Group Processes & Intergroup Relations

G P I R

GPIR 25th Anniversary Special Issue

Why leaders can be bad: Linking rigor with relevance using machine learning analysis to test the transgression credit theory of leadership

Group Processes & Intergroup Relations
1–20

DOI: 10.1177/13684302241242095
journals.sagepub.com/home/gpi
Sage

Ben Davies,[1] Dominic Abrams[1] and Carola Leicht[1]

Note: This article is part of a special issue entitled [Addressing Critical Social Issues From the Perspective of Group Processes and Intergroup Relations: GPIR 25th Anniversary Special Issue], guest edited by [Amber M. Gaffney, Michael A. Hogg, and Dominic Abrams].

Abstract
Transgression credit is a form of deviance credit that occurs when people are more permissive towards transgressions by in-group leaders than by in-group nonleaders and out-group members and leaders. Despite rigorous experimental and simulation evidence for transgression credit, the ability to make such group processes research relevant to organizations and wider policy requires evidence with greater ecological validity. We examine transgression credit using spontaneously arising data from Twitter (now X) to test theoretically specified reactions to instances of transgressive leadership by the UK Prime Minister Boris Johnson. Studies 1a and 1b compared Conservative and Labour Members of Parliament's (MPs') tweets in response to Boris Johnson's unlawful prorogation of Parliament (Study 1a) and his publication of an Internal Market Bill that would break international law (Study 1b) with tweets responding to a nonleader, Dominic Cummings, breaking coronavirus lockdown rules. Conservative, but not Labour, MPs were more permissive of Johnson's, but not Cummings', transgression. Study 2 examined the semantic themes occurring among supportive and unsupportive tweets posted by the UK general public in response to Boris Johnson's unlawful prorogation of Parliament. Across studies, the evidence is consistent with propositions from deviance credit and social identity theories.

Keywords
group deviance, leadership, social identity, transgression

Paper received 23 November 2023; revised version accepted 02 March 2024.

The idea that leaders have special roles and opportunities within groups is a long-accepted tenet in social and organizational psychology (e.g., Hollander, 1958). However, this is often viewed through the lens of the leader as a moral,

[1]University of Kent, UK

Corresponding author:
Dominic Abrams, School of Psychology, University of Kent, Canterbury, CT2 7NP, UK.
Email: D.Abrams@kent.ac.uk

intellectual, or respected authority who is perceived as an ideal representative of the group as a whole (Haslam et al., 2011, Haslam, Reicher et al., 2022, Haslam et al., in press; Hogg et al., 2012). In reality, leaders are (at least) as fallible as other group members, and when they are in various ways corrupt, immoral, or incompetent, group members' continued support and faith in such leaders can be perilous.

Transgressive leadership involves leader behaviors that violate established rules, norms, or laws (Abrams et al., 2013). Such leadership can be observed in a variety of social groups and organizations and has potentially significant consequences for the group. When political leaders transgress, these consequences not only impact the group but also have implications for the wider society. In recent history, UK Prime Ministers Boris Johnson and Elizabeth Truss, and U.S. President Donald Trump all represent national leaders who have transgressed by breaking rules or conventions, each case with significantly damaging consequences for the societal stability of their countries.

Problematically, transgressive leaders set a precedent for others to behave similarly, potentially spreading transgressive behavior within society. For example, hate crimes increased in the months following Donald Trump's 2016 election win (Edwards & Rushin, 2018). Clearly, transgressive leadership threatens to disrupt societal functioning. Indeed, populist leaders may even make a virtue of this, for example, the recently elected President of Argentina, Javier Milei, whose plans include smashing the country's central bank. Similarly, despite his facing innumerable lawsuits for various forms of corruption, including fraud, withholding national defense information, racketeering, and interfering with electoral results (Protess et al., 2023), it seems likely that Republicans in the US have selected Donald Trump to be their nominee for the 2024 presidential election (Green, 2023).

In the worlds of business, media, and entertainment, it is easy to find examples of CEOs and leaders who have engaged in deviant behavior, ranging from sexual harassment to fraud, embezzlement, false accounting, bullying, and deception, often over periods of many years before ultimately resigning in disgrace. Harvey Weinstein, Timothy J. Sloan (Wells Fargo), Steve Easterbrook (McDonald's), Tony Danker (Director General of the Confederation of British Industry), Mike Jeffries (CEO of Abercrombie & Fitch), Sam Bankman-Fried (FTX CEO), and Elizabeth Holmes (CEO of blood testing company Theranos) are all noteworthy examples of industry leaders exposed for their nefarious behavior. Therefore, it appears that transgressive leadership may be an enduring societal challenge. Understanding why and how such leaders might hold appeal to their group's members, and why they are often able to maintain their tenure for so long beyond the point when their transgressions are creating serious dysfunction for their organizations, is an important theoretical challenge for research on group processes. It also has clear political, social, and economic significance. Awareness of these processes is necessary if organizations and groups are to be able to effectively hold their leaders to account and prevent or respond quickly to their transgressions.

One notable pattern in recent examples of transgressive national leaders, such as Boris Johnson, Elizabeth Truss, and Donald Trump, is that other less powerful members within the group are often expelled from their positions first. For example, Dominic Cummings, Boris Johnson's senior advisor until 2020, was expelled for violating coronavirus lockdown restrictions. Kwasi Kwarteng, Chancellor of the Exchequer under Elizabeth Truss, was dismissed after his and her proposed package of tax cuts (which bypassed the usual controls such as assessment by the Office of Budget Responsibility and was preceded by the extraordinary dismissal of the Chief Secretary to the Treasury), prompting a very negative response from financial markets. In the case of Donald Trump, several of his advisors and inner circle members, such as Rudolf Giuliani, Steve Bannon, Michael Cohen, and Paul Manafort, have been arrested and charged with various crimes whilst Trump himself continues to thrive as a politician.

This pattern of medium-term survival seems indicative of transgression credit (Abrams et al., 2013): the tendency for in-group transgressive leaders to be treated more leniently than others who engage in similarly transgressive behavior. One challenge for research on transgressive leadership is that it is inherently difficult to study in realistic settings and is therefore largely limited to experimental simulations or studies involving vignettes or self-reports. Yet if social psychology is to be of use to tackle transgressive leadership, it is essential to test and demonstrate these features of transgression credit in the contexts and populations in which they arise. In the present article, we illustrate how data from social media can inform a real-world examination of transgression credit and reveal the way people justify the acceptance or rejection of transgressive leaders.

Subjective Group Dynamics and Deviant Leaders

Deviant group members generally attract harsh derogation from others within their group. People typically desire their groups to be homogeneous and cohesive (Abrams & Hogg, 1988; Hogg, 1992), which ensures their distinctiveness from opposing groups with differing values and norms (Tajfel & Turner, 1979; Turner, 1978, 1985). According to subjective group dynamics theory (Marques et al., 1998), deviant group members threaten this desired homogeneity and validity of the in-group's norms (Tajfel & Turner, 1979; Turner et al., 1987), which reduces intergroup distinctiveness (Abrams et al., 2000). Group members consequently derogate deviants as a means of symbolically marginalizing them from the rest of the group (Eidelman et al., 2006). This mitigates the deviant's potential to undermine the group norm and thereby protects or reinforces the subjective validity of the group's values (see also Anjewierden et al., 2024).

An important exception to this pattern is that in-group leaders are granted license to deviate from their group's norms more than other group members. The phenomenon of transgression credit occurs when in-group leaders are judged more positively than transgressive in-group members or transgressive out-group leaders and members (Abrams et al., 2013). In-group transgressive leaders create a psychological dilemma for followers, who must choose between upholding the normative standards of the group and continuing to perceive their leader as representative. Granting leniency in the form of transgression credit resolves this dilemma (Abrams et al., 2013). By contrast, transgressive in-group members who do not occupy leadership positions pose relatively less threat to the in-group's normative standards, and transgressive out-group leaders and members pose no threat to the in-group. Consequently, it is only in-group leaders that attract transgression credit.

The deviance credit model (Abrams et al., 2018) proposes two mechanisms that underlie this leniency for in-group leaders: their perceived prototypicality (centrality to the image of the group) and the normative conferral of a "right to lead" as part of the leadership role. When social identity is salient, judgements of leaders become increasingly dependent on how much they represent the group prototype (Barreto & Hogg, 2017; Hogg, 2001), and people are generally motivated to perceive their leader as prototypical. People may also consider that, in principle, a new leader should have the right to strike out in new directions because of their de facto status (Abrams et al., 2008). Consequently, leaders may be conferred a right to lead and act as they please.

Across a series of experiments, Abrams et al. (2018) demonstrated that perceptions of prototypicality and conferral of a right to lead both mediate the favorable perceptions of transgressive in-group leaders relative to others. Recent evidence from Syfers et al. (2022) further supports these mechanisms. In a series of studies, Syfers et al. found that deviant election candidates were viewed as more prototypical and legitimate once they had secured their new leadership position, indicating that once leaders are firmly in their role, they are viewed as more prototypical and conferred a license to deviate from the established norms of the group. On the other hand, there are clear limits to these effects. For example, when a leader transgresses in such a way that it is harmful to the group, such as by harassing a subordinate in public, people may be less willing to support their behavior. Chang (2022) found that the high prototypicality of group leaders amplified the threat that their harmful transgressions posed to the image of the group and promoted more punitive reactions. Abrams et al. (2014) showed that transgression credit was curtailed when a leader's motivation was perceived as crossing strong boundaries of morality (racism).

Research Methods for Real-World Settings

Prior research has provided valuable insight into the mechanisms behind the support and rejection of transgressive leaders, but it has important limitations. Most studies have used experimental designs with newly created groups, fictitious leaders, or transgressions. For example, Abrams et al. (2013) had university students read vignettes depicting transgressive behaviors from sports team captains, and Shapiro et al. (2011) examined self-reports from employees who were asked to recall a time when their leader transgressed. These methods may have strong internal validity but have limited ecological validity because of low experimental realism (Aronson & Carlsmith, 1969; Blascovich et al., 2002), memory effects (DePrince et al., 2004), social desirability or recall biases (van de Mortel, 2008), or potentially poor generalizability to real-world settings (Osborne-Crowley, 2020). Thus, there remains a clear need for evidence that show real-time behavioral expression of transgression credit in a real-world setting.

One way to address these issues is to examine social media data. Twitter (now rebranded as X) had over 330 million active users at the time this research was conducted (Statista, 2019), and provides ample opportunities for an analysis of real-world data in response to social events. For the present studies, we use sentiment analysis (Liu, 2012) to examine the sentiment of tweets. Lexical-based sentiment analysis (cf. Zhang et al., 2011) has previously been applied to social media data in a range of fields, such as business and politics (Ceron et al., 2014; Pang & Lee, 2008). For example, Tumasjan et al. (2010) used the "Linguistic Inquiry and Word Count" (LIWC) package (Tausczik & Pennebaker, 2010) to test whether the sentiment of tweets referencing German political figures would predict election outcomes, and Georgiadou et al. (2020) utilized the "Valence Aware Dictionary and Sentiment Reasoner" (VADER) package (Hutto & Gilbert, 2014) to analyze sentiment towards Brexit. Sentiment analysis therefore represents a suitable method for examining the opinion expressed towards transgressive group leaders in online settings.

Utilizing Twitter data relies on the spontaneous occurrence of an event relevant to the research question and hypothesis, and which subsequently garners enough attention for users to generate tweets. Although Twitter (now X) data are less freely available at the present time, we had been able to assess tweets for two transgressions committed by both a leader and a member of the same in-group at a time when it was still possible to gather all relevant tweets. In September 2019, the UK Supreme Court ruled that Prime Minister Boris Johnson had acted unlawfully in proroguing (suspending) Parliament. Boris Johnson claimed that this act was necessary to allow sufficient preparation time for the Queen's speech (setting out the legislation to be enacted in the forthcoming year), but the Court ruled that, amidst protracted Brexit negotiations, the suspension frustrated the ability of Parliament to carry out its function. A year later, Boris Johnson again transgressed by publishing his Internal Market Bill, which would violate international law. In May 2020, Dominic Cummings, a senior aide to Boris Johnson, was caught breaching coronavirus lockdown rules by travelling with his wife and son to his parents' home in Durham from his family home in London. Cummings had made the trip whilst self-isolating with symptoms of the virus, and he argued in a statement that the trip was necessary to ensure his parents could care for his son in the event that he and his wife became ill. An investigation by Durham police concluded that Cummings had likely breached lockdown rules. These transgressions, two by a leader and one by a nonleader, offered the opportunity to assess transgression credit within a real-world context.

Overview of Studies

The present studies test deviance credit theory (Abrams et al., 2018) applied to transgression credit utilizing social media data. Studies 1a and 1b examine the sentiment of tweets collected from UK Labour and Conservative Members of Parliament (MPs) in response to three different transgressive events: two by Boris Johnson (Conservative Party Leader and UK Prime Minister at the time of writing) and one by his senior aide, Dominic Cummings. We predicted that in-group members (Conservative MPs) would have posted a greater proportion of positive sentiment tweets in response to Boris Johnson's than to Dominic Cummings's transgressions. We further expect out-group members (Labour MPs) to post low proportions of positive sentiment tweets for both Boris Johnson and Dominic Cummings, and that there would be no differences in the proportions of positive tweets about Johnson and Cummings.

To explore why people express supportive or unsupportive opinions in response to transgressive leadership, Study 2 uses classification and clustering methods to examine the content of tweets from the general public in response to Boris Johnson's unlawful suspension of Parliament. Although we primarily investigated this in an exploratory manner, we did expect processes specified by leadership theories based on social identity theory (Steffens et al., 2014) and deviance credit (Abrams et al., 2018) to be evident within the data. Specifically, we expected themes concerning issues of prototypicality and conferral to be present, as well as other social identity leadership themes such as identity advancement, entrepreneurship, or impresarioship (Steffens et al., 2014). For all studies, any excluded observations and the reasons for making those exclusions are reported in the Method sections.

Study 1a

Study 1a examined transgression credit via tweets from UK Conservative and Labour MPs in the days following two transgressive events: Boris Johnson's unlawful prorogation of Parliament and Dominic Cummings's breaking of coronavirus lockdown rules. We note that Dominic Cummings is not a regular member of the Conservative Party (i.e., not an MP), but that for several years he was a key aide to Boris Johnson and was his senior advisor during the period in question. We therefore assumed that Cummings would be viewed as an in-group member. To verify this assumption, we conducted a further empirical study which is described in the supplemental material. This study confirmed that 96% of participants ($N = 56$) regarded Dominic Cummings as a member of the Conservative Party. We expected that Conservative MPs would have posted more positive sentiment tweets in response to Boris Johnson than to Dominic Cummings, and that Labour MPs would have posted a similar (low) number of positive tweets in response to both Boris Johnson and Dominic Cummings.

Method

Tweets were collected from the Twitter accounts of Conservative and Labour MPs. At the time of data collection, relevant accounts were identified from (no longer) publicly available data using the website MPs on Twitter (https://www.mpsontwitter.co.uk) and cross-checked with government MP listings to confirm the accounts were held by current MPs. Tweets were collected in the 48-hour period following two separate events: (a) A ruling by the Supreme Court that Boris Johnson had acted unlawfully in his prorogation of Parliament (occurred on September 24, 2019, with tweets collected until September 26), and (b) Dominic Cummings breaking the coronavirus lockdown rules by driving from London to Durham with his family (occurred on May 23, 2020, with tweets collected until May 25). The sample size of the collected tweets is therefore determined on the basis of convenience.

Data were collected using the Python package "Tweepy" (Roesslein, 2020), which interfaces with Twitter's application programming interface (API) to collect tweets. We first collected tweets from each individual MP's timeline dating back to the beginning of each transgressive event. To ensure that tweets were directly posted by the MP and specifically referred to the event of interest, retweets were removed and tweets were then filtered on the basis of keywords relating to each specific event. For Boris Johnson, these terms included "Boris Johnson," "PM," "Prime Minister," "prorogue," "prorogation," "court," "ruling," and "ruled." For Dominic Cummings, these terms were "Dominic Cummings," "breaking," "lockdown," "guidelines," "coronavirus," "virus," "Durham," "family," "son," and "child." The remaining tweets were then cleaned using typical natural language processing methods. Specifically, text was converted to lowercase and punctuation, special characters, and stop words were removed.

Results

Preliminary analysis. Excluding retweets, there were a total of 945 tweets posted by Conservative MPs during the Boris Johnson event and 783 during the Dominic Cummings event. There were a total of 1,973 tweets posted by Labour MPs during the Boris Johnson event and 1,532 posted during the Dominic Cummings event. After filtering tweets to include only those that were directed at Boris Johnson and Dominic Cummings, and that referenced each respective event, there were 18 tweets (2%) posted by 13 individual Conservative MPs in response to Boris Johnson and 26 (3%) posted by 20 individual Conservative MPs in response to Dominic Cummings. There were 132 tweets (7%) posted by 83 Labour MPs in response to Boris Johnson and 116 (8%) posted by 67 Labour MPs in response to Dominic Cummings.

Sentiment analysis. A sentiment analysis was conducted to assess how favorably Boris Johnson and Dominic Cummings were perceived by Conservative and Labour MPs. We used the "VADER" Python module (Hutto & Gilbert, 2014), which provides a compound polarity score ranging from -1 (negative sentiment) to $+1$ (positive sentiment). We classified tweets scoring above zero as positive, and those scoring below zero as negative. Tweets scoring zero were classified as neutral.

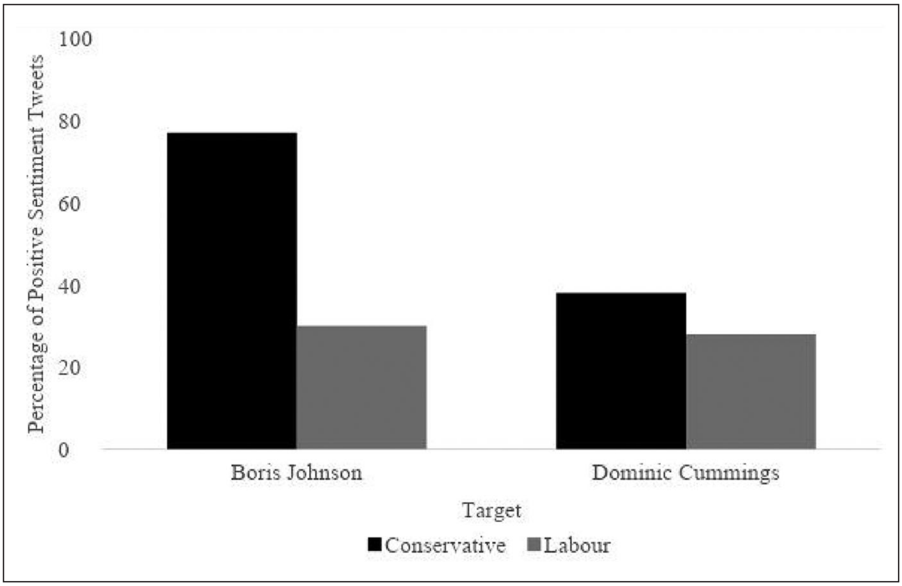
In line with the transgression credit effect, the analysis revealed that 77% of Conservative tweets discussing Boris Johnson were classified as positive, whereas only 38% discussing Dominic Cummings were classified as positive. In contrast, 30% of Labour tweets discussing Boris Johnson were classified as positive, and 28% discussing Dominic Cummings were classified as positive. Figure 1 displays the proportion of positive tweets posted by Conservative and Labour MPs in response to each event.

To further assess the robustness of this split, we conducted a three-way log-linear analysis (excluding tweets classified as neutral) to assess the association between political party (Conservative vs. Labour), target (Boris Johnson vs. Dominic Cummings), and sentiment (negative vs. positive).[1] Sensitivity power analysis indicated that the sample size ($N = 261$) was sufficient to detect effect sizes of $\phi = .17$ with 1 degree of freedom at 80% power. Backwards elimination produced a final model that retained the Political Party × Target × Sentiment association, $\chi^2(1, N = 261) = 3.93$, $p = .047$, $\phi = .12$. To break down this three-way effect, we conducted separate chi-square tests on target and sentiment within Conservative and Labour party levels. For the Conservative Party, there was a significant association between target and sentiment, $\chi^2(1, N = 40) = 4.31$, $p = .038$, $\phi = .33$, but there was no significant association between target and sentiment within the Labour Party, $\chi^2(1, N = 221) = 0.15$, $p = .903$, $\phi = .01$. Odds ratios indicated that the odds of Conservatives posting a positive sentiment tweet were 4.20 times higher when tweeted in response to Boris Johnson than to Dominic Cummings.

Study 1b

Study 1a provided initial evidence of transgression credit occurring within a real-world setting. However, a limitation of this study was the small sample size, particularly within the Conservative

Figure 1. Percentage of positive sentiment tweets by target and political party: Study 1a.



leader cell ($n = 18$). Opportunely, Boris Johnson subsequently engaged in a second widely discussed transgressive behavior. In September 2020, Boris Johnson published his UK Internal Market Bill, which set out legislation for trading between the UK's four countries. Controversially, the bill included legislation that was incompatible with the already agreed Withdrawal Agreement with the EU following Brexit. This meant that the bill would break international law. To assess the consistency of the transgression credit effect, we conducted additional analysis of this new transgression and again compared the sentiment of Conservative and Labour MPs' tweets with their responses to the Dominic Cummings event assessed in Study 1a.

Method

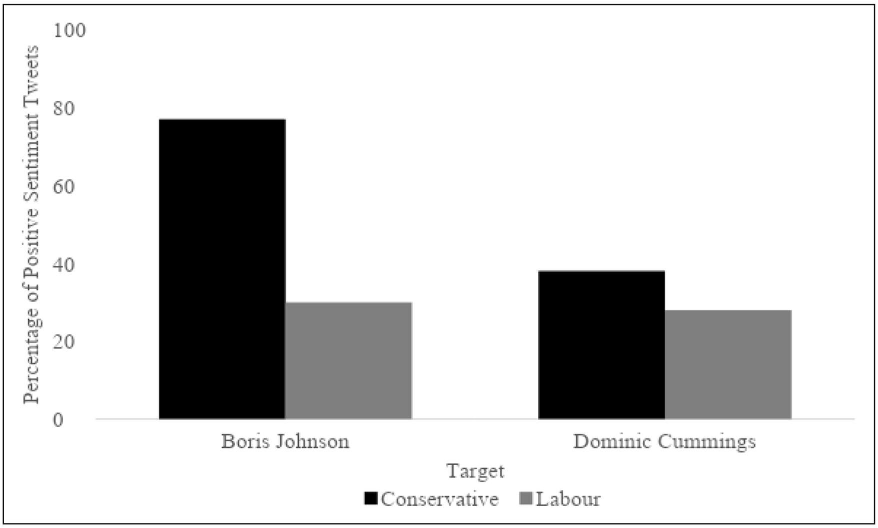
We obtained tweets using the same method as in Study 1a. Tweets were collected for a period of 1 week from September 8, 2020 (the date the Internal Market Bill was published) until September 15, 2020. To ensure relevance, we again filtered tweets using the following keywords: "Boris Johnson," "PM," "Prime Minister," "internal," "market," "withdrawal," bill," "breaking," "international," and "law."

Results

Preliminary analysis. Excluding retweets, there were a total of 2,960 tweets posted by Conservative MPs and 4,212 tweets posted by Labour MPs. After filtering tweets to include only those that were directed at Boris Johnson and that referenced the Internal Market Bill, there were 42 tweets (1%) posted by 35 individual Conservative MPs and 125 tweets (3%) posted by 58 Labour MPs.

Sentiment analysis. We again used the "VADER" Python package (Hutto & Gilbert, 2014) to conduct a sentiment analysis of MP's tweets in response to Boris Johnson's Internal Market Bill. To demonstrate the transgression credit effect, we compared these sentiment responses to those from the Dominic Cummings data from Study 1a. In line with the transgression credit effect, 93% of Conservative MPs' tweets were classified as positive in response to Boris Johnson, whereas only 38% were positive for Dominic Cummings. In contrast, only 35% of Labour MP's tweets were

Figure 2. Percentage of positive sentiment tweets by target and political party: Study 1b.



classified as positive in response to Boris Johnson, with only 28% discussing Dominic Cummings as positive. The proportion of positive tweets posted by Conservative and Labour MPs in response to each event is displayed in Figure 2.

We conducted a log-linear analysis to test the association between political party (Conservative vs. Labour), target (Boris Johnson vs. Dominic Cummings), and sentiment (negative vs. positive). Sensitivity power analysis indicated that the sample size of 275 was sufficient to detect effect sizes of $\varphi = .17$ with 1 degree of freedom at 80% power. Backwards elimination produced a final model that retained the Political Party × Target × Sentiment association, $\chi^2(1, N = 275) = 16.38$, $p < .001$, $\varphi = .24$. To break down this three-way effect, we conducted separate chi-square tests on target and sentiment within Conservative and Labour party levels. For the Conservative Party, there was a significant association between target and sentiment, $\chi^2(1, N = 62) = 23.20$, $p < .001$, $\varphi = .61$, but there was no significant association between target and sentiment within the Labour Party, $\chi^2(1, N = 213) = 1.35$, $p = .245$, $\varphi = .08$. Odds ratios indicated that the odds of Conservatives posting a positive sentiment tweet were 46.80 times higher when tweeted in response to Boris Johnson than to Dominic Cummings.

Study 2

Studies 1a and 1b revealed clear instances of transgression credit within an applied and naturally occurring context. However, given the small sample size, it was not feasible to pursue further insight into why people express different opinions in response to transgressive leadership. In Study 2, we use tweets from the general public to explore what themes occur in reactions to transgressive leadership, in an attempt to understand the reasons people offer when supporting or rejecting transgressive leaders. Study 2 investigated a selection of tweets from the general public in response to Boris Johnson's unlawful prorogation of Parliament. We used machine-learning methods to classify tweets posted in response to Boris Johnson's prorogation of Parliament as either supportive or unsupportive, and we then used clustering techniques to uncover themes and topics underlying supportive and unsupportive reactions.

Method

Data were collected from Twitter using the same method as in Studies 1a and 1b. To ensure that tweets were relevant to the context and directed towards Boris Johnson, we used convenience sampling to sample tweets that were posted as a reply to a tweet Boris Johnson had sent out himself which contained a video of his response to the Supreme Court ruling that his prorogation was illegal. A total of 4,511 replies were collected. No tweets were excluded from the analysis. As with Studies 1a and 1b, stop words and special characters were removed. In addition to the removal of stop words, we also removed several words that were frequent across all tweets ("keep," "go," and "going") and words that were specific to the context ("pm," "Boris," "prime," "minister," "ruling," "court," "prorogue," "prorogation," "ruled," "supreme," "Johnson," "judgement") to avoid overlapping clusters.

Tweet classification. For comparison, we first classified the sample of tweets into two categories: tweets expressing a supportive stance towards Boris Johnson versus tweets expressing an unsupportive stance. We classified tweets using a naïve Bayes classifier with Python's "SciKit Learn" module (Pedregosa et al., 2011). A subset of 450 tweets (approximately 10% of the full sample) were manually labelled as either expressing a supportive or unsupportive stance, vectorized into a bag-of-words model, and then used as training data for the classifier. The model achieved 85% accuracy in classification. We used this model to classify the remaining unlabeled tweets as supportive or unsupportive (for a similar method of classifying tweets, see Oscar et al., 2017). Only the tweets classified by the model were included in the clustering analysis ($N = 4,061$).

Results

Tweet clustering. To investigate the themes occurring in the supportive and unsupportive tweets, we used a KMeans clustering algorithm[2] to identify clusters of similar tweets. Each tweet was first cleaned (converted to lowercase, stop words and special characters removed) and then vectorized using a bag-of-words model. We then computed the cosine similarity between each vectorized tweet, which was used to fit the KMeans clustering model. This algorithm was run on both supportive and unsupportive tweets. To assess the most appropriate number of clusters in the data, we first iterated over several KMeans models with k ranging from 2 to 20. We used the elbow method (Thorndike, 1953) to determine the most suitable number of clusters based on the point at which inertia stabilized. This indicated that five clusters underlined the supportive tweets and seven clusters underlined the unsupportive tweets.

However, an analysis of the distribution of tweets across clusters revealed that a substantial proportion of tweets were located within one cluster for both supportive and unsupportive tweets; 57% in the supportive and 54% in the unsupportive. Having one cluster contain such a large proportion of the data is problematic as the KMeans algorithm assumes equal density within each of the clusters and, all things being equal, will attempt to split the data into roughly equal-sized clusters (Raykov et al., 2016). The large clusters in this dataset suggested that several data points from overlapping clusters had been assigned to the same cluster. Based on an additional cluster analysis and key phrase analysis of these predominant clusters (reported in the supplemental material), we determined that tweets in these clusters did not discern any clear topic, and largely overlapped with the other smaller clusters. We therefore removed them from the analysis as noise. We reran the KMeans clustering algorithm using four clusters for the supportive and six clusters for the unsupportive tweets, which showed a more acceptable distribution of tweets across clusters.

Key phrase extraction. To gain insight into what these clusters of similar tweets may represent, we conducted a key phrase analysis on each cluster within the supportive and unsupportive groups. Specifically, we examined the 10 most frequent unigrams, bigrams, and trigrams within each cluster. The full output of this analysis is reported in Table 1 (supportive tweets) and Table 2 (unsupportive tweets). From the tables, several key themes occur. Table 3 provides example tweets from each theme.

Table 1. N-gram analysis for supportive Tweet clusters.

		N-gram level		
		Unigrams	Bigrams	Trigrams
Topic 1				
	Cluster 1			
		get (200)	get us (101)	get us eu (15)
		us (173)	us eu (21)	get us get (9)
		eu (45)	get done (15)	get brexit done (8)
		brexit (44)	us get (13)	get us 31st (6)
		deal (38)	please get (9)	please get us (6)
		please (35)	get brexit (9)	get job done (5)
		done (33)	brexit done (8)	take us eu (5)
		don't (29)	lets get (7)	us get us (5)
		want (19)	us 31st (7)	us eu 31st (4)
		leave (19)	31st oct (6)	get us don't (4)
	Cluster 2			
		people (181)	people voted (20)	174 million people (12)
		voted (35)	174 million (17)	deliver people voted (5)
		don't (25)	million people (16)	people voted leave (4)
		million (24)	british people (11)	voted leave eu (2)
		leave (23)	voted leave (9)	people voted british independence (2)
		brexit (22)	people support (7)	real people country (2)
		parliament (22)	deliver people (6)	don't give people (2)
		back (18)	many people (6)	get brexit done (2)
		174 (18)	people people (6)	many people don't (2)
		vote (17)	people country (5)	million people voted (2)
Topic 2				
	Cluster 3			
		don't (237)	don't let (60)	please don't give (10)
		let (66)	don't give (51)	don't let us (9)
		give (58)	please don't (30)	don't give don't (8)
		please (47)	don't resign (23)	please don't resign (6)
		behind (34)	174 million (14)	don't let bastards (6)
		resign (34)	don't want (11)	don't dare resign (6)
		get (30)	let us (9)	don't give please (5)
		us (28)	give don't (9)	don't give 174 (5)
		need (28)	people behind (8)	give 174 million (5)
		brexit (27)	don't dare (8)	please don't let (4)
	Cluster 4			
		behind (163)	people behind (23)	people behind behind (6)
		people (39)	right behind (19)	behind 100 behind (5)
		100 (20)	behind behind (14)	people right behind (5)
		right (19)	100 behind (14)	behind take us (2)
		million (18)	174 million (14)	behind good work (2)
		country (16)	behind 100 (11)	174 million behind (2)
		174 (15)	behind way (9)	behind people behind (2)

(continued)

Table 1. (Continued)

	N-gram level		
	Unigrams	Bigrams	Trigrams
	get (14)	still behind (5)	im 100 behind (2)
	way (13)	behind peoples primeminister (5)	behind never surrender (2)
	us (11)	fully behind (5)	many us behind (2)

Note. Term frequencies are shown in parentheses.

Table 2. N-gram analysis of unsupportive Tweet clusters.

		N-gram level		
		Unigrams	Bigrams	Trigrams
Topic 1				
	Cluster 1			
		brexit (211)	nothing brexit (82)	parliament nothing brexit (19)
		nothing (92)	parliament nothing (19)	said nothing brexit (14)
		said (37)	proroguing parliament (16)	proroguing parliament nothing (12)
		thought (35)	said nothing (14)	thought nothing brexit (11)
		parliament (33)	thought nothing (13)	ha ha ha (6)
		proroguing (26)	brexit said (9)	brexit said nothing (5)
		queen (15)	thought proroguing (9)	thought proroguing parliament (5)
		deal (12)	wasn't brexit (8)	thought said nothing (4)
		youre (11)	brexit thought (8)	proroguing nothing brexit (4)
		say (11)	queens speech (7)	thought proroguing nothing (4)
	Cluster 2			
		people (254)	british people (57)	working class people (3)
		british (67)	people people (10)	lies british public (2)
		parliament (25)	people don't (7)	british people don't (2)
		brexit (24)	people want (6)	don't know british (2)
		leave (23)	leave eu (6)	know british people (2)
		eu (21)	people uk (6)	british people brexit (2)
		deal (21)	still people (5)	3 years ago (2)
		don't (20)	people changed (5)	half british people (2)
		referendum (19)	want leave (5)	british people people (2)
		one (18)	3 years (5)	one british people (2)
Topic 2				
	Cluster 3			
		law (214)	broke law (39)	fought law law (6)
		broke (39)	break law (23)	law broke law (6)
		break (27)	breaking law (23)	lied queen broke (3)
		breaking (27)	law law (11)	queen broke law (3)

(continued)

Table 2. (Continued)

		N-gram level		
		Unigrams	Bigrams	Trigrams
		disagree (23)	fought law (7)	broke law broke (3)
		don't (19)	broken law (7)	youll break law (3)
		get (18)	rule law (7)	land broke law (2)
		deal (18)	law don't (6)	need get deal (2)
		youre (17)	law broke (6)	deterred breaking law (2)
		parliament (15)	obey law (6)	way broke law (2)
	Cluster 4			
		liar (70)	liar liar (32)	liar liar liar (19)
		pants (5)	pants fire (5)	liar liar pants (4)
		fire (5)	liar pants (4)	liar pants fire (4)
		criminal (5)	proven liar (3)	liar liar ciminal (2)
		go (4)	serial liar (3)	pathological liar liar (1)
		proven (3)	go liar (3)	liar liar according (1)
		serial (3)	liar criminal (2)	liar according nothing (1)
		lock (3)	criminal liar (2)	according nothing liar (1)
		ha (3)	compulsive liar (2)	nothing liar lying (1)
		lies (2)	lock lock (2)	liar lying liar (1)
Topic 3				
	Cluster 5			
		resign (177)	resign resign (17)	decent thing resign (4)
		lied (33)	lied queen (14)	honourable thing resign (4)
		people (32)	broke law (13)	broke law resign (4)
		queen (28)	thing resign (11)	resign resign liar (3)
		country (27)	would resign (9)	lied queen resign (3)
		youre (25)	people resign (8)	integrity would resign (3)
		liar (25)	need resign (7)	decency would resign (3)
		law (24)	decent thing (6)	right thing resign (2)
		parliament (23)	resign youre (6)	lied parliament lied (2)
		thing (17)	british people (6)	lied queen lied (2)
	Cluster 6			
		resign (130)	resign resign (96)	resign resign resign (71)
		crook (3)	resign crook (3)	resign resign crook (3)
		buffoon (2)	crook resign (3)	resign crook resign (3)
		ffs (1)	resign buffoon (2)	resign resign buffoon (2)
		man (1)	buffoon resign (2)	resign buffoon resign (2)
		humility (1)	ffs resign (1)	buffoon resign resign (2)
		resign boris (1)	resign man (1)	ffs resign man (1)
		get boris out (1)	man resign (1)	resign man resign (1)
		protofascist (1)	resign humility (1)	man resign humility (1)
		drivel (1)	humility resign (1)	resign humility resign (1)

Note. Term frequencies are included in parentheses.

Table 3. Example tweets representative of each supportive and unsupportive theme.

Theme	Exemplar tweet
	Supportive tweets
Brexit and the EU	"Love and respect you Boris please take us out of the EU on the 31 October with no deal we never wanted a deal we want our own sovereignty back" (Cluster 1)
	"Thank you Boris for sticking up for the 17.4m and our democratic vote you are the peoples prime minister the pm we want and need to get brexit done" (Cluster 2)
General support	"Please don't resign press on Boris you will reap the rewards at the ballot box if that means anything anymore" (Cluster 3)
	"The majority of the people are behind you Boris keep going with your agenda we are with you" (Cluster 4)
	Unsupportive tweets
Brexit	"but hang on Boris you said that proroguing parliament was nothing to do with Brexit hmmm" (Cluster 1)
	"Stop saying the British people it was only just over a quarter of the population at best that was 3 years ago as well" (Cluster 2)
Transgression focus	"you broke the law you crook" (Cluster 3)
	"liar liar unlawful pants on fire" (Cluster 4)
Desire for resignation	"Do the honourable thing and resign you lied to the monarch and the public #johnsonmustresign" (Cluster 5)
	"resign you criminal" (Cluster 6)

Supportive group themes. Within the supportive group, each cluster represented a specific theme, but based on the key phrase analysis, we grouped these clusters together into two overarching semantic topics. Topic 1 concerns Brexit and leaving the European Union and consists of Clusters 1 and 2, which together contain 49% of supportive tweets. Within this topic, it appears that people expressing a supportive opinion discuss Brexit in different ways. For example, Cluster 1 includes the key phrases "get Brexit done" and "get us [out of the] EU,"[3] indicating a generic desire to leave the European Union by the "31st October." However, Cluster 2 includes the terms "17.4 million people," "deliver [what the] people voted [for]," and "people voted leave," framing a desire for Brexit specifically in relation to the majority vote of the 2016 referendum. This topic, given the desire for Brexit, appeared consistent with themes of identity advancement (Steffens et al., 2014).

Consistent with themes of conferral of a right to lead (Abrams et al., 2013), Topic 2 consists of Clusters 3 and 4 and concerns general statements or expressions of support for Boris Johnson. For example, Cluster 3 includes the terms "please don't give [up]" and "please don't resign," and Cluster 4 includes "people right behind [you]," "behind good work," and "behind never surrender." This topic contained 51% of supportive tweets.

These clusters indicate that, despite his unlawful behavior, supportive tweets given in response to Boris Johnson's prorogation of Parliament continued to encourage his trajectory at that time, suggesting a conferred right for Johnson to act as he pleased. Additionally, it appears that people who supported Johnson's unlawful behavior were willing to overlook it in the name of securing Brexit.

Unsupportive group themes. Within the unsupportive group, there again appeared to be a semantic consistency between several clusters, and so we grouped them into three overarching topics. Topic 1, consisting of Clusters 1 and 2

(44% of unsupportive tweets), again concerns Brexit but focuses on different aspects than the supportive group. Cluster 1 within this topic includes the terms "parliament nothing [to do with] Brexit," "proroguing nothing [to do with] Brexit," and "said nothing Brexit," reflecting previous comments and justification from Boris Johnson that the prorogation of Parliament was "nothing to do with Brexit" and was instead to offer sufficient time to prepare for the Queen's speech. Cluster 2 includes the phrases "British people don't," "British people Brexit," and "3 years ago." This cluster largely reflects ideas that "people [had] changed" and that only "half [of the] British people" had voted to leave the EU. We discern that this cluster represents the idea that the referendum vote from "3 years ago" was outdated and no longer represented the interests of the British people. This topic largely speaks to the counterfactual of accrual of prototypicality (Abrams et al., 2018; Hogg, 2001); leaders who are viewed as unrepresentative are not supported.

Topic 2 consists of Clusters 3 and 4 (together accounting for 26% of unsupportive tweets) and concerns aspects of the transgression. Specifically, Cluster 3 includes the terms "broke law" and "law broke law," and Cluster 4 includes the terms "liar liar liar," "liar pants [on] fire," and "compulsive liar," indicating that people expressing unsupportive opinions focused on the fact that Boris had lied to the Queen and that he had broken the law. Finally, Topic 3 concerns calls for Boris's resignation, with Cluster 5 including the terms "decent thing resign" and "honourable thing resign," and Cluster 6 including the terms "resign crook resign" and "resign buffoon resign." This topic accounted for 30% of unsupportive tweets.

Discussion

Study 2 examined reasons why people support or reject a transgressive leader by exploring the responses they gave to an instance of transgressive leadership. The cluster analysis indicates that people express several reasons, consistent with social identity theorizing and deviance credit, for supporting or opposing transgressive leaders. People expressing a nonsupportive stance refer to the 2016 EU referendum vote as outdated and unrepresentative. This is largely in line with the accrual (or lack) of group prototypicality, indicating that when leaders are viewed as nonprototypical or unrepresentative, they are not endorsed. Those expressing supportive stances tended to do so with general statements of approval, consistent with a conferral process in which people express unconditional support for leaders to act as they please.

Other themes of social identity leadership were also present within the data. Specifically, those expressing a supportive opinion referred to upholding the vote of the EU referendum and ensuring democracy was upheld. This is largely consistent with the dimension of identity advancement (Steffens et al., 2014) and the idea that not only must leaders be one of us but also act for our interests (Platow & van Knippenberg, 2001). Finally, those expressing an unsupportive opinion also directly focused on the transgression, confirming the illegality of the behavior and consequently that Boris Johnson was not legitimate as a leader.

General Discussion

The present article highlights the substantial disruption that groups, organizations, and society may experience as a result of leaders' transgressive behavior. It also sought to demonstrate how real-world evidence can be used to test propositions from theories founded on experimental studies of group processes. We aimed to show how social psychological theory can offer insight into why people choose to support or oppose transgressive leaders in real-world settings.

In line with the deviance credit model, we observed that Conservative MPs posted a greater proportion of positive tweets in response to Boris Johnson's transgressions than to Dominic Cummings's, whereas Labour MPs posted similarly low proportions of positive tweets for both Johnson and Cummings. This is consistent with the expected transgression credit effect. Study 2 provided further insight, suggesting that British people who rallied around Boris Johnson primarily did so over a desire for Brexit, whereas those who opposed him judged him to be unrepresentative and as having uncontestably broken the law.

Studies 1a and 1b demonstrate transgression credit in a real-world setting, increasing our confidence that transgression credit is a robust and ecologically valid phenomenon. Study 2 probed the basis for support of transgressive leadership. In line with deviance credit (Abrams et al., 2018), cluster analysis revealed themes of conferral and prototypicality in over 4,000 tweets regarding a leader's clear transgression. The deviance credit model proposes that one reason leaders receive lenient evaluations is due to their group prototypicality (Hogg, 2001; Hogg et al., 2012; Syfers et al., 2022). The present findings speak to the counterfactual; that leaders who are seen as unrepresentative of the group's position receive more negative evaluations. We also revealed that some people expressed unwavering support for Boris Johnson despite his unlawful behavior. In line with the conferral component of deviance credit, these individuals held that Boris Johnson, as the UK Prime Minister, had the right to break conventional norms (laws in this case) and act as he pleased. Revealing these theoretically specified social identity mechanisms within our Twitter data provides a novel extension of the existing research and important bolstering of theory, illustrating how social identity processes are enacted in a real-world setting.

Our exploratory analysis in Study 2 also highlights themes consistent with theorizing from social identity leadership research. Specifically, the cluster analysis suggests that a common theme among people supportive of Boris Johnson was their support for Brexit; both an inherent desire for Brexit to be delivered and to see democracy upheld by delivering the outcome of the 2016 EU referendum vote. Boris Johnson largely personified the Brexit movement, being a key figurehead of the "Vote Leave" campaign in 2016 and having his 2019 general election campaign revolve around the slogan "Get Brexit Done." Indeed, Boris Johnson's unlawful prorogation of Parliament was arguably a direct expression of his purpose of stifling opponents of his Brexit policy to facilitate its passing into law. This is in line with previous research findings that support for uncivil (i.e., transgressive) behavior from politicians may be context-dependent, and that this type of behavior is more likely to be supported when directed at out-group opponents (Walter & Kutlaca, 2024).

These themes are consistent with previous studies indicating that leaders must act for the group (Platow & van Knippenberg, 2001) or advance the group's identity (Steffens et al., 2014) to receive support. For example, B. van Knippenberg and van Knippenberg (2005) found that leaders can maintain support by engaging in self-sacrificial behaviors in the name of the group, and Giessner and van Knippenberg (2008) found that undergraduate students were willing to overlook a leader's nonprototypicality providing they brought success to the group. Specifically in the context of transgression, Abrams et al. (2013, Study 5) also established that leaders only receive transgression credit if their behavior is for the good of the group rather than for self-serving motives. The present studies bolster this theorizing. Our data suggest that not only do people who support their leader construe the leader's behavior as beneficial for the group, but they actively overlook the leader's transgressive behavior in the name of advancing in-group interests.

One unexpected finding arising from our analysis was the focus that different individuals placed on the transgression. One theme among those who adopted a nonsupportive stance was to draw attention to the fact that Boris Johnson had broken the law. In contrast, references to this transgressive behavior were scarce among those expressing a supportive opinion. Although prior experimental studies precluded this possibility by design, it remains likely that an additional driving force in reactions to transgressive leaders, currently unaccounted for by existing research, is that people differ in the extent to which they see the same behavior as transgressive, despite clear evidence or legal judgments. This is likely something that occurs as a function of group membership. For example, in-group members may downplay the threat that the deviant's misconduct represents to the group (Davies et al., 2022; Otten & Gordijn, 2014), morally disengage from the deviant's behavior (Aguiar et al., 2017), or even not view the behavior as transgressive at all.

It is unclear to what extent these differences in focus reflect cognitive distortions or are motivated by a desire to protect the in-group. In experiments on deviance credit, it appears that even when participants could accurately report the objective differences between deviant and nondeviant group members, their evaluations focused much more on differentiating individuals who deviated in ways that contrasted with their own group (antinorm deviants) from those whose deviance was manifested as an extremization of the group norm (pro-norm deviants; Abrams et al., 2008, 2013, 2018; see also Anjewierden et al., 2024).

We acknowledge several limitations with the present research. Firstly, the in-group sample size in Study 1a was small, with only 18 Conservative tweets in response to Boris Johnson's transgression. We accept that this may limit the statistical power of our analyses, and that the sensitivity power analysis may be inappropriate. Whilst this is somewhat remedied by Study 1b, the overall sample size remains relatively small. Indeed, a major limitation of social media research is that sample sizes cannot be controlled, especially in niche contexts such as that of the present study. Countering these disadvantages, the novelty and generalizability of these data largely come from their real-world origin, rather than their statistical power. The fact that our results replicate across studies and are theoretically consistent with what would be expected from deviance credit theory gives us further confidence in our data.

We are also aware that the contexts of the two transgressions in Study 1a differ. Data collection for Boris Johnson's transgression occurred amidst protracted Brexit negotiations, whereas the data collected for Dominic Cummings's occurred amidst a global pandemic. Unavoidably, there are possible confounds. Johnson's transgression may have been construed as serving the group, whereas Dominic Cummings's transgression appears much more self-serving, which may produce differences in evaluation (see Abrams et al., 2013, Study 5). Whilst a general limitation of naturally occurring data is the inability to control such confounds, there are also conceptual grounds for tolerating them. In the present case, the two events both represented highly salient breaches of rules, both protagonists were high-status members of the Conservative government, but only the leader, Johnson, benefitted from transgression credit. Indeed, Cummings seemed to be derogated as we would expect of a central in-group member, but not a leader (Pinto et al., 2010).

Twitter, X, and other social media data are inherently unstructured. This unavoidably results in several challenges in ensuring scientific rigor. However, the use of social media data and the computational methods for analysis allows the extension of more traditional research in numerous ways. For one, such data offer the unique chance to observe and examine social psychological processes in a real-world context. Not only does this establish the external validity of social psychological processes, but also allows for the advancement and development of theory. The present research reinforces the conclusion that transgression credit is a process with real-world significance and extends deviance credit theory by highlighting identity advancement as a key mechanism that operates in the support of transgressive leadership. Importantly, the potential struggle between rigor and relevance (Brewer, 1985) can be resolved not by victory for one over the other, and not within studies, but through the convergence of laboratory/experimental and real-world/observational evidence.

In revealing transgression credit and some of the group processes that bear on people's evaluations of transgressive leadership, the three studies support the robustness and ecological validity of the deviance credit model. The evidence also underlines the importance of transgression credit as a serious risk for groups and society. When political parties or other organizations that have substantial power or influence become permissive towards their leaders' transgressive, illegal, or unethical acts, the wider sustainability of widely valued standards may be imperiled. Permissiveness towards transgressive leadership may facilitate extremist, undemocratic, irrational, and potentially dangerous decisions that could ultimately harm the wider population. Indeed, events such as the U.S. Capitol riots by pro-Trump supporters point to the crucial need to understand how the continued support of transgressive leaders operates outside of the laboratory (Haslam Gaffney et al., 2022). The experimental and naturalistic study of leader and followership from the perspective of group processes and intergroup relations will continue to be a vital part of strategies to manage such risks.

Data availability

Due to confidentiality concerns and the potential identification of participants via Twitter data, the data supporting these studies are unavailable for sharing. However, the Python and R code used in our analyses are available from the corresponding author upon request.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was partially supported by funding from the Economic and Social Research Council in the form of a PhD studentship grant (ES/P00072X/1).

ORCID iD

Dominic Abrams https://orcid.org/0000-0002-2113-4572

Supplemental material

Supplemental material for this article is available online.

Notes

1. Given the nested nature of the data, additional analyses (described in the supplemental material) were conducted to assess whether a multilevel model was needed. These indicated that a multilevel model was not necessary.
2. We also explored the use of another popular clustering method known as Latent Dirichlet Allocation (LDA) modelling for clustering our data. However, LDA is conventionally used for longer textual documents, such as entire news articles or novels, and often performs poorly with shorter texts such as tweets (Yan et al., 2013). For this reason, we opted to use the KMeans algorithm instead.
3. Words in brackets represent stop words that were removed as part of the data-cleaning process.

References

Abrams, D., & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology*, *18*(4), 317–334. https://doi.org/10.1002/ejsp.2420180403

Abrams, D., Marques, J. M., Bown, N., & Henson, M. (2000). Pro-norm and anti-norm deviance within and between groups. *Journal of Personality and Social Psychology*, *78*(5), 906–912. https://doi.org/10.1037/0022-3514.78.5.906

Abrams, D., Randsley de Moura, G., Marques, J., & Hutchison, P. (2008). Innovation credit: When can leaders oppose their group's norms? *Journal of Personality and Social Psychology*, *95*(3), 662–678. https://doi.org/10.1037/0022-3514.95.3.662

Abrams, D., Randsley de Moura, G., & Travaglino, G. A. (2013). A double standard when group members behave badly: Transgression credit to ingroup leaders. *Journal of Personality and Social Psychology*, *105*(5), 799–815. https://doi.org/10.1037/a0033600

Abrams, D., Travaglino, G. A., Marques, J. M., Pinto, I., & Levine, J. M. (2018). Deviance credit: Tolerance of deviant ingroup leaders is mediated by their accrual of prototypicality and conferral of their right to be supported. *Journal of Social Issues*, *74*(1), 36–55. https://doi.org/10.1111/josi.2018

Abrams, D., Travaglino, G. A., Randsley de Moura, G., & May, P. J. (2014). A step too far? Leader racism inhibits transgression credit. *European Journal of Social Psychology*, *44*(7), 730–735. https://doi.org/10.1002/ejsp.2063

Aguiar, T., Campos, M., Pinto, I. R., & Marques, J. M. (2017). Tolerance of effective ingroup deviants as a function of moral disengagement. *Revista de*

Psicología Social, *32*(3), 659–678. https://doi.org/10.1080/02134748.2017.1352169

Anjewierden, B. J., Syfers, L., Pinto, I. R., Gaffney, A. M., & Hogg, M. A. (2024). *Group responses to deviance: Disentangling the motivational roles of collective enhancement and self-uncertainty reduction* [Conference session]. SPSP Group Processes and Intergroup Relations Preconference, February 8th, San Diego, CA.

Aronson, E., & Carlsmith, J. M. (1969). Experimentation in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology: Vol. 2. Research methods* (2nd ed., pp. 1–79). Addison-Wesley.

Barreto, N. B., & Hogg, M. A. (2017). Evaluation of and support for group prototypical leaders: A meta-analysis of twenty years of empirical research. *Social Influence*, *12*(1), 41–55. https://doi.org/10.1080/15534510.2017.1316771

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*(2), 103–124. https://doi.org/ 10.1207/S15327965PLI1302_01

Brewer, M. B. (1985). Experimental research and social policy: Must it be rigor versus relevance? *Journal of Social Issues*, *41*(4), 159–176. https://doi.org/10.1111/j.1540-4560.1985.tb01149.x

Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, *16*(2), 340–358. https://doi.org/10.1177/1461444813480466

Chang, J. W. (2022). Leader group prototypicality and reactions to leader transgressions. *Group Processes & Intergroup Relations*, *25*(7), 1880–1904. https://doi.org/10.1177/1368430221103228

Davies, B., Leicht, C., & Abrams, D. (2022). Donald Trump and the rationalization of transgressive behavior: The role of group prototypicality and identity advancement. *Journal of Applied Social Psychology*, *52*(7), 481–495. https://doi.org/10.1111/jasp.12873

DePrince, A. P., Allard, C. B., Oh, H., & Freyd, J. J. (2004). What's in a name for memory errors? Implications and ethical issues arising from the use of the term "false memory" for errors in memory for details. *Ethics & Behavior*, *14*(3), 201–233. https://doi.org/10.1207/s15327019eb1403_1

Edwards, G. S., & Rushin, S. (2018). *The effect of President Trump's election on hate crimes*. SSRN. https://doi.org/10.2139/ssrn.3102652

Eidelman, S., Silvia, P. J., & Biernat, M. (2006). Responding to deviance: Target exclusion and differential devaluation. *Personality and Social Psychology Bulletin*, *32*(9), 1153–1164. https://doi.org/10.1177/0146167206288720

Georgiadou, E., Angelopoulos, S., & Drake, H. (2020). Big data analytics and international negotiations: Sentiment analysis of Brexit negotiating outcomes. *International Journal of Information Management*, *51*, Article 102048. https://doi.org/10.1016/j.ijinfomgt.2019.102048

Giessner, S. R., & van Knippenberg, D. (2008). "License to fail": Goal definition, leader group prototypicality, and perceptions of leadership effectiveness after leader failure. *Organizational Behavior and Human Decision Processes*, *105*(1), 14–35. https://doi.org/10.1016/j.obhdp.2007.04.002

Green, L. (2023, August 28). Whether or not he is convicted, Trump will be the Republican nominee for president. *The Guardian*. https://www.theguardian.com/commentisfree/2023/aug/28/trump-republican-presidential-nominee-conviction

Haslam, S. A., Gaffney, A. M., Hogg, M. A., Rast, D. E., III, & Steffens, N. K. (2022). *Reconciling identity leadership and leader identity: A dual-identity framework*. PsyArXiv. https://doi.org/10.31234/osf.io/du24t

Haslam, S. A., Reicher, S. D., & Platow, M. J. (2011). *The new psychology of leadership: Identity, influence and power*. Psychology Press.

Haslam, S. A., Reicher, S. D., Selvanathan, H. P., Gaffney, A. M., Steffens, N. K., Packer, D., van Bavel, J. J., Ntontis, E., Neville, F., Vestergren, S., Jurstakova, K., & Platow, M. J. (2022). Examining the role of Donald Trump and his supporters in the 2021 assault on the U.S. Capitol: A dual-agency model of identity leadership and engaged followership. *The Leadership Quarterly*, *34*(2), Article 101622. https://doi.org/10.1016/j.leaqua.2022.101622

Hogg, M. A. (1992). *The social psychology of group cohesiveness: From attraction to social identity*. Harvester Wheatsheaf.

Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review*, *5*(3), 184–200. https://doi.org/10.1207/S15327957PSPR0503_1

Hogg, M. A., van Knippenberg, D., & Rast, D. I. (2012). The social identity theory of leadership: Theoretical origins, research findings, and conceptual developments. *European Review of Social*

Psychology, *23*(1), 258–304. https://doi.org/10.1080/10463283.2012.741134

Hollander, E. P. (1958). Conformity status and idiosyncrasy credit. *Psychological Review*, *65*(2), 117–127. https://doi.org/10.1037/h0042501

Hutto, C. J., & Gilbert, E. (2014, June 1–4). *VADER: A parsimonious rule-based model for sentiment analysis of social media text* [Paper presentation]. Eighth International AAAI Conference on Weblogs and Social Media, Ann Arbor, MI.

Liu, B. (2012). *Sentiment analysis and opinion mining*. Springer. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Marques, J. M., Páez, D., & Abrams, D. (1998). Social identity and intragroup differentiation as subjective social control. In S. Worchel, J. F. Morales, D. Páez & J. Deschamps (Eds.), *Social identity: International perspectives* (pp. 124–141). Sage.

Osborne-Crowley, K. (2020). Social cognition in the real world: Reconnecting the study of social cognition with social reality. *Review of General Psychology*, *24*(2), 144–158. https://doi.org/10.1177/1089268020906483

Oscar, N., Fox, P. A., Croucher, R., Wernick, R., Keune, J., & Hooker, K. (2017). Machine learning, sentiment analysis, and tweets: An examination of Alzheimer's disease stigma on Twitter. *The Journals of Gerontology Series B*, *72*(5), 742–751. https://doi.org/10.1093/geronb/gbx014

Otten, S., & Gordijn, E. H. (2014). Was it one of us? How people cope with misconduct by fellow ingroup members. *Social and Personality Psychology Compass*, *8*(4), 165–177. https://doi.org/10.1111/spc3.12098

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1–135. https://doi.org/10.1561/1500000001

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*(85), 2825–2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https:/

Pinto, I. R., Marques, J. M., Levine, J. M., & Abrams, D. (2010). Membership status and subjective group dynamics: Who triggers the black sheep effect? *Journal of Personality and Social Psychology*, *99*(1), 107–119. https://doi.org/10.1037/a0018187

Platow, M. J., & van Knippenberg, D. (2001). A social identity analysis of leadership endorsement: Effects of leader ingroup prototypicality and distributive intergroup fairness. *Personality and Social Psychology Bulletin*, *27*(11), 1508–1519. https://doi.org/10.1177/01461672012711011

Protess, B., Feuer, A., & Hakim, D. (2023, October 24). Catch up on where the Trump investigations stand. *The New York Times*. https://www.nytimes.com/article/trump-investigations-civil-criminal.html

Raykov, Y. P., Boukouvalas, A., Baig, F., & Little, M. A. (2016). What to do when K-means clustering fails: A simple yet principled alternative algorithm. *PLoS One*, *11*(9), Article e0162259. https://doi.org/10.1371/journal.pone.0162259

Roesslein, J. (2020). *Tweepy: Twitter for Python!* (Version 3.10) [Computer software]. https://github.com/Tweepy/Tweepy

Shapiro, D. L., Boss, A. D., Salas, S., Tangirala, S., & von Glinow, M. A. (2011). When are transgressing leaders punitively judged? An empirical test. *Journal of Applied Psychology*, *96*(2), 412–422. https://doi.org/10.1037/a0021442

Statista. (2019). *Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019* [Data set]. https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

Steffens, N. K., Haslam, S. A., Reicher, S. D., Platow, M. J., Fransen, K., Yang, J., Ryan, M. K., Jetten, J., Peters, K., & Boen, F. (2014). Leadership as social identity management: Introducing the Identity Leadership Inventory (ILI) to assess and validate a four-dimensional model. *The Leadership Quarterly*, *25*(5), 1001–1024. https://doi.org/10.1016/j.leaqua.2014.05.002

Syfers, L., Gaffney, A. M., III, Rast, D. E., & Estrada, D. A. (2022). Communicating group norms through election results. *British Journal of Social Psychology*, *61*(1), 300–321. https://doi.org/10.1111/bjso.12481

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In S. Worchel & W. G. Austin (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks/Cole Publishing Company.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal* *of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, *18*, 267–276. https://doi.org/10.1007/BF02289263

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010, May 23–26). *Predicting elections with Twitter: What 140 characters reveal about political sentiment* [Paper presentation]. Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC.

Turner, J. C. (1978). Social categorization and social discrimination in the minimal group paradigm. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 101–140). Academic Press.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research* (pp. 77–122). JAI Press.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.

van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *The Australian Journal of Advanced Nursing*, *25*(4), 40–48. https://www.ajan.com.au/archive/Vol25/Vol_25-4_vandeMortel.pdf

van Knippenberg, B., & van Knippenberg, D. (2005). Leader self-sacrifice and leadership effectiveness: The moderating role of leader prototypicality. *Journal of Applied Psychology*, *90*(1), 25–37. https://doi.org/10.1037/0021-9010.90.1.25

Walter, L., & Kutlaca, M. (2024). Tolerance of political intolerance: The impact of context and partisanship on public approval of politicians' uncivil behavior. *Group Processes & Intergroup Relations*, *27*(1), 158–177. https://doi.org/10.1177/13684302231156719

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013, May 13–17). *A biterm topic model for short texts* [Paper presentation]. 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil.

Zhang, L., Ghosh, R., Dekhil, M., Hsu, M., & Liu, B. (2011). *Combining lexicon-based and learning-based methods for Twitter sentiment analysis* (Technical report). HP Laboratories. https://www.researchgate.net/publication/228978262_Combining_Lexicon-based_and_Learning-based_Methods_for_Twitter_Sentiment_Analysis