In this way, the only aspect that differed between the two conditions was the moral grandstanding variable. For example, while the commenter in the control condition stated, “However, if we have the ability to prevent genetic disorders and reduce suffering, shouldn’t we at least consider it?”, the grandstander said, “How can anyone in good conscience stand in the way of such obvious progress?” The exact comments used can be found in Appendix B. Participants were able to download the image of the discussion so that they could refer back to it at later points in the study.

Perceived Consensus

Following the exposure to either the control or the experimental condition, perceived consensus was measured using a single item, on which participants could indicate on a percentage slider ranging from 0 to 100, what proportion of people they estimated to share the grandstander’s views. This is a commonly used approach for measuring perceived consensus (Akiyama et al., 2016; Traberg et al.,2024; Wallace et al., 2001). The item read, “Think back to the last comment in the discussion (the one you were asked to focus on). Please estimate the percentage of the general population you believe to share the last commenter’s view on gene modification. There is no right or wrong answer.” This item was inspired by the consensus perception measure described by Wallace et al. (2001, p. 77).

Manipulation Check

A manipulation check was added at the end of the survey to assess whether the experimental and control conditions differed in the noticeability of moral grandstanding. For

this, participants were asked the following question on a 7-point Likert scale: “To what degree do you agree with the following statement? The last person to comment (the one you were asked to focus on) was commenting primarily out of a desire to come across as moral just, rather than out of genuine moral concern”.

Lastly, participants were asked about their demographics, including age and gender. This was followed by a short debrief about the study's true purpose, meaning they were informed that the study was about the effects of a phenomenon called “moral grandstanding” on the quality of public moral discourse. After the debrief, participants were asked whether they wished their data to be withdrawn from the study.

Results

For the analysis of our data, SPSS and the PROCESS macro by Hayes (2013) were used. The necessary assumptions were tested, and no notable violations were found. The normality of the residuals was mildly violated. However, the kurtosis and skewness values did not exceed the -2 to +2 range, and it was thus still acceptable to continue the analysis without further correction (George & Mallery, 2010).

Manipulation Check

A manipulation check was used to assess whether our manipulation had worked as intended. To do this, an analysis of variance (ANOVA) was conducted, which used the grandstanding manipulation as the independent variable and the manipulation check item as the dependent variable. In line with our expectations, participants in the grandstanding condition ($M = 4.25$, $SD = 1.32$) scored higher on the manipulation check than participants in the control condition ($M = 3.15$, $SD = 1.40$). This positive effect was highly significant at $F(1,146) = 24.12$, $p < .001$).

Hypotheses Test

To test our hypotheses that moral grandstanding increased perceived consensus, and that this effect would be larger for people who agreed with the grandstander, we ran a single analysis using the PROCESS macro (Hayes, 2013), with our manipulation (grandstanding vs. control condition) as the independent variable, perceived consensus as the dependent variable, and agreement with the topic as the moderator.

Contrary to our hypothesis, participants in the control condition (M = 48.09, SD = 16.94) scored higher on perceived consensus than those in the experimental condition (M = 40.81, SD = 18.10). This effect was significant, $t(144) = -2.45$, $p = .02$.

However, the analysis yielded no support for a moderating effect of agreement with the topic on perceived consensus, $t(144) = -1.15$, $p = .25$. Additionally, there was no significant main effect of the moderator, $t(144) = 0.50$, $p = .62$.

Discussion

We hypothesized that moral grandstanding would increase the perception of consensus. However, contrary to this hypothesis, the results showed a significant, negative effect of moral grandstanding on perceived consensus. Furthermore, we hypothesized that the relationship between moral grandstanding and perceived consensus would be moderated by a person's agreement with the topic discussed by the grandstander, with those agreeing perceiving higher consensus than those disagreeing. This hypothesis was also not supported, as we found no significant moderation effect.

While there is no direct literature available regarding the relationship between moral grandstanding and perceived consensus, these findings were still surprising given our theoretical reasoning. We theorized, based on the availability heuristic (Tversky & Kahneman, 1973), that moral grandstanding would serve as a salient cue that would come to mind as a reference when assessing consensus, since this heuristic predicts that salient cues come to mind more easily when assessing prevalence. In our case, grandstanding was the

salient cue based on which we predicted participants would overestimate the consensus of the speaker's view. Similarly, the absence of a moderation effect was unexpected, given prior research suggesting that people tend to perceive more consensus for their own views (Ross et al., 1977) and that they tend to pay selective attention to information confirming these attitudes, known as the confirmation bias (Nickerson, 1998).

Two possible explanations of the negative effects of moral grandstanding on perceived consensus will be discussed. The first explanation pertains to our methods and could shed light on why we did not find the hypothesized positive effect, although it might exist. For the measurement of the independent variable, participants were instructed to actively think about, based on the last comment, how much they thought the general population agreed with the commenter. Additionally, they were able to re-read the comment. This method was chosen to ensure participants based their judgments only on the manipulation (i.e., grandstanding vs. control comment), instead of the entire discussion. Due to these instructions, participants might have scrutinized the information more to arrive at their estimate of consensus. It is proposed that there are two ways to process information: either relatively automatically, for example, based on stereotypes or heuristics, or by actively analyzing and scrutinizing the information (Evans, 2008). There has been some debate about the exact processes, some researchers suggesting parallel (i.e., exclusive) routes, while others argue that the two processes can happen after one another (for a review, see Evans, 2008). This discussion is beyond the scope of this paper. However, irrespective of the exact underlying mechanisms, the more analytical route might have been activated in this study due to the instructions received by participants. While we theorized that the moral grandstanding expression would serve as a heuristic cue, this did not quite align with how they were instructed to estimate the consensus. By encouraging them to actively consider the last speaker and review the comment, a more deliberate analysis might have been activated. In this way, the

grandstanding comment did not achieve heuristic value, leading participants to not generalize a single opinion to a broader consensus.

A second explanation for our findings might be that moral grandstanding indeed lowers the perception of consensus. People draw information about norms from what they observe in others (Sherif, 1936). Although more speculative, the fact that participants noticed the grandstander's ulterior motives (as confirmed by the manipulation check) makes it appear plausible that they may not have viewed his claims as a valid reflection of norms. If they doubted the grandstander genuinely believed in his own message, it might not have served to inform the estimation of the general frequency of these views. It is suggested that information that stems from an unreliable person might be disregarded (Sperber et al., 2010). However, the context differed considerably from that in the current study, and, to our knowledge, no direct research is available addressing this specific concern. Because of this, the current interpretation remains speculative, but of interest for further investigation.

There are also three plausible reasons why no moderation effect was found**.** Firstly, a theoretical explanation could lie in the unique rhetoric of moral grandstanding, which might have shifted the focus away from the grandstander's actual views. If participants focused more on the phenomenon of grandstanding than on the grandstanding content, then the level of agreement with the topic would have little influence on subsequent estimates of consensus. Secondly, a methodological explanation, connected to the one discussed regarding the main hypothesis, concerns how participants were instructed. By asking them explicitly to focus on the last speaker, we might have evoked analytical processing (Evans, 2008) in which participants actively tried to follow these instructions, possibly downregulating the influence of their agreement to only use the comment as a relevant reference. Lastly, although our scale showed good face validity and internal consistency, it was not a validated measure. It is possible that our agreement measure targeted a different aspect of agreement than was later

relevant for the manipulation. That is, while we only marginally addressed ethics and focused on a more superficial agreement with gene modification, the comments had a very strong moral focus. This might have led to the agreement we originally assessed to not be significantly relevant for the perception of agreement based on the comment. While slightly more speculative, this explanation might additionally help to explain the surprising finding of the missing main effect of agreement on perceived consensus, although previously established in research (e.g., Ross et al., 1977). If people did not access the same aspect of agreement as the grandstander, there should logically not be a significant connection between their agreement and the perceived consensus of the commenter's views.

Taken together, these findings suggest that moral grandstanding reduces the perception of agreement for the grandstander's views. This has several important implications, three of which will be discussed. Firstly, witnessing grandstanding could foster feelings of polarization, even if there is no actual increase in disagreement in society. Hearing a grandstander advocate for a certain point, while perceiving decreased agreement from most others based on this, could lead to a mistaken impression of division. Furthermore, people who agree with the grandstander but, due to the grandstanding, perceive little consensus on that opinion, might become hesitant to speak out. This aligns with the concept of pluralistic ignorance, a psychological phenomenon in which people privately agree with a certain standpoint but might feel hesitant to speak up because they falsely perceive that most others do not feel this way (Prentice & Miller, 1993). Paradoxically, those who agree with the grandstander might end up discouraged from speaking out if the perceived consensus is equally lowered for those who agree and those who disagree. This could negatively impact genuine efforts of change. Lastly, as previously mentioned, the perception of consensus influences people's problem-solving strategies, so a misconception of consensus on one's standpoint might lead to maladaptive problem-solving strategies (Whitney & Miller, 2002).

For example, someone who is in the majority might mistakenly believe they are part of the minority after hearing the grandstander. This person might then agree to a compromise they did not have to take, leading to comparatively adverse outcomes during moral discussions.

Limitations

Although this study was carefully designed, it had several limitations that might affect its generalizability and interpretability, four of which will be discussed. Firstly, the sample of this study might not be representative of the general population due to self-selective sampling, as those who decided to participate might systematically differ from those who did not (Stone et al., 2023). This might introduce bias by overrepresenting people from certain educational, cultural, or social backgrounds, while underrepresenting others (Bethlehem, 2010) in ways we cannot account for. This might limit the generalizability of our findings, as our sample might have lacked representativeness (Bethlehem, 2010; Stone et al., 2023). To improve this, more diverse sampling strategies should be used if feasible. Secondly, given financial and time constraints, self-reports were the most practical and suitable choice. Overall, self-reports are often used in psychological research and have many benefits, however, they also bear the possibility of bias (Paulhus & Vazire, 2007). Despite efforts to minimize biased responding (e.g., by informing participants that there were no right or wrong answers), some degree of influence might have remained. According to our manipulation check, participants perceived the grandstander's ulterior motives. Possibly then, after being exposed to the grandstanding condition, participants might have tried to "correct" for what they may have perceived to be an attempted manipulation by adjusting their consensus ratings. If people perceive possible bias, in the present case by perception of the grandstanding manipulation, they might have been motivated to apply corrections to account for such biases (Wegener & Petty, 1997), for example, as not to appear gullible, falling for the manipulation. These biases might impact the interpretability of the findings, as we cannot be sure if people might have corrected and, if so,

if they might have overcorrected, which can occur in these situations (Wegener & Petty, 1997). To address this issue, future studies could make use of more implicit measures, over which participants have limited control, so that their perceptions of perceived bias and subsequent correction are less likely to influence the outcomes. A third limitation might be that the post was not real but created as part of the experiment. This allowed us to manipulate the conditions and to draw causal conclusions from our findings. However, experimental control often comes at the cost of ecological validity, which might reduce the generalizability of findings to real-world settings (Kihlstrom, 2021). In the future, a "reality check" could be implemented at the end of the study, where participants would be asked how real and believable they thought the post was, to alleviate some of the concern about ecological validity. A final limitation is that the study only used one topic (gene modification). This approach was chosen because the sample size was not sufficiently large to introduce multiple topics. However, because of this, we cannot be sure whether our findings generalize to discussions about other subjects or if the topic influenced the perception in some unique way. To improve this, either multiple subjects should be tested in the same study, or future studies should make use of different topics to research whether moral grandstanding effects differ per subject or if certain subjects are more susceptible to its influences.

Despite these limitations, this study was able to give further evidence that moral grandstanding can be manipulated in experimental settings, which will be helpful for future research.

Future Research and Practical Implications

For future research, it might be interesting to explore whether different grandstanding manifestations influence perceived consensus uniquely. If it were true that people underestimated consensus due to noticing the ulterior motives of the speaker, it might be insightful to investigate whether this is true for all of the different manifestations of moral

grandstanding equally or if certain manifestations show a different or no effect on the perception of consensus. Furthermore, future research could investigate whether the effects of moral grandstanding on perceived consensus remain the same when people are part of a face-to-face discussion, rather than witnessing it online. It would be interesting to see whether tone of voice, appearance, and non-verbal communication of the grandstander, not present in our online discussion, affect the effect of moral grandstanding on perceived consensus. Lastly, future studies might consider using constructed or fictional topics, i.e., scenarios specifically designed for the context of the study. By introducing participants to a novel moral issue, equal levels of knowledge could be ensured, and preconceptions about general consensus could be reduced. This would allow for further isolation of the effects of the manipulation (grandstanding vs. control) on perceived consensus.

The findings of this study not only have theoretical implications for our knowledge about the relationship between moral grandstanding and perceived consensus but also have practical consequences, three of which will be elaborated on. Firstly, speakers should be aware of their use of moral grandstanding rhetoric and its possible negative consequences. People may grandstand unintentionally if they truly care about the subject, but also want to be morally acknowledged. This may reduce their credibility and undermine their goal of advocating for the topic. Secondly, public discourse might benefit from educational initiatives to raise awareness about moral grandstanding. As moral grandstanding might have a multitude of negative effects, people should be aware of its existence and possible manifestations. However, at the same time, we should be wary about perpetuating a false idea of certain detection of grandstanding based on typical manifestations, as the motivation behind it remains the most important factor and cannot be easily determined. Lastly, it should be considered by social media platforms to avoid incentivizing grandstanding behavior. If

performative moral talk is disproportionately rewarded (e.g., through likes and engagement), this might further increase the issue of moral grandstanding and its negative effects.

Conclusion

To conclude, the present study explored the effect of moral grandstanding on perceived consensus and how this relationship might differ by preexisting agreement. The findings suggest that moral grandstanding might decrease the perception of consensus, regardless of the audience's preexisting attitudes. While future research is still needed to solidify these findings, it is important to acknowledge the possibly destructive effects of moral grandstanding on moral discourse. It might lead to important issues not being discussed by silencing those who truly care, to fulfill the self-serving motives of a few. Thus, moral grandstanding should be taken seriously as a potential harm to public moral discussions.

References

Akiyama, Y., Nolan, J., Darrah, M., Abdal Rahem, M., & Wang, L. (2016). A method for measuring consensus within groups: An index of disagreement via conditional probability. *Information Sciences*, *345*, 116-128. https://doi.org/10.1016/j.ins.2016.01.052

Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, *78*(2), 161-188. https://doi.org/10.1111/j.1751-5823.2010.00112.x

Evans, J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255-278. https://doi.org/10.1146/annurev.psych.59.103006.093629

George, D., & Mallery, P. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update* (10th ed.). Boston: Pearson.

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. https://doi.org/10.1016/j.jrp.2020.104009

Grubbs, J. B., Warmke, B., Tosi, J., James, A. S., & Campbell, W. K. (2019). Moral Grandstanding in public discourse: Status-seeking motives as a potential explanatory mechanism in predicting conflict. https://doi.org/10.31234/osf.io/gnaj5

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Kihlstrom, J. F. (2021). Ecological validity and "Ecological validity". *Perspectives on Psychological Science*, *16*(2), 466-471. https://doi.org/10.1177/1745691620966791

Krueger, J. (1998). On the perception of social consensus. *Advances in Experimental Social Psychology*, 163-240. https://doi.org/10.1016/s0065-2601(08)60384-6

Miller, D. T. (2023). A century of pluralistic Ignorance: What we have learned about its origins, forms, and consequences. *Frontiers in Social Psychology*, *1*. https://doi.org/10.3389/frsps.2023.1260896

Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E., & Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, *21*(3), 262-283. https://doi.org/10.1016/0022-1031(85)90020-4

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175-220. https://doi.org/10.1037//1089-2680.2.2.175

Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of research methods in personality psychology*, *1*(2007), 224-239.

Prentice, D. A., & Miller, D. T. (1993). Pluralistic Ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, *64*(2), 243-256. https://doi.org/10.1037//0022-3514.64.2.243

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279-301. https://doi.org/10.1016/0022-1031(77)90049-x

Savejnarong, T., Pornsukjantra, P., & Manley, H. (2021). The interpersonal consequences of moral Grandstanding. https://doi.org/10.31234/osf.io/2jdp6

Sherif, M. (1936). *The psychology of social norms*. New York, Harper.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*(4), 359-393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Stone, A. A., Schneider, S., Smyth, J. M., Junghaenel, D. U., Couper, M. P., Wen, C., Mendez, M., Velasco, S., & Goldstein, S. (2023). A population-based investigation of participation rate and self-selection bias in momentary data capture and survey studies. *Current Psychology*, *43*(3), 2074-2090. https://doi.org/10.1007/s12144-023-04426-2

Tosi, J., & Warmke, B. (2016). Moral Grandstanding. *Philosophy & Public Affairs*, *44*(3), 197-217. https://doi.org/10.1111/papa.12075

Traberg, C. S., Harjani, T., Roozenbeek, J., & Van der Linden, S. (2024). The persuasive effects of social cues and source effects on misinformation susceptibility. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-54030-y

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232. https://doi.org/10.1016/0010-0285(73)90033-9

Wallace, D. S., Desforges, D. M., Thomas, J. C., Lord, C. G., & Sia, T. L. (2001). Effects of perceived consensus on selective reinforcement of other people's stated opinions. *Basic and Applied Social Psychology*, *23*(2), 75-85. https://doi.org/10.1207/153248301300148845

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. *Advances in Experimental Social Psychology*, 141-208. https://doi.org/10.1016/s0065-2601(08)60017-9

Whitney, C., & Miller, C. T. (2002). Goal interference, perceived consensus, and preferences for conflict-resolution strategies. *Journal of Applied Social Psychology*, *32*(4), 834-850. https://doi.org/10.1111/j.1559-1816.2002.tb00244.x

Zou, X., Tam, K., Morris, M. W., Lee, S., Lau, I. Y., & Chiu, C. (2009). Culture as common sense: Perceived consensus versus personal beliefs as mechanisms of cultural

influence. *Journal of Personality and Social Psychology*, *97*(4), 579597. https://doi.org/10.1037/a0016399

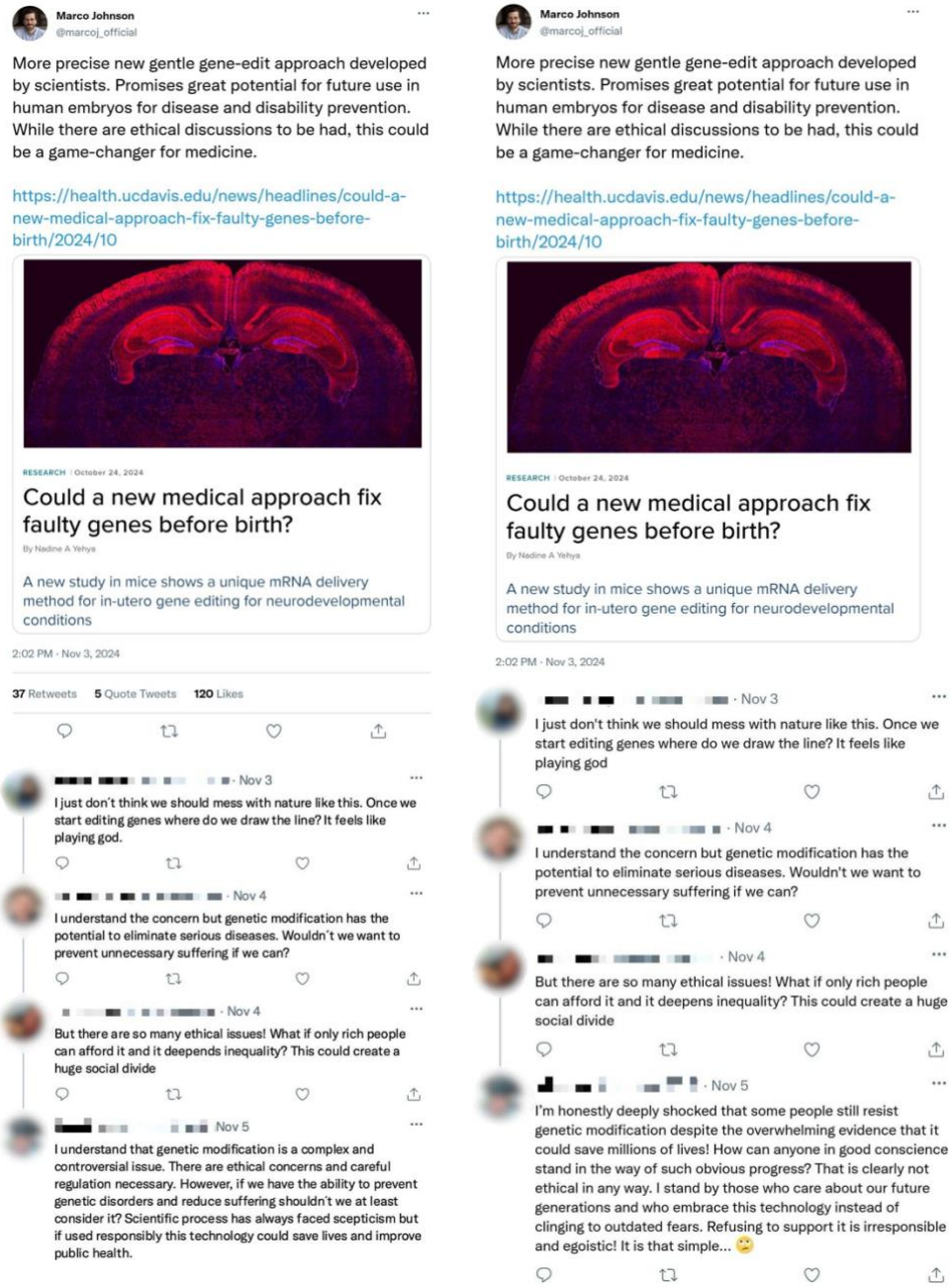
Appendix A

Overview of Study Variables

The table shows an overview of all the models that were part of this study.

Independent Variable	**Moderator Variables**	**Dependent Varibales**
Moral Grandstanding	Agreement with the topic	Perceived consensus
Moral Grandstanding	Agreeableness	Likability
Moral Grandstanding	Moral cynicism	Perception of moral grandstanding
Moral Grandstanding	Political identity strength	Ideological extremism
Moral Grandstanding	Agreement with the topic	Credibility of the speaker
Moral Grandstanding	Moral sensitivity	Engagement with topic

Appendix B

Experimental and Control Condition

The images below show the experimental (moral grandstanding) condition on the left, and the control condition on the right side.