

In press at Psychoneuroendocrinology

Testosterone fluctuations in response to a democratic election predict partisan attitudes toward the elected leader

Smrithi Prasad^{*ab}, Erik L. Knight^{*ac}, Amar Sarkar^{def}, Keith M. Welker^g, Bethany Lassetter^h, Pranjali H. Mehta^{ai}

^a Department of Psychology, University of Oregon, Eugene, OR 97403, USA

^b Division of Social Sciences, Yale-NUS College, Singapore, 138527, Singapore

^c Department of Psychology and Neuroscience, University of Colorado, Boulder, CO 80309, USA

^d Trinity College, Trinity Street, University of Cambridge, Cambridge, CB2 1TQ, UK

^e Leverhulme Centre for Human Evolutionary Studies, Department of Archaeology, Fitzwilliam Street, University of Cambridge, Cambridge, CB2 1QH, UK

^f Department of Human Evolutionary Biology, Harvard University, 11 Divinity Avenue, Cambridge, MA 02138, USA.

^g College of Liberal Arts, University of Massachusetts Boston, Boston, MA 02125, USA

^h Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, CA

ⁱ Department of Experimental Psychology, University College London, London, WC1H0AP, UK

*shared first authorship

Conflict of interest and financial disclosures: The authors report no financial interests or potential conflicts of interest. This research was supported by NSF Grants #1451848 and #1303743 awarded to PHM.

Correspondence to: Smrithi Prasad, Division of Social Sciences, Yale-NUS College, smrithi.prasad@yale-nus.edu.sg or Erik L. Knight, Department of Psychology and Neuroscience, University of Colorado, Boulder, Erik.Knight@colorado.edu, or Pranjali H. Mehta, Department of Experimental Psychology, University College London, pranj.mehta@ucl.ac.uk.

Abstract

Intergroup competitions such as democratic elections can intensify intergroup polarization and conflict. Partisan attitudes toward the elected leader can also shift from before to after an election, but the biology underlying these attitudinal shifts remains largely unknown. An important factor could be the hormone testosterone, which is theorized to fluctuate during competition and to influence status seeking. In a naturalistic study of 113 registered voters, we measured changes in testosterone levels and attitudes toward the winner of the 2012 US Presidential Election. We found that supporters of the losing candidate (Mitt Romney) showed acute increases in testosterone levels compared to supporters of the winner (Barack Obama) on the evening of Election Day. Supporters of the losing candidate also demonstrated flatter diurnal testosterone slopes on Election Day that persisted up to two days after the election. Furthermore, greater increases in acute testosterone levels and flatter diurnal slopes among supporters of the losing candidate were associated with less positive evaluations of the winning candidate. These testosterone-moderated attitudinal shifts observed in the days after the election showed a directionally similar pattern with a weaker effect size six months later. Finally, we confirmed that the main results were robust to alternative data analytic choices using multiverse specification curve analysis. The findings from this paper suggest that hormonal responses to large-scale intergroup competitions may shape how we perceive our elected leaders, shedding light on the biology of intergroup relations.

Keywords: intergroup competition; testosterone; social status; reverse winner-loser effect; diurnal rhythms

Testosterone fluctuations in response to a democratic election predict partisan attitudes toward the elected leader

1. Introduction

Across the animal kingdom, groups compete for control over valued resources (Packer & Pusey, 1982; Williams et al., 2002; Wilson & Wrangham, 2003), resulting in status hierarchies composed of dominant and subordinate groups. Democratic elections are uniquely human intergroup competitions in which individuals representing their political group's interests vie for leadership. Winning an election enhances the group's status and increases the leader's power to promote policies reflecting the group's values. A political leader's ability to sway the electorate's attitudes is critical to winning an election, but the importance of these attitudes extends well beyond the election. Partisan attitudes toward the elected leader not only influence the efficacy of his or her tenure via presidential approval ratings (Barrett & Eshbaugh-Soha, 2007) but also impact subsequent intergroup relations by exacerbating partisanship and deepening political and ideological divisions (Abramowitz & McCoy, 2019; Oc et al., 2018).

Despite the importance of partisan attitudes toward elected leaders and their implications for intergroup conflict, the biological factors associated with these attitudes remain largely unknown. Characterizing the biological underpinnings of attitudinal shifts toward elected leaders may provide insights into partisan tendencies that sustain or intensify intergroup conflict (Chang et al., 2016; Tajfel & Turner, 1979; Van Bavel & Pereira, 2018). In the present research, we use the 2012 United States (US) presidential election to investigate how victory or defeat in an election is linked to changes in voters' concentrations of testosterone—a steroid hormone theorized to fluctuate during competition and to influence status seeking (Carré et al., 2009; Casto & Edwards, 2016b; Mehta & Josepfs, 2006). We further examine how election outcome-related changes in testosterone are associated with shifts in attitudes toward the elected leader.

The biosocial model of status predicts that status gained from winning a competition triggers increases in testosterone concentrations, whereas the loss of status following defeat triggers decreases in testosterone concentrations (i.e., the winner-loser effect; Casto & Edwards, 2016b; Mazur & Booth,

1998). Empirical support for the winner-loser effect derives from laboratory and naturalistic competitions in competitors themselves, and the effect also occurs vicariously in competition spectators (Bernhardt et al., 1998; Stanton et al., 2009). For instance, during the 2008 US presidential election, the winning candidate's supporters experienced acute elevations in testosterone concentrations 40 minutes after the outcome was declared, relative to supporters of losing candidates (Stanton et al., 2009; see Apicella & Cesarini, 2011 for similar pattern¹). However, the winner-loser effect is heterogeneous across studies, with several studies finding effect sizes that were small or close to zero and others reporting a *reversal* of the winner-loser effect (Geniole et al., 2017; Oliveira et al., 2013; Oliveira et al., 2014; Vongas & Hajj, 2017; Wu et al., 2017; Zilioli et al., 2014). For example, in one laboratory study, losers of close competitions experienced acute *elevations* in testosterone concentrations relative to winners (Zilioli et al., 2014). Although these findings highlight substantial variability in the magnitude and direction of testosterone responses to competitive outcomes, this evidence for variability comes primarily from sports and laboratory competitions. To date, the 2008 US presidential election is the only societal intergroup competition in which the effect of the competition outcome on testosterone changes has been examined (Stanton et al., 2009). Therefore, the generalizability of these results to different societal intergroup competitions remains an open question.

The biosocial model of status also predicts that competition-related changes in testosterone should influence subsequent status-seeking behaviors. This prediction has received support in laboratory studies, particularly when examining losers' behaviors toward winners (Carré et al., 2009; Casto & Edwards, 2016b; Mehta & Josephs, 2006). Losers who experience testosterone elevations are more likely to re-challenge winners to a second competition and to behave more aggressively toward winners compared to losers who experience testosterone decreases (Carré et al., 2009; Mehta & Josephs, 2006). Although these laboratory studies suggest that competition-related changes in testosterone among losers are linked to status-seeking behavior directed toward winners, the function of testosterone changes in

¹This study was also conducted during the 2008 US presidential election and reported in a book chapter. Personal communication with the first author confirmed that there was an error in the primary figure and that the hormonal pattern found was indeed consistent with Stanton et al. (2009).

large-scale, societal competitions—for instance, among voters in democratic elections—remains unknown.

Extending these laboratory studies to a political context, we propose that testosterone fluctuations among supporters of a defeated political candidate may relate to shifts in attitudes toward the winning candidate. In particular, supporters of the defeated political candidate who experience greater *increases* in testosterone may react to the loss of status by displaying more negative attitudes toward the elected leader (Branscombe & Wann, 1994; Cikara et al., 2011; Oc et al., 2018; Riek et al., 2006). This possibility is implied by research suggesting that losing a competition may not only potentiate aggressive and dominant behaviors toward winners but also increase derogation of winners via the expression of negative attitudes (Branscombe & Wann, 1994; Riek et al., 2006). A tendency to derogate winners via the expression of negative attitudes after a competitive loss is theorized to stem from the motivation to enhance one's status after experiencing a status threat (Branscombe & Wann, 1994; Fein & Spencer, 1997), and testosterone is theorized to increase this motivation (Mehta & Josephs, 2006). On the other hand, supporters of the defeated candidate who experience greater *decreases* in testosterone may experience a reduction in negative attitudes toward the elected leader, perhaps signaling acceptance of their lower status in the new political hierarchy (Alabastro et al., 2013; Rand et al., 2009). An election provides an opportunity to extend the predictions of the biosocial model of status into the domain of attitudinal shifts as a manifestation of status-seeking motivation during a large-scale, naturalistic intergroup competition.

The present research also examined the association between testosterone responses and attitudes toward the winner among the winner's supporters. According to the biosocial model of status, testosterone responses to competition should promote dominance directed toward opponents in pursuit of social status (Carré et al., 2013; Casto et al., 2020; Oyegbile & Marler, 2005; *cf.* Apicella et al., 2014). For supporters of a winning political candidate, attitudes toward the winner are expected to reflect attitudes toward an ingroup leader, not an opponent. Thus, testosterone responses to the election outcome among supporters of the winner may be unrelated to shifts in attitudes toward the winner. However, a previous study of soccer fans found that higher baseline testosterone was related to more positive ingroup

cooperation during an intergroup competition, which suggests that testosterone may also function to promote prosocial behavior toward one's own group (Diekhof et al., 2014; Reimers & Diekhof, 2015; for further evidence related to prosocial behavior: Eisenegger et al., 2010; *cf.* Boksem et al., 2013; Dreher et al., 2016). This prior work did not examine testosterone responses to competition and was conducted outside a political context, but the findings raise the possibility that increases in testosterone among supporters of a winning politician may relate to more positive attitudes toward the winning candidate as a manifestation of ingroup positivity (Cialdini et al., 1976). Nevertheless, a prosocial function of testosterone responses to victory in a societal intergroup competition remains speculative, given little empirical research on this topic.

Beyond testing our primary research questions that address gaps in knowledge about associations between testosterone responses and attitudes toward the elected leader in a political context, the present study also addresses secondary questions about the timing of endocrine and psychological responses to competitions. Human studies have primarily examined the psychological consequences of testosterone changes just minutes after a competition has ended (Carré et al., 2009; Mehta & Josephs, 2006). However, animal studies suggest that competition-related testosterone responses produce neural changes that in turn influence behavior several days later (Cunningham et al., 2012; Fuxjager et al., 2010; Oyegbile & Marler, 2005; for initial laboratory work in humans, see also Ziloli & Watson, 2014). Further, shifts in attitudes toward elected leaders are present days after a democratic election (Alabastro et al., 2013). Informed by this preliminary evidence, we explored whether acute testosterone responses to an election's result would predict shifts in attitudes the day after the election as well as several days later. Because little work has examined longer time periods, we measured attitudes six months after the election to explore whether testosterone-related attitudinal shifts in the days after the election would weaken or persist over this longer time scale.

Finally, researchers have generally assumed that testosterone responses to competition occur minutes after a competition and are short-lived (Casto & Edwards, 2016a; Geniole et al., 2017; Gleason et al., 2009). But the societal importance of elections suggests that their outcomes may exert persistent

physiological effects hours or even days later. We therefore measured testosterone levels at multiple time points on election night, capturing testosterone changes up to several hours after results were announced, and determined whether these acute hormone responses were related to subsequent shifts in attitudes toward the elected leader. To capture hormone changes across days, we measured testosterone's diurnal rhythm over five consecutive days. Testosterone levels follow a diurnal pattern with a peak in the morning followed by decline throughout the day (Gorman & Lee, 2002). However, little work has focused on shifts in testosterone's diurnal patterns following competitions, and prior theorizing (e.g., the biosocial model of status) has not accounted for longer-term testosterone dynamics. The functional significance of testosterone's diurnal patterns is poorly understood, but changes in diurnal rhythms may help the individual adapt to and coordinate behavior in response to changing social and physical environments (Gorman & Lee, 2002; Gray et al. 2004), such as changes in social status (Mazur & Booth, 1998). Initial evidence suggests that status-relevant situations can disrupt diurnal testosterone's rhythms (Trawalter et al., 2012), and that variations in diurnal rhythms are also linked to antisocial status-relevant behaviors (Granger et al., 2003; Peckins & Susman, 2015). Building on these initial findings, we explored the extent to which the presidential election predicted subsequent shifts in testosterone's diurnal rhythms in the days following the result and the effects of these diurnal rhythm changes on shifts in attitudes toward the elected leader.

2. Method

2.1. Study design. To test our research questions, we conducted a longitudinal field study in which we measured acute testosterone responses to the election outcome, changes in testosterone's diurnal rhythms, and short- and long-term shifts in attitudes toward the elected leader in a sample of registered voters. Data were collected across a five-day naturalistic sampling period, at a pre- and post-election laboratory session, and six months after the election in a subset of participants (see Fig. 1).

To examine changes in testosterone, saliva was collected on each of the five days in the naturalistic sampling period when participants woke up (*Wake-up*), 30 minutes after they woke up (*Wake-up+30 mins*), at 3pm (Afternoon), and when they went to bed (*Bedtime*). Diurnal rhythms were indexed

by the magnitude of the slope connecting the morning peak (via the two wake-up samples), the afternoon plateau (via the 3pm sample), and the end-of-day nadir (via the *Bedtime* sample; Adam & Kumari, 2009; Granger et al., 2003; Kuzawa, et al. 2016; Trawalter et al., 2012). Additional samples were collected at 5pm, 7pm, and 9pm on Election Day (*ED*). The period from 7pm to *Bedtime* was determined as the competition-outcome window because, on average, participants reported learning about the final outcome of the election at 8:50 pm PST (networks called the election approximately at 8:12 pm PST; Ariens, 2016). Hence, these samples allowed us to capture acute testosterone reactivity across several hours, from before to after the results were announced (see Fig. S2 for affective changes that occurred during the competition-outcome window).

Attitudes toward the elected leader were measured via surveys administered in the laboratory at the *Pre-election Lab* session, the day before (*ED-1*) and day after (*ED+1*) the election, at the *Post-election Lab* session, and at the *six-month follow-up*. Because participants were scheduled to come in for the laboratory sessions based on their availability in the days before and after the election, on average, the *Pre-election Lab* session was conducted 4.37 ($SD=1.14$) days before the election and the *Post-election Lab* session 4.46 ($SD=1.33$) days after the election.

Key strengths of this study design include: (i) the ability to test associations between testosterone responses to competition and shifts in attitudes toward an elected leader; (ii) the measurement of testosterone responses over hours, diurnal hormone rhythms over days, and attitudinal changes over days and months, which permit examination of the temporal trajectory of hormonal and psychological responses to competition; and (iii) the use of a naturalistic competitive setting with significant societal importance to address our research questions (Table S6).

2.2. Participants. We recruited 113 registered voters ($M_{\text{age}} = 24.50$ years, $SD = 8.49$ years; 57.52% female; 3.67% African/African-American, 4.59% Asian/Asian-American, 76.15% European/European-American, 6.42% Hispanic/Latino, 0.92% Middle Eastern, 3.67% Native American, 1.83% Pacific Islander, 2.75% Multiracial) from Eugene and Springfield, Oregon. Voter registration was

verified against available public records. Participants provided informed consent prior to involvement in the research. The research was approved by the local research ethics committee.

Because Obama and Romney were the two main contenders in the 2012 US presidential race, we recruited and classified voters as either Obama supporters (i.e., supporters of a winning candidate; $n=62$) or Romney supporters (i.e., supporters of a losing candidate; $n=33$).² We also included a third group of undecided voters (based on their lack of preference for any candidate two weeks before the election; $n=18$) for exploratory purposes, because of their considerable power in swinging outcomes and because they serve as a less partisan reference group to facilitate the interpretation of differences between Obama and Romney supporters (see SI-Methods Sections 1 and 2 for information about classification of political groups, voting details, and confirmation of voting status). A subset of our sample ($n=66$; $M_{\text{age}}=24.48$; $SD=8.02$) also responded to a longitudinal follow-up survey and consisted of 31 Obama supporters, 21 Romney supporters, and 14 undecided voters.

Our sample has several strengths compared to prior studies examining testosterone and psychological responses to competitive outcomes: (i) our sample was larger ($n=113$) than most previous studies (e.g., $n_{\text{median}}=39$ in a recent meta-analysis, Geniole et al., 2017; $n=50$ in Mehta & Josephs, 2006); and (ii) our sample included an approximately equal representation of males and females, unlike prior studies that disproportionately focused on males (Casto & Prasad, 2017; Geniole et al., 2017).

Our final sample size was within the range of our targeted sample size of 100-120 participants indicated by power simulation analyses for the principal multilevel models (see below; see OSF for power simulation code)³.

2.3. Salivary testosterone. Saliva samples were collected via passive drool using standard procedures (Schultheiss & Stanton, 2009). At-home saliva sampling, collection, and storage adhered to standard procedures for diurnal hormone measurement (Adam & Kumari, 2009; Gildner, 2021; Gorman & Lee, 2002; Granger et al., 2003; Kudielka et al., 2012; Kuzawa, et al. 2016; Trawalter et al., 2012) and

² The skewed distribution of political groups in our sample was representative of the political ideologies of those in the Eugene/Springfield area in Oregon—a liberal college city.

³ We ran these power analyses after collecting the data, but we did so in a manner consistent with an *a priori* approach (e.g., no values were extracted from the data to inform the simulations).

were further reinforced through additional compliance measures. Some of these compliance measures included sending participants automated text messages prior to each sample's collection time as both a reminder and a means to increase likelihood of compliance, and having participants indicate the time they provided each saliva sample by using an electronic time stamper (see SI-Methods Section 3.1 and 3.2 for more information). Samples were assayed in duplicate using commercially available enzyme-linked immunoassay (EIA) kits (Salimetrics LLC; State College, PA) following standard protocols and practices (Schultheiss & Stanton, 2009; see SI-Methods Section 3.3). The average intra- and inter-assay coefficients of variation (CVs) were 6.45% and 8.07%, respectively. All raw testosterone values 3 *SDs* above the mean (within sex) at each time point were winsorized to the value at 3 *SD*.

The ease of using immunoassays, alongside their accessibility and cost-effectiveness allowed us to measure testosterone concentrations in over 2500 samples. Most studies in psychoneuroendocrinology have measured testosterone levels using immunoassays but have not directly evaluated the measurement validity of the immunoassays. We confirmed the validity of testosterone concentrations obtained from EIAs in our study by assessing them against a highly accurate reference method (liquid chromatography tandem mass spectrometry, LC-MS/MS; Prasad et al., 2019). In 96 randomly selected samples, testosterone concentrations from the EIAs used in this study significantly correlated with those from LC-MS/MS (overall sample: $r_{(94)}=.80$, 95%CI[.72, .86], $p<.001$; males: $r_{(40)}=.54$, 95%CI[.28, .72], $p<.001$; females: $r_{(52)}=.66$, 95%CI[.47, .79], $p<.001$). We also conducted Deming regression analysis, which determines how closely the relationship between two methods conforms to a line of identity that assumes equality between methods. Deming regression analysis in males and the upper 50% of the distribution showed good convergence with the line of identity (Table S5). Deming regressions in females and the lower 50% of the distribution showed some deviations from the line of identity that are expected in low-concentration samples (Herold & Fitzgerald, 2003; Schultheiss & Stanton, 2009), but these deviations were smaller, and the confidence intervals were narrower for samples from the current study compared to previous research (Welker et al., 2016; see SI-Methods Section 3.4 for more information). Fixed bias and proportional bias results are also reported in the supplemental material (Table S5). Collectively, these

analyses provide evidence of improved testosterone measurement with immunoassays in the present study compared to previous work⁴ (Welker et al., 2016). In the current study's discussion, we offer recommendations for evaluating and improving hormone measurement validity in future work.

2.4. Attitudes toward the elected leader. Attitudes toward the elected leader were assessed via scales adapted from polling surveys that resemble presidential approval ratings (Gallup, 2012a; Gallup, 2012b). Our measure of attitudes consisted of twelve items that measured judgments of Obama's leadership capabilities, and ten items that measured his ability to handle social, economic, and political issues. Responses were provided on 5-point Likert scales and were averaged across all 22 items (see SI-Methods Section 4.1. for individual items). In a subset of our sample ($n=92$), we also assessed online behaviors on social media (i.e., Facebook) that reflected attitudes toward the elected leader (see SI-Methods Section 4.2. for more information about how attitudes on social media were coded).

2.5. Statistical analysis strategy. Multilevel linear regression models (MLM) were used to conduct the main analyses because of the hierarchical structure of our data wherein observations across time were nested within participants, and participants were further nested within political groups. MLMs provide several advantages over traditional methods of analysis (e.g., ANOVAs) including: (i) the capability to model fixed and random effects to better account for individual variation around average fixed effects, (ii) no listwise deletion of data, thereby boosting statistical power, (iii) the ability to focus on pre-planned contrasts that test differences in linear and curvilinear responses across specific political groups as opposed to unfocused omnibus tests that require follow-up post hoc tests to determine where specific effects may occur, and (iv) greater flexibility in the modeling approach (e.g., not requiring assumptions of sphericity, or homogeneity of regression slopes; Raudenbush & Bryk, 2002; see SI-Methods Section 6.1 for more detail).⁵

⁴ Testosterone concentrations from EIAs reported in Welker et al. (2016) were obtained from re-assaying the random subset of samples for which we report methodological validity in the current article and in Prasad et al. (2019). The discussion in Prasad et al. (2019) highlights possible explanations (e.g., reduced matrix interference) for improved measurement of testosterone in the current study compared to Welker et al. (2016).

⁵ MLMs do not require that assumptions pertaining to the distribution of outcome variables be met, and therefore we do not transform testosterone data. This approach is consistent with published recommendations (Feng et al., 2014; Stroup et al., 2015).

In all our models, at Level 1, linear and quadratic time parameters were used to predict outcome variables. The inclusion of these time parameters allowed us to model the linear and curvilinear (quadratic) temporal trajectory of testosterone responses to the election outcome and testosterone-moderated shifts in attitudes. For example, this approach allowed us to use the linear effect to assess if testosterone levels rose across the timepoints, the quadratic effect to assess a return to baseline, or some combination of the two response patterns. The intercepts and slopes of the models at Level 1 were allowed to vary randomly, as far as possible (see SI-Methods Section 6.2 for equations). At Level 2, differences across political groups were tested via two orthogonal dummy codes. We treated supporters of the losing candidate (Romney supporters) as the “reference group” (1st contrast: Romney supporters= 0, Obama supporters= 1; 2nd contrast: Romney supporters= 0, undecided voters= 1) because we expected differences in testosterone concentrations between losers and winners, and that losers’ changes in testosterone would predict shifts in attitudes toward the elected leader (Carré et al., 2009; Geniole et al., 2017; Mehta & Josephs, 2006). The orthogonal contrasts were allowed to interact with lower-order linear and quadratic time parameters. In the main manuscript we report findings from the comparison between Romney and Obama supporters. We included undecided voters in all models (see SI-Results). Sex was entered as a covariate (Female=0 and Male=1) in all models. In models testing for sex differences and similarities, the sex variable interacted with other variables of interest. Finally, we performed simple slopes analysis to decompose and interpret the magnitude and direction of overarching interactions.⁶

In the model assessing acute changes in testosterone on the night of the election, raw testosterone concentrations during the theorized competition-outcome window (*7pm*, *9pm*, and *Bedtime*) were treated as outcome variables. In the model examining changes in testosterone diurnal slopes, empirical Bayesian estimates of participants’ linear diurnal slopes served as outcome variables (see Marceau et al., 2015 for similar methodological approach). Specifically, we performed individual MLMs for each day of the

⁶ We do not include *p* values for the simple slope estimates in our models because *p* values cannot be reliably calculated for simple slopes produced within longitudinal MLMs using the lme4 package in R. Therefore, we encourage readers to consider the size and direction of the confidence intervals instead.

naturalistic sample period using the *Wake-up*, *Wake-up+30 mins*, *Afternoon*, and *Bedtime* samples, and extracted empirical Bayesian estimates for the linear slopes for every participant.

In models examining testosterone-moderated shifts in attitudes, attitudes toward the elected leader were treated as outcome variables. Short-term shifts in attitudes were assessed prior to the announcement of the election outcome (*Pre-Election Lab and ED-1*) and then after the outcome was known (*ED+1 and Post-Election Lab*). Long-term shifts in attitudes were assessed across ratings of the elected leader from before the election (*Pre-Election Lab and ED-1*) to six months after it. Testosterone change predictors—acute testosterone changes (standardized within sex; Mehta et al., 2009) and changes in diurnal slopes from *ED-1 to ED* (standardized within sex)—were entered into their respective models and were allowed to interact with the lower-order time parameters (at Level 1) and political group contrasts (at Level 2).

All statistical modeling was conducted in R (v.4.0.3; R Core Team, 2020) using the lme4 package (version 1.1-18-1; Bates et al., 2015). See SI-Method Section 6.3 for more information on other packages used in our data analysis.

2.6. Power Simulations. Our power simulations focused on determining the power of our models to detect small, medium, or large effect sizes with varying sample sizes and varying degrees of stability of our outcome variables. Simulations of testosterone responses to the election outcome indicated that the study was approximately 80% powered to detect a moderate effect size assuming a total sample size of $n=100$ and moderately high ICC for testosterone ($r \geq .7$). At lower ICCs (e.g., due to less reliable testosterone assays), these simulations indicated that a total sample size of 120 or greater was necessary to achieve 80% power for a moderate effect size. Simulations of testosterone responses predicting attitude changes also indicated that the study was approximately powered at 80% to detect moderate effect sizes with a sample size of at least $n=100$, assuming a relatively large ICC ($ICC \geq .7$). With a total sample size of 120, moderate effect sizes were approximately 80% powered at all but the lowest ICC ($r = .3$). See SI-Methods (Section 7 and Fig. S1) for full a description and results related to these power simulations.

2.7. Data availability and pre-printing. Data and relevant code are available on the Open Science Framework (<https://osf.io/mjavw/>). This article was also available as a pre-print prior to submission (<https://psyarxiv.com/w6rz9>).

3. Results

3.1. Primary analyses

3.1.1. Acute testosterone change. First, we tested the effect of the electoral outcome on acute testosterone changes, across the theorized competition-outcome window (*7pm*, *9pm*, and *Bedtime*). The MLM analysis revealed a significant Time (Linear) \times Group (Romney versus Obama) interaction ($B=-14.60$, $95\%CI[-26.00, -3.21]$, $\beta=-.33$ $p=.013$; see Tables S9 and S10). Simple slopes analysis indicated that supporters of the winning candidate experienced a decrease in testosterone levels on election night, consistent with typical circadian decline (Time (Linear): $B= -10.68$, $95\%CI[-17.25, -4.10]$; Fig. 2- Left Panel). However, supporters of the losing candidate resisted the circadian decline (Time (Linear): $B= 3.93$, $95\%CI[-5.39, 13.24]$), instead experiencing acute increases in testosterone levels, a pattern consistent with the reverse winner-loser effect (Zilioli et al., 2014).

To confirm this interpretation, we calculated acute testosterone reactivity from before to after the election results were announced (*Bedtime* minus *7pm* testosterone levels). We then conducted a regression analysis comparing these testosterone reactivity scores between groups, controlling for sex. A significant difference again emerged for testosterone reactivity scores between supporters of the winning and losing candidate, in support of the reverse winner-loser effect ($B=-19.36$, $95\%CI[-36.07, -2.65]$, $\beta=-.51$, $p=.024$, $d=0.48$). Supporters of the losing candidate showed acute increases in testosterone levels after the outcome was declared ($M=4.27$ pg/mL; $SE=6.91$), compared to supporters of the winning candidate ($M=-15.09$ pg/mL; $SE= 4.80$; Fig. 2- Right Panel; also see Fig. S3 for pirate plot).

3.1.2. Acute testosterone change predicting short-term shifts in attitudes toward the elected leader. Next, we tested whether acute changes in testosterone (raw difference score from *7pm* to *Bedtime*, standardized within sex) on election night predicted short-term shifts in attitudes toward the elected leader across four time points: *Pre-Election Lab* session, *ED-1*, *ED+1* and *Post-election Lab* session. Our

analysis revealed a significant three-way Time (Linear) \times Acute Testosterone Reactivity \times Group interaction (Romney versus Obama: $B=0.29$, 95% CI[0.17, 0.41], $\beta=.31$, $p<.001$; Table S11). Simple slopes analyses revealed that among supporters of the losing candidate, larger increases in testosterone levels predicted less positive attitudes toward the elected leader (+1SD Acute Testosterone Reactivity, Time (Linear): $B=-0.20$, 95%CI[-0.31, -0.09]), whereas weaker testosterone reactions predicted more positive attitudes toward the elected leader from before to after the election (-1SD Acute Testosterone Reactivity, Time (Linear): $B=0.34$, 95%CI[0.19, 0.49]; see Fig. 3- Left Panel). Among supporters of the winning candidate, testosterone reactivity did not moderate shifts in attitudes. A follow-up linear regression that examined shifts in attitudes from before (average of attitudes from *Pre-Election Lab* and *ED-1*) to after the election (average of attitudes from *ED+1* and the *Post-Election Lab*) corroborated our primary results: the Acute Testosterone Reactivity \times Group (Romney versus Obama) interaction predicted attitude shifts toward the elected leader ($B=0.25$, 95%CI[0.13, 0.37], $\beta=.91$, $p<.001$, $d=0.98$; see Fig. S4 and SI-Results Section 2).

3.2. Secondary analyses

3.2.1. Changes in diurnal slopes. We next determined whether the election outcome predicted changes in the magnitude of testosterone diurnal linear slopes. Because testosterone levels are the highest in the morning and drop to their nadir at night, we expected, on average, a negative slope for each day of the sampling period (see Table S13). Larger negative slope values are indicative of steeper slopes, whereas less negative slope values are indicative of flatter slopes. The MLM revealed a significant Day (Linear) \times Group (Romney versus Obama supporters) interaction ($B=-7.30$, 95% CI[-12.17, -2.43], $\beta=-.38$, $p=.004$; Table S14). Simple slopes analysis revealed that supporters of the losing candidate exhibited flatter slopes across the sampling period (see Fig. 4; Time (Linear): $B=7.63$, 95% CI[3.80, 11.46]). These flatter diurnal slopes among supporters of the losing candidate were evident on Election Day (see SI-Results Section 5, Table S16, and Fig. S5 for analysis only on Election Day) and persisted two days after the election (see SI-Results for other follow up analyses). On the other hand, supporters of the winning candidate showed no change in their diurnal rhythms across the sampling period (Time (Linear): $B=0.33$,

95% $CI[-2.46, 3.12]$). Given that diurnal rhythms can be affected by sleep patterns, we reran the MLM analysis controlling for hours of sleep. The electoral outcome robustly predicted shifts in diurnal slopes even after controlling for hours slept (Day (Linear) \times Group (Romney versus Obama): $B=-7.25$, 95% $CI[-12.12, -2.37]$, $\beta=-.38$, $p=.004$).

3.2.2. Changes in diurnal slopes predicting shifts in attitudes toward the elected leader.

Next, we explored if competition-related changes in diurnal slopes (raw slope difference from *ED-1* to *ED*, standardized within sex) predicted short-term shifts in attitudes toward the elected leader. We found a significant Time (Linear) \times Diurnal Slope Change \times Group (Romney versus Obama) interaction ($B=0.13$, 95% $CI[0.02, 0.23]$, $\beta=.13$, $p=.019$; see Table S17 and Fig. S6). Similar to the pattern of results with acute testosterone reactivity, simple slopes analysis indicated that among supporters of the losing candidate, flatter slopes predicted less positive attitudes toward the elected leader (+1SD Diurnal Slope Change, Time (Linear): $B=-0.13$, 95% $CI[-0.24, -0.01]$) whereas steeper diurnal slopes predicted more positive attitudes toward the elected leader (-1SD Diurnal Slope Change, Time (Linear): $B=0.13$, 95% $CI[0.00, 0.25]$). A follow-up linear regression that examined shifts in attitudes from before (average of attitudes from *Pre-Election Lab* and *ED-1*) to after the election (average of attitudes from *ED+1* and the *Post-Election Lab*) corroborated these results: The Diurnal Slope Change \times Group (Romney versus Obama) interaction also predicted attitude shifts toward the elected leader ($B=0.17$, 95% $CI[0.07, 0.27]$, $\beta=.61$, $p=.001$, $d=0.78$; see Fig. S7 and SI-Results Section 6).

3.3. Robustness checks

3.3.1. Robustness of testosterone-change predictors. We sought to determine which of the testosterone change predictors (i.e., acute reactivity or diurnal slope change) more robustly explained shifts in attitudes. We compared the model fit indices of an expanded model that included both testosterone-change predictors and their corresponding interactions (Table S19) to indices from separate models with only acute testosterone reactivity or diurnal slope change. Whereas the model that included only acute testosterone reactivity demonstrated comparable model fit to the expanded model, the model that used only diurnal slope change showed poorer model fit (see SI-Results Section 9.1). This suggests

that acute testosterone reactivity on election night was a more robust predictor of shifts in partisan attitudes compared to diurnal slope change.

3.3.2. Alternative quantification of attitudes using self-reports and behaviors on social media. We examined the association between testosterone change predictors and shifts in attitudes toward the elected leader using a different quantification of attitudes that included both self-reports and behaviors on social media (i.e., Facebook; see Table S20 for correlations between self-reports and social media behaviors). Once again, the pattern of our primary result linking acute testosterone reactivity to attitudinal shifts was observed in these analyses (see SI-Results Section 9.2.1, Table S22, and Fig. S9). A similar but weaker pattern was also found with diurnal slope changes predicting shifts in composite attitudes (see SI-Results Section 9.2.2 and Table S23).

3.3.3. Other robustness checks. We conducted additional robustness checks by: (i) including participant age and socioeconomic status as covariates in all models (see SI-Results Models B and Section 9.8 for specification curve analysis); and (ii) entering changes in positive and negative affect during the competition outcome window as covariates (see SI-Results Section 9.3). The inclusion of these covariates did not substantively change our main results. Finally, our main results were also robust to controlling for when participants learned about the election outcome, and for when they completed the *Pre-Election Lab* and *Post-Election Lab* (see SI-Results Sections 9.5 and 9.6, and Tables S24 and S25).

3.3.4. Correction for multiple comparisons. We used Benjamini's and Hochberg's (1995) false discovery rate (FDR) to correct for multiple comparisons across all our main results in this article. We found that all our statistically significant results remained statistically significant after correcting for FDR (see SI-Results Section 9.7 and Table S26).

3.3.5. Multiverse specification curve analysis. To further assess the robustness of our findings we also conducted specification curve analysis for our main results (Simonsohn et al., 2020). We identified the following data analytic choices that determined reasonable alternative specifications to our original models: (i) the choice of the regression model and analytic choices within those models, (ii) the treatment of outliers, and (iii) the manner in which diurnal slopes were calculated (see Table S27). The

results from the specification curve analysis indicated that the median effect sizes from the multiverse of each of the main findings are statistically significant (acute change in testosterone: $p < .001$; acute change in testosterone predicting shifts in attitudes: $p < .001$; changes in diurnal testosterone slopes: $p = .006$; and changes in diurnal testosterone slopes predicting shifts in attitudes: $p = .002$). Overall, this analysis indicates that our findings are robust to alternative analytic approaches (see SI-Results Section 9.8, Table S28, and Figures S10 to S13).

3.4. Summary of key supplementary analyses

3.4.1. Testosterone changes predicting long-term shifts in attitudes toward the elected

leader: In the subset of our sample that responded to a longitudinal survey ($n = 66$), we tested whether testosterone responses moderated long-term shifts in attitudes from before the election (*Pre-Election Lab* and *ED-1*) to six months after it. The association between acute and diurnal testosterone responses to the election and long-term shifts in attitudes showed a similar directional pattern albeit with smaller effect sizes compared to short-term shifts in attitudes (see SI-Results Sections 3 and 7, Tables S12 and S18, and Figs. 3 and S6; also see Table S29 for additional robustness checks). We note, however, that the confidence intervals for testosterone-moderated short- and long-term shifts overlapped, and the magnitudes of the point estimates in the analyses of the long-term effects leave open the possibility that there may be weak long-term term effects.

3.4.2. Sex differences and similarities: Competition-related acute and diurnal changes in testosterone were larger in males than females (though the effects were in the same direction for both sexes). However, we did not find evidence of sex differences in the associations between testosterone change predictors and attitudinal shifts toward the elected leader (see SI-Results Section 8; also see Table S29 for additional robustness checks).

3.4.3. Undecided voters: We included undecided voters as an exploratory comparison group and investigated differences between supporters of the losing candidate and undecided voters, across all analyses (see SI-Results). Broadly, we found that supporters of the losing candidate differed from undecided voters in the pattern of: (i) testosterone responses to the election, and (ii) testosterone-

moderated shifts in attitudes toward the elected leader. These differences between supporters of the losing candidate and undecided voters, although directionally similar to the differences between supporters of the losing and winning candidates, were not as robust or large in their effect sizes. Given these preliminary results and the importance of undecided voters in determining electoral outcomes, we encourage researchers to incorporate undecided voters when studying hormonal variation and attitudinal changes during elections.

4. Discussion

Hyper-partisan political interactions are a common and prominent feature of present-day democratic politics (Abramowitz & McCoy, 2019). We investigated the extent to which a high-stakes intergroup competition—a US presidential election— predicted changes in testosterone levels, and how competition-related testosterone responses in turn predicted shifts in attitudes toward the elected leader. We found that: (i) supporters of the losing candidate experienced acute elevations in testosterone concentrations on the night of the election compared to supporters of the winner; (ii) supporters of the losing candidate also exhibited flatter diurnal slopes starting on Election Day, and these flatter slopes persisted up to two days after the election; and (iii) among the losing candidate’s supporters, greater testosterone reactivity on election-night and flatter diurnal slopes predicted less positive attitudes toward the winning candidate after the election. Even though these two indices of testosterone change predicted a similar attitudinal shift, acute testosterone reactivity was a stronger predictor of this shift in attitudes toward the elected leader.

Previous research found evidence of the standard winner-loser effect during the 2008 US presidential election (i.e., an increase in testosterone in supporters of the winning candidate; Apicella & Cesarini, 2011; Stanton et al., 2009). In contrast, the present study highlights heterogeneity in the winner-loser effect by providing evidence of a reversal of the effect (i.e., an increase in testosterone in supporters of the losing candidate) in the 2012 US presidential election (Oliveira et al., 2013; Oliveira et al., 2014; Vongas & Hajj, 2017; Wu et al., 2017; Zilioli et al., 2014). These divergent winner-loser patterns may highlight the presence of moderators that account for variability in testosterone responses (Geniole et al.,

2017; also see Table S30 highlighting other differences between the current study and study conducted during the 2008 election; Stanton et al., 2009).

One possible moderator could be differences in the perceived closeness of the two races. Prior laboratory work suggests that the standard winner-loser effect emerges more readily in competitions perceived as decisive (i.e., the winner dominates the competition with a clear margin, e.g., Denson et al., 2013) and that the reverse winner-loser effect may occur in competitions perceived as close (i.e., there is uncertainty surrounding who might win the competition; Zilioli et al., 2014). In this vein, the 2008 election was predicted to be a relatively decisive race, with early polls favoring Obama over McCain with a 55% versus 44% chance of victory, respectively (Newport et al., 2019). By contrast, the 2012 election was expected to be a much closer race (Pew Research, 2018), with pre-election polls predicting a 49% chance of a Romney victory relative to a 48% chance of an Obama victory, even a day before the election (Gallup, 2012c). Thus, in the 2012 election, testosterone increases among Romney supporters may be attributed to the expectation of a close race. In this context, increased testosterone concentrations could reflect a motivation to regain status that is lost after an unexpected defeat (Mazur & Booth, 1998). It is, however, important to note that the electoral outcomes of both the 2008 and 2012 races turned out to be decisive Obama victories (per electoral college and popular vote percentages; Federal Election Commission, 2008, 2012). Therefore, it is possible that the *perceived* closeness of the competition rather than the *actual* closeness of the outcomes may have affected changes in testosterone concentrations. To test this experimentally, future laboratory research could manipulate predicted margins of victory prior to the outcome compared to actual margins of victory and evaluate differences in their effects on testosterone concentrations during competitive encounters.

Another possible moderator that could be explored in future work is the social status of the competitors prior to the competition. Obama rose in status—from Senator to President—with his victory in 2008, which may explain why his supporters rose in testosterone (Casto & Edwards, 2016b; Mazur & Booth, 1998). By contrast, Obama's status as US president did not change with his 2012 re-election, which may explain why we did not observe a similar rise in testosterone amongst Obama supporters in

the 2012 election (see Zilioli & Watson, 2014 and Knight & Mehta, 2017 for related laboratory research). This argument, however, does not easily explain why we detected a rise in testosterone among supporters of the losing candidate as part of the reverse winner-loser effect. Thus, it is possible that perceived closeness and prior social status operate in concert to explain the divergent winner-loser patterns (for further discussion of possible moderators of the winner-loser effect, see Casto & Edwards, 2016b; Geniole et al., 2017).

A rise in testosterone following a competitive loss may function to promote status seeking, often expressed in overt behaviors such as challenging the winner to a re-match (Carré et al., 2009; Mehta & Josephs, 2006). Because democratic elections do not allow for immediate rematches, acute testosterone increases in supporters of the losing candidate may instead manifest as negative attitudes toward the newly elected leader. Given that shifts in individual attitudes in democratic organizations have the potential to influence a leader's performance in office (Barrett & Eshbaugh-Soha, 2007), this study's results raise the possibility that acute testosterone reactivity could impact the efficacy of democratically elected leaders. Furthermore, elevated testosterone concentrations may also inhibit intergroup cooperation (Dickof, et al., 2014; Mehta et al., 2017). Therefore, an increase in acute testosterone levels following defeat in a societal intergroup competition may have broader implications for intergroup relations by impairing cooperation and deepening partisan divisions among members of competing groups.

This study's findings also suggest that the outcome of a large-scale societal competition not only predicts acute testosterone reactivity within hours of a competition but also predicts changes in day-to-day physiological functioning (i.e., diurnal rhythms). Diurnal endocrine patterns may help individuals coordinate behavior and physiology with seasonal and daily changes in environmental demands (Gorman & Lee, 2002). Here, supporters of the losing candidate exhibited flatter diurnal slopes on Election Day and at least up to two days thereafter. Moreover, changes in diurnal rhythms among the losing candidate's supporters were associated with reduced positive attitudes toward the elected leader. These findings extend the biosocial model of status by delineating longer-term impacts of intergroup competition outcomes on diurnal testosterone functioning and suggest additional directions for future work involving

diurnal testosterone. For example, studies have found that more negative attitudes toward election results are correlated with symptoms of poor mental health (e.g., event-related distress; Hagan et al., 2020). Given that variations in diurnal rhythms of testosterone have also been linked with psychopathology (e.g., anxiety-depression, behavioral problems; Granger et al., 2003), future work could consider the impact of election-related changes in testosterone's diurnal rhythms on attitudes toward elected leaders and their downstream effects on mental health outcomes.

By using a longitudinal design, we found that testosterone responses predicted shifts in attitudes toward the winner in the days after the election, but that these testosterone-moderated attitudinal shifts were smaller in magnitude six months after the election, albeit in the same direction as the short-term shifts in attitudes. Previous studies have found that short-term acute testosterone reactivity after competitive encounters predicts status-relevant behaviors one to six days after the encounter in non-human animals (Fuxjager et al., 2010; Oyegbile & Marler, 2005; Trainor et al., 2004), and one laboratory study showed a similar effect in humans (Zilioli & Watson, 2014). This timescale over days is consistent with theories of neuroendocrine function which suggest that acute hormonal effects on the brain and psychological outcomes may be short-term but not long-term (Marler et al., 2005; Oliveira et al., 2009). However, one study in military personnel found that testosterone reactivity to a laboratory stressor before deployment predicted downstream health outcomes months later (Cobb et al., 2018). The effect size in this war zone study was similar in size to the testosterone-moderated attitudinal shifts we observed six months after the election. Thus, our results leave open the possibility that there could be long-term hormonal effects, albeit weaker than the short-term effects. To explore this possibility, we recommend additional studies with larger sample sizes to detect small effects. These studies should also consider environmental factors that may influence the magnitude of long-term effects, such as factors that could extend political uncertainty surrounding electoral outcomes well past election day (e.g., vote recounts, misinformation about election results, impeachment, insurrections).

We found non-significant sex differences in testosterone-moderated attitudinal shifts. That is, testosterone responses to the election's result showed a similar association with attitude shifts in males

and females. Several previous studies have also found similar testosterone-behavior associations across sexes (Mehta et al., 2009; Stanton et al., 2011). Together with this previous work, our findings suggest that testosterone has important psychological consequences for males and females (Casto & Prasad, 2017). However, the effect of the election outcome on acute and diurnal changes in testosterone was stronger in males than in females, which is consistent with previous work that found stronger effects of competition outcomes on acute testosterone responses in males (Geniole et al., 2017; Stanton et al., 2009). One possible explanation for this sex difference is that the sex hormone estradiol may be more responsive than testosterone to competition outcomes in females (Casto & Prasad, 2017), which can be tested in new studies by quantifying estradiol levels. Another explanation is that EIAs may have been more susceptible to measurement error in females and individuals with lower testosterone concentrations. In a random subset of our samples, we assessed the validity of testosterone concentrations obtained from EIAs in comparison to those obtained from a highly accurate reference method, LC-MS/MS, across samples with lower and higher concentrations of testosterone. We note that samples with lower concentrations of testosterone (females and in the lower 50% of the sampling distribution) demonstrated less robust validity compared to samples with higher concentrations (males and upper 50%), a finding consistent with prior work (Welker et al., 2016). Future studies may benefit from using the more accurate LC-MS/MS, especially in all female or mixed-sex samples. We return to this point below.

Previous studies that have examined associations between testosterone responses to competition and psychological outcome measures have focused primarily on laboratory models (e.g., Carré et al., 2009; Mehta & Josephs, 2006), which have limited ecological validity. The present study extends this work to a naturalistic setting with significant societal importance. However, the advantages associated with an ecologically valid setting come with some loss of experimental control that a laboratory study would have when examining differences across winners and losers. Although we did control for some factors (e.g., SES) that may vary across political parties in the US political system, there may have been other ideological factors that distinguished the two groups (e.g., System justification, Right Wing Authoritarianism; Jost et al., 2009; Womick et al., 2019) that can be controlled in future work. Further,

because we tested our research questions in an ecological setting, we were unable to provide evidence of causality. Future research could pharmacologically manipulate testosterone levels to test its causal impact on competition-related shifts in intergroup attitudes and behaviors.

We chose to use immunoassays in this study to measure salivary testosterone for several reasons: cost-effectiveness for the number of samples (~2500), ease of use, expertise in conducting the hormonal analyses, and standards in the field at the time of data collection (see Yasuda et al., 2008). We were able to confirm the validity of testosterone concentrations obtained from EIAs in a random subset of our samples against a highly accurate reference method (liquid chromatography tandem mass spectrometry, LC-MS/MS). In future studies, we recommend that researchers use larger mixed-sex samples to conduct validity assessments of their assay method, which will thereby allow them to draw meaningful conclusions of their study's assay validity, especially given concerns with testosterone measurement at lower concentrations (see discussion above; Welker et al., 2016; see McCullough et al., 2013 for related discussion with oxytocin assays). We also encourage researchers to transparently report the methodological details of their assay methods and associated limitations in their manuscripts, to independently test the validity of immunoassays they are using (e.g., radioimmunoassays and chemiluminescence immunoassays), and to prioritize the use of LC-MS/MS as a method for sex-hormone analyses, whenever feasible (see discussion in Prasad et al., 2019 for more details).

This study provides initial estimates of testosterone-moderated attitudinal shifts, but future work will be helpful in improving these estimates. The effect sizes reported in this field study are consistent with effect sizes reported in other field studies (Jones & Josephs, 2006) as well as with meta-analytic evidence of larger effect sizes in the field relative to laboratory settings (Geniole et al., 2017). We hope that our detailed methods section (also see SI Materials) and freely accessible dataset will support future replication attempts and meta-analyses.

Finally, we assessed the robustness of our main findings via specification curve (multiverse) analysis. The results from this analysis indicated that our main results are robust to alternative data analytic specifications and are not dependent on the choices we made in our original analytic approach.

We encourage future researchers to also test the limits and dependencies of their own findings across reasonable alternative data analytic choices to facilitate transparency and replicability of research findings.

Humans operating within social organizations must often adapt to leaders from outgroups who are perceived not to represent the ingroup's interests. This study's results suggest that the fluctuations in testosterone resulting from intergroup competitive encounters are related to an individual's attitudinal responses toward leaders from one's outgroup. Ultimately, free and fair elections and the peaceful transfer of power are the hallmarks of democratic governance. We offer evidence toward explaining how individual-level biological activity may be associated with the dynamics of intergroup competitions and us-versus-them divides, social phenomena that can greatly dampen progress toward shared societal goals.

Author contributions statement. P.H.M., S.P., and B.L. designed the study. P.H.M., S.P., and B.L. collected the data. B.L. and S.P. immuno-assayed the saliva samples. S.P., E.L.K., and K.W. analyzed the data. S.P., E.L.K., A.S., and P.H.M. wrote the manuscript. K.W. and B.L. provided comments on the manuscript.

Acknowledgements. We are indebted to all the research assistants involved in this project. We also thank Richard Slatcher, Robert Josephs, Matthias Mehl, Megan Robbins, Colton Christian, Jeffrey Whitaker, Ron Nieva, and Sarah Donaldson.

References

- Abramowitz, A., McCoy, J., 2019. United States: Racial Resentment, Negative Partisanship, and Polarization in Trump's America. *Ann. Am. Acad. Pol. Soc. Sci.*
<https://doi.org/10.1177/0002716218811309>
- Adam, E.K., Kumari, M., 2009. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2009.06.011>
- Alabastro, A., Rast, D.E., Lac, A., Hogg, M.A., Crano, W.D., 2013. Intergroup bias and perceived similarity: Effects of successes and failures on support for in- and outgroup political leaders. *Gr. Process. Intergr. Relations*. <https://doi.org/10.1177/1368430212437212>
- Apicella, C. L., & Cesarini, D. A. 2011. Testosterone response and political donations: Winner loser effects in the 2008 presidential election. *Man is by nature a political animal: Evolution, biology and politics*, 261-272.
- Apicella, C.L., Dreber, A., Mollerstrom, J., 2014. Salivary testosterone change following monetary wins and losses predicts future financial risk-taking. *Psychoneuroendocrinology*.
<https://doi.org/10.1016/j.psyneuen.2013.09.025>
- Ariens, C. 2016. Here's What Time the Networks Called the 2012 Election. URL:
<http://www.adweek.com/tvnewser/heres-what-time-the-networks-called-the-2012-election/31011>
- Barrett, A.W., Eshbaugh-Soha, M., 2007. Presidential success on the substance of legislation. *Polit. Res. Q.* <https://doi.org/10.1177/1065912906298605>
- Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v067.i01>
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B.* <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bernhardt, P.C., Dabbs, J.M., Fielden, J.A., Lutter, C.D., 1998. Testosterone changes during vicarious experiences of winning and losing among fans at sporting events. *Physiol. Behav.*
[https://doi.org/10.1016/S0031-9384\(98\)00147-4](https://doi.org/10.1016/S0031-9384(98)00147-4)
- Boksem, M.A.S., Mehta, P.H., Van den Bergh, B., van Son, V., Trautmann, S.T., Roelofs, K., Smidts, A., Sanfey, A.G., 2013. Testosterone Inhibits Trust but Promotes Reciprocity. *Psychol. Sci.*
<https://doi.org/10.1177/0956797613495063>
- Branscombe, N.R., Wann, D.L., 1994. Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *Eur. J. Soc. Psychol.* <https://doi.org/10.1002/ejsp.2420240603>

- Carré, J.M., Campbell, J.A., Lozoya, E., Goetz, S.M.M., Welker, K.M., 2013. Changes in testosterone mediate the effect of winning on subsequent aggressive behaviour. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2013.03.008>
- Carré, J.M., Putnam, S.K., McCormick, C.M., 2009. Testosterone responses to competition predict future aggressive behaviour at a cost to reward in men. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2008.10.018>
- Casto, K. V., Edwards, D.A., 2016a. Before, During, and After: How Phases of Competition Differentially Affect Testosterone, Cortisol, and Estradiol Levels in Women Athletes. *Adapt. Hum. Behav. Physiol.* <https://doi.org/10.1007/s40750-015-0028-2>
- Casto, K. V., Edwards, D.A., 2016b. Testosterone, cortisol, and human competition. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2016.04.004>
- Casto, K. V., Edwards, D.A., Akinola, M., Davis, C., Mehta, P.H., 2020. Testosterone reactivity to competition and competitive endurance in men and women. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2019.104665>
- Casto, K. V., Prasad, S., 2017. Recommendations for the study of women in hormones and competition research. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2017.05.009>
- Chang, L.W., Krosch, A.R., Cikara, M., 2016. Effects of intergroup threat on mind, brain, and behavior. *Curr. Opin. Psychol.* <https://doi.org/10.1016/j.copsyc.2016.06.004>
- Cialdini, R.B., et al, 1976. Basking in reflected glory: Three (football) field studies. *J. Pers. Soc. Psychol.* <https://doi.org/10.1037/0022-3514.34.3.366>
- Cikara, M., Botvinick, M.M., Fiske, S.T., 2011. Us versus them: Social identity shapes neural responses to intergroup competition and harm. *Psychol. Sci.* <https://doi.org/10.1177/0956797610397667>
- Cobb, A.R., Josephs, R.A., Lancaster, C.L., Lee, H.J., Telch, M.J., 2018. Cortisol, Testosterone, and Prospective Risk for War-zone Stress-Evoked Depression. *Mil. Med.* <https://doi.org/10.1093/milmed/usy065>
- Cunningham, R.L., Lumia, A.R., McGinnis, M.Y., 2012. Androgen receptors, sex behavior, and aggression. *Neuroendocrinology*. <https://doi.org/10.1159/000337663>
- Denson, T.F., Mehta, P.H., Ho Tan, D., 2013. Endogenous testosterone and cortisol jointly influence reactive aggression in women. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2012.07.003>
- Diekhof, E.K., Wittmer, S., Reimers, L., 2014. Does competition really bring out the worst? testosterone, social distance and inter-male competition shape parochial altruism in human males. *PLoS One*. <https://doi.org/10.1371/journal.pone.0098977>
- Dreher, J.C., Dunne, S., Pazderska, A., Frodl, T., Nolan, J.J., O'Doherty, J.P., 2016. Testosterone causes both prosocial and antisocial status-enhancing behaviors in human males. *Proc. Natl. Acad. Sci. U. S. A.* <https://doi.org/10.1073/pnas.1608085113>

- Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., Fehr, E., 2010. Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature*. <https://doi.org/10.1038/nature08711>
- Federal Election Commission, 2008. Federal Elections 2008. <https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/federal-elections-2008/> (accessed 3 Dec 2020)
- Federal Election Commission, 2012. Federal Elections 2012. <https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/federal-elections-2012> (accessed 3 Dec 2020).
- Fein, S., Spencer, S.J., 1997. Prejudice as Self-Image Maintenance: Affirming the Self Through Derogating Others. *J. Pers. Soc. Psychol.* <https://doi.org/10.1037/0022-3514.73.1.31>
- Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., Tu, X.M., 2014. Log-transformation and its implications for data analysis. *Shanghai Arch. Psychiatry*. <https://doi.org/10.3969/j.issn.1002-0829.2014.02>
- Fuxjager, M.J., Forbes-Lorman, R.M., Coss, D.J., Auger, C.J., Auger, A.P., Marler, C.A., 2010. Winning territorial disputes selectively enhances androgen sensitivity in neural pathways related to motivation and social aggression. *Proc. Natl. Acad. Sci. U. S. A.* <https://doi.org/10.1073/pnas.1001394107>
- Gallup Inc. 2012a. Presidential Ratings -- Issues Approval. URL: <http://www.gallup.com/poll/1726/Presidential-Ratings-Issues-Approval.aspx>.
- Gallup Inc. 2012b. Presidential Ratings -- Personal Characteristics. Gallup. URL: <http://www.gallup.com/poll/1726/Presidential-Ratings-Issues-Approval.aspx>.
- Gallup, Inc. 2012c. Romney 49%, Obama 48% in Gallup's Final Election Survey. Gallup. URL: <http://www.gallup.com/poll/158519/romney-obama-gallup-final-election-survey.aspx>
- Geniole, S.N., Bird, B.M., Ruddick, E.L., Carré, J.M., 2017. Effects of competition outcome on testosterone concentrations in humans: An updated meta-analysis. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2016.10.002>
- Gildner, T.E., 2021. Reproductive hormone measurement from minimally invasive sample types: Methodological considerations and anthropological importance. *Am. J. Hum. Biol.* <https://doi.org/10.1002/ajhb.23535>
- Gleason, E.D., Fuxjager, M.J., Oyegbile, T.O., Marler, C.A., 2009. Testosterone release and social context: When it occurs and why. *Front. Neuroendocrinol.* <https://doi.org/10.1016/j.yfrne.2009.04.009>
- Gorman, M. R., and T. M. Lee 2002 Hormones and Biological Rhythms. In *Behavioral Endocrinology*, second ed., J. B. Becker, S. M. Breedlove, D. Crews, and M. M. McCarthy, eds. Pp. 451- 494. Cambridge: MIT Press.
- Granger, D.A., Shirtcliff, E.A., Zahn-Waxler, C., Usher, B., Klimes-Dougan, B., Hastings, P., 2003. Salivary testosterone diurnal variation and psychopathology in adolescent males and females:

- Individual differences and developmental effects. *Dev. Psychopathol.*
<https://doi.org/10.1017/S0954579403000233>
- Gray, P.B., Chapman, J.F., Burnham, T.C., McIntyre, M.H., Lipson, S.F., Ellison, P.T., 2004. Human male pair bonding and testosterone. *Hum. Nat.* <https://doi.org/10.1007/s12110-004-1016-6>
- Hagan, M.J., Sladek, M.R., Luecken, L.J., Doane, L.D., 2020. Event-related clinical distress in college students: Responses to the 2016 U.S. Presidential election. *J. Am. Coll. Heal.*
<https://doi.org/10.1080/07448481.2018.1515763>
- Herold, D.A., Fitzgerald, R.L., 2003. Immunoassays for testosterone in women: Better than a guess? *Clin. Chem.* <https://doi.org/10.1373/49.8.1250>
- Jones, A.C., Josephs, R.A., 2006. Interspecies hormonal interactions between man and the domestic dog (*Canis familiaris*). *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2006.04.007>
- Jost, J.T., West, T. V., Gosling, S.D., 2009. Personality and ideology as determinants of candidate preferences and “obama conversion” in the 2008 U.S. presidential election. *Du Bois Rev.*
<https://doi.org/10.1017/S1742058X09090109>
- Knight, E.L., Mehta, P.H., 2017. Hierarchy stability moderates the effect of status on stress and performance in humans. *Proc. Natl. Acad. Sci. U. S. A.* <https://doi.org/10.1073/pnas.1609811114>
- Kudielka, B.M., Gierens, A., Hellhammer, D.H., Wüst, S., Schlotz, W., 2012. Salivary cortisol in ambulatory assessment-some dos, some don'ts, and some open questions. *Psychosom. Med.*
<https://doi.org/10.1097/PSY.0b013e31825434c7>
- Kuzawa, C.W., Georgiev, A. V., McDade, T.W., Bechayda, S.A., Gettler, L.T., 2016. Is There a Testosterone Awakening Response in Humans? *Adapt. Hum. Behav. Physiol.*
<https://doi.org/10.1007/s40750-015-0038-0>
- Marceau, K., Ruttle, P.L., Shirtcliff, E.A., Hastings, P.D., Klimes-Dougan, B., Zahn-Waxler, C., 2015. Within-person coupling of changes in cortisol, testosterone, and DHEA across the day in adolescents. *Dev. Psychobiol.* <https://doi.org/10.1002/dev.21173>
- Marler, C.A., Oyegbile, T.O., Plavicki, J., Trainor, B.C., 2005. Response to Wingfield's commentary on “A continuing saga: The role of testosterone in aggression.” *Horm. Behav.*
<https://doi.org/10.1016/j.yhbeh.2005.05.010>
- Mazur, A., Booth, A., 1998. Testosterone and dominance in men. *Behav. Brain Sci.*
<https://doi.org/10.1017/S0140525X98001228>
- McCullough, M.E., Churchland, P.S., Mendez, A.J., 2013. Problems with measuring peripheral oxytocin: Can the data on oxytocin and human behavior be trusted? *Neurosci. Biobehav. Rev.*
<https://doi.org/10.1016/j.neubiorev.2013.04.018>
- Mehta, P.H., Josephs, R.A., 2006. Testosterone change after losing predicts the decision to compete again. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2006.07.001>

- Mehta, P.H., Lawless DesJardins, N.M., van Vugt, M., Josephs, R.A., 2017. Hormonal underpinnings of status conflict: Testosterone and cortisol are related to decisions and satisfaction in the hawk-dove game. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2017.03.009>
- Mehta, P.H., Wuehrmann, E. V., Josephs, R.A., 2009. When are low testosterone levels advantageous? The moderating role of individual versus intergroup competition. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2009.04.001>
- Newport, F., Jones, J. M. & Saad, L. 2019. Final Presidential Estimate: Obama 55%, McCain 44%. URL: <https://news.gallup.com/poll/111703/final-presidential-estimate-obama-55-mccain-44.aspx>.
- Oc, B., Moore, C., Bashshur, M.R., 2018. When the tables are turned: The effects of the 2016 U.S. Presidential election on in-group favoritism and out-group hostility. *PLoS One*. <https://doi.org/10.1371/journal.pone.0197848>
- Oliveira, G.A., Uceda, S., Oliveira, T., Fernandes, A., Garcia-Marques, T., Oliveira, R.F., 2013. Threat perception and familiarity moderate the androgen response to competition in women. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2013.00389>
- Oliveira, G.A., Uceda, S., Oliveira, T.F., Fernandes, A.C., Garcia-Marques, T., Oliveira, R.F., 2014. Testosterone response to competition in males is unrelated to opponent familiarity or threat appraisal. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2014.01240>
- Oliveira, R.F., 2009. Social behavior in context: Hormonal modulation of behavioral plasticity and social competence. *Integr. Comp. Biol.* <https://doi.org/10.1093/icb/icp055>
- Oyegbile, T.O., Marler, C.A., 2005. Winning fights elevates testosterone levels in California mice and enhances future ability to win fights. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2005.04.007>
- Packer, C., Pusey, A.E., 1982. Cooperation and competition within coalitions of male lions: kin selection or game theory? *Nature*. <https://doi.org/10.1038/296740a0>
- Peckins, M.K., Susman, E.J., 2015. Variability in diurnal testosterone, exposure to violence, and antisocial behavior in young adolescents. *Dev. Psychopathol.* <https://doi.org/10.1017/S095457941400145X>
- Pew Research Center for the People and the Press. 2018. In Deadlocked Race, Neither Side Has Ground Game Advantage. URL: <http://www.people-press.org/2012/10/31/in-deadlocked-race-neither-side-has-ground-game-advantage/>.
- Prasad, S., Lassetter, B., Welker, K.M., Mehta, P.H., 2019. Unstable correspondence between salivary testosterone measured with enzyme immunoassays and tandem mass spectrometry. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2019.104373>
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- Rand, D.G., Pfeiffer, T., Dreber, A., Sheketoff, R.W., Wernerfelt, N.C., Benkler, Y., 2009. Dynamic remodeling of in-group bias during the 2008 residential election. *Proc. Natl. Acad. Sci. U. S. A.* <https://doi.org/10.1073/pnas.0811552106>
- Raudenbush, S. W., & Bryk, A. S. 2002. *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- Reimers, L., Diekhof, E.K., 2015. Testosterone is associated with cooperation during intergroup competition by enhancing parochial altruism. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2015.00183>
- Riek, B.M., Mania, E.W., Gaertner, S.L., 2006. Intergroup threat and outgroup attitudes: A meta-analytic review. *Personal. Soc. Psychol. Rev.* https://doi.org/10.1207/s15327957pspr1004_4
- Schultheiss, O. C., & Stanton, S. J. 2009. Assessment of salivary hormones. In Harmon-Jones, E. & Beer, J. (Eds.) *Methods in Social Neuroscience*. Guilford Publications, New York.
- Simonsohn, U., Simmons, J.P., Nelson, L.D., 2020. Specification curve analysis. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-020-0912-z>
- Stanton, S. J., Lienesch, S. H., & Schultheiss, O. C. 2011. Testosterone is positively associated with risk taking in the Iowa Gambling Task. *Hormones and behavior*, 59(2), 252-256.
- Stanton, S.J., Beehner, J.C., Saini, E.K., Kuhn, C.M., LaBar, K.S., 2009. Dominance, politics, and physiology: Voters' testosterone changes on the night of the 2008 United States presidential election. *PLoS One*. <https://doi.org/10.1371/journal.pone.0007543>
- Stroup, W.W., 2015. Rethinking the analysis of non-normal data in plant and soil science. *Agron. J.* <https://doi.org/10.2134/agronj2013.0342>
- Tajfel, H., Turner, J., 1979. An integrative theory of inter-group conflict, in: *The Social Psychology of Intergroup Relations*.
- Trainor, B.C., Bird, I.M., Marler, C.A., 2004. Opposing hormonal mechanisms of aggression revealed through short-lived testosterone manipulations and multiple winning experiences. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2003.09.006>
- Trawalter, S., Chung, V.S., DeSantis, A.S., Simon, C.D., Adam, E.K., 2012. Physiological stress responses to the 2008 U.S. presidential election: The role of policy preferences and social dominance orientation. *Gr. Process. Intergr. Relations*. <https://doi.org/10.1177/1368430211428163>
- Van Bavel, J.J., Pereira, A., 2018. The Partisan Brain: An Identity-Based Model of Political Belief. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2018.01.004>
- Vongas, J.G., Al Hajj, R., 2017. The effects of competition and implicit power motive on men's testosterone, emotion recognition, and aggression. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2017.04.005>

- Welker, K.M., Lassetter, B., Brandes, C.M., Prasad, S., Koop, D.R., Mehta, P.H., 2016. A comparison of salivary testosterone measurement using immunoassays and tandem mass spectrometry. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2016.05.022>
- Williams, J.M., Pusey, A.E., Carlis, J. V., Farm, B.P., Goodall, J., 2002. Female competition and male territorial behaviour influence female chimpanzees' ranging patterns. *Anim. Behav.* <https://doi.org/10.1006/anbe.2001.1916>
- Wilson, M.L., Wrangham, R.W., 2003. Intergroup relations in Chimpanzees. *Annu. Rev. Anthropol.* <https://doi.org/10.1146/annurev.anthro.32.061002.120046>
- Womick, J., Rothmund, T., Azevedo, F., King, L.A., Jost, J.T., 2019. Group-Based Dominance and Authoritarian Aggression Predict Support for Donald Trump in the 2016 U.S. Presidential Election. *Soc. Psychol. Personal. Sci.* <https://doi.org/10.1177/1948550618778290>
- Wu, Y., Eisenegger, C., Zilioli, S., Watson, N. V., Clark, L., 2017. Comparison of clear and narrow outcomes on testosterone levels in social competition. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2016.05.016>
- Yasuda, M., Honma, S., Furuya, K., Yoshii, T., Kamiyama, Y., Ide, H., Muto, S., Horie, S., 2008. Diagnostic significance of salivary testosterone measurement revisited: using liquid chromatography/mass spectrometry and enzyme-linked immunosorbent assay. *J. Mens. health.* <https://doi.org/10.1016/j.jomh.2007.12.004>
- Zilioli, S., Mehta, P.H., Watson, N. V., 2014. Losing the battle but winning the war: Uncertain outcomes reverse the usual effect of winning on testosterone. *Biol. Psychol.* <https://doi.org/10.1016/j.biopsycho.2014.07.022>
- Zilioli, S., Watson, N. V., 2014. Testosterone across successive competitions: Evidence for a “winner effect” in humans? *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2014.05.001>

Figures

Pre-election lab	ED – 2 days	ED – 1 day	Election Day (ED)	ED + 1 day	ED + 2 days	Post-election lab	6-month follow up
	Salivary testosterone	Salivary testosterone	Salivary testosterone	Salivary testosterone	Salivary testosterone		
Attitudes toward the elected leader*		Attitudes toward the elected leader*		Attitudes toward the elected leader		Attitudes toward the elected leader	Attitudes toward the elected leader

Fig. 1. Timeline of the study. *Salivary testosterone* was measured when participants woke up (*Wake-up*), 30 minutes after they woke up (*Wake-up+30 mins*), at 3 pm (*Afternoon*), and when they went to bed (*Bedtime*). Additional samples were collected at 5pm, 7pm, and 9pm on Election Day (ED). *Attitudes toward the elected leader* (*then presidential candidate) were measured at five time points: the Pre-Election lab session conducted on average 4.37 ($SD=1.14$) days before the election (*Pre-Election Lab*), one day before the election (*ED-1 day*), one day after the election (*ED+1 day*), the Post-Election lab session (*Post-Election Lab*) conducted on average 4.46 ($SD=1.33$) after the election, and the six-month follow up.

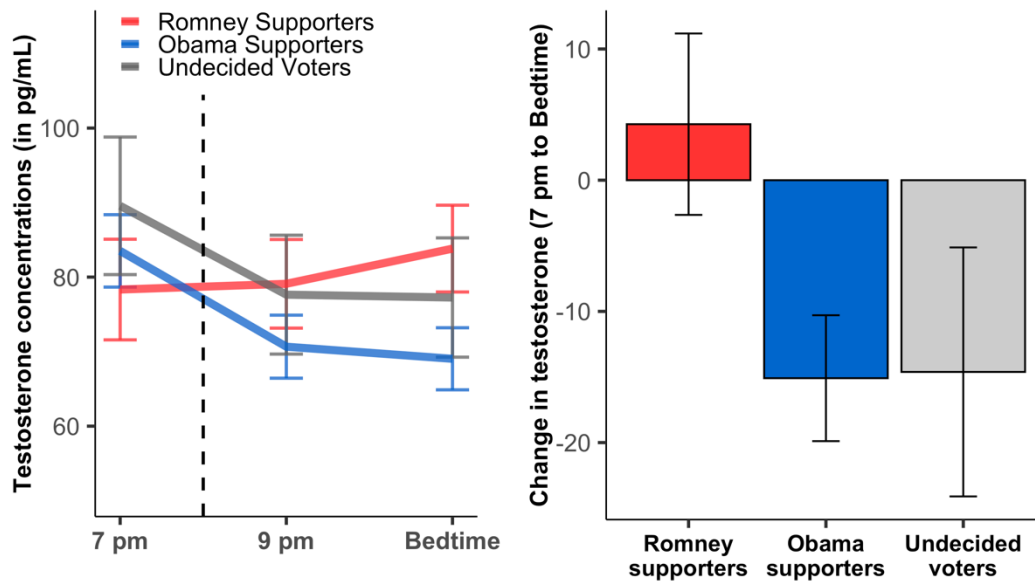


Fig. 2. Acute testosterone reactivity to the election outcome. **Left Panel:** Testosterone levels (in pg/mL) during the competition-outcome window: 7pm, 9pm, and at *Bedtime* on election night. Solid lines represent estimated marginal means (controlling for sex) and error bars = SEs. The vertical black dashed line indicates the time at which the election results were declared. **Right Panel:** Acute change in testosterone (in pg/mL) from before to after the election results were announced (*Bedtime* minus 7pm). Bars represent estimated marginal means of acute change (controlling for sex) and error bars = $\pm 1SEs$.

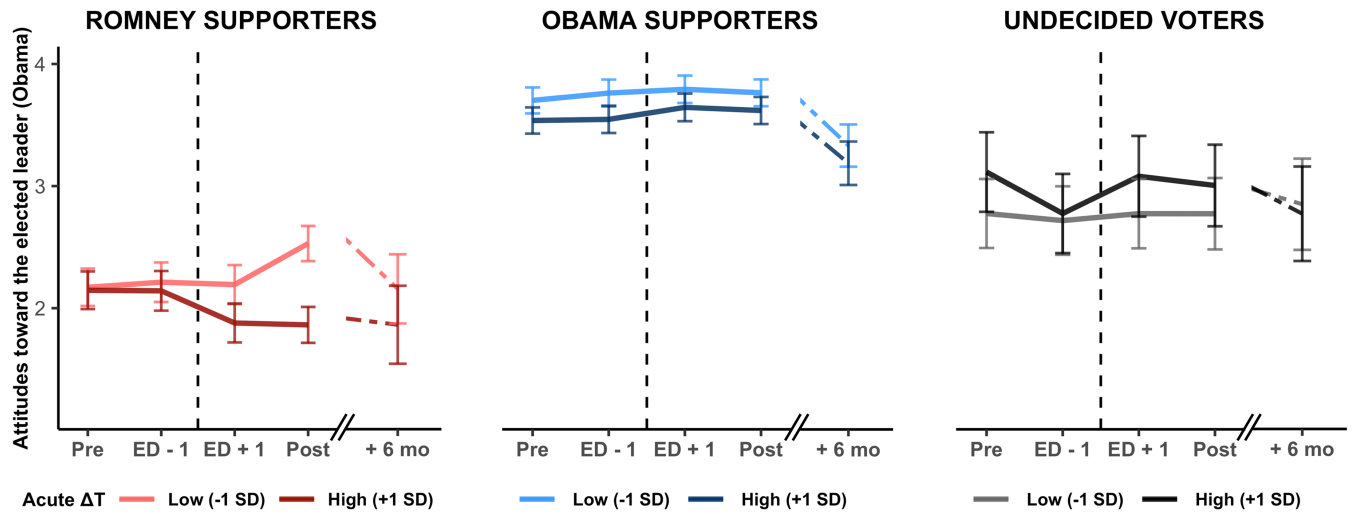


Fig. 3. Attitudes toward the elected leader (Obama) as a function of acute testosterone reactivity $\pm 1 SD$ around the mean for each political group. Attitudes toward the elected leader were reported at the pre-election lab session (Pre), one day before the election (ED-1), one day after the election (ED+1), during the post-election lab session (Post), and in a subset of our sample at a six-month follow-up (+6mo). Colored lines represent estimated marginal means (controlling for sex) and error bars = $\pm 1 SE$. The vertical black dashed line between ED-1 and ED+1 indicates when election results were declared.

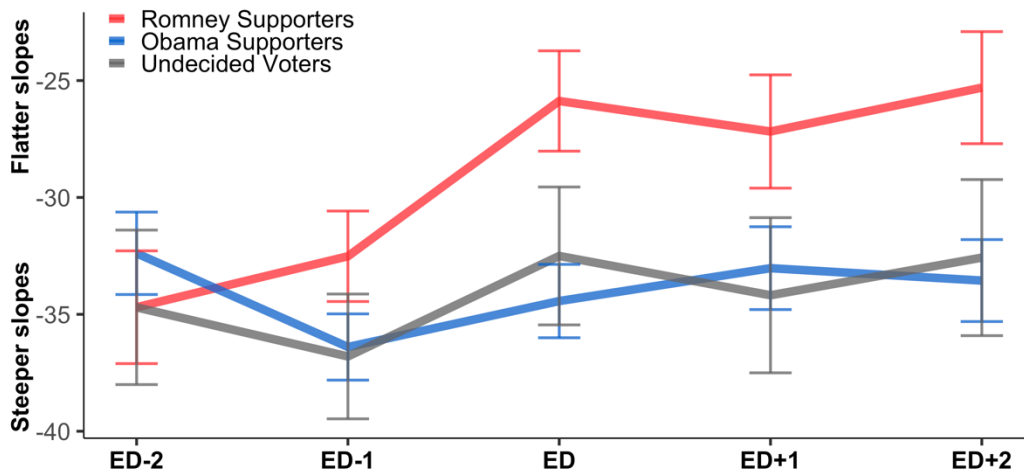


Fig. 4. Testosterone diurnal slopes across the naturalistic sampling period: 2 days before the election (ED - 2), 1 day before the election (ED-1), Election Day (ED), 1 day after the election (ED+1), and two days after the election (ED + 2). Solid lines represent estimated marginal means (controlling for sex) and error bars = ± 1 SE.

Supplementary Information (SI)

**Testosterone fluctuations in response to a democratic election predict partisan attitudes
toward the elected leader**

Smrithi Prasad*, Erik L. Knight*, Amar Sarkar, Keith M. Welker, Bethany Lassetter, Pranjali H.
Mehta

*shared first authorship

Table of Contents

SI METHODS.....	5
1. Political group classification	5
Table S1. Means and SDs of Age, SES, and Political Investment across the three political groups.	6
Table S2. Frequencies (in % of total available data) of race within each political group.	6
2. Confirmation of voting status and voting details	6
Table S3. Frequencies (in % of available data) for the reported voting process and the context within which participants found out about the election.	7
3. Salivary testosterone.....	7
Table S4. Average time (in hours) after awakening for each sample across the naturalistic sampling period	8
Table S5. Correlations, Deming Regression (DR) Analysis, fixed and proportional bias estimates of Salimetrics EIAs compared to mass spectrometry (LC-MS/MS; also reported in Prasad et al., 2019).	9
4. Attitudes toward the elected leader	9
5. Affective states	11
6. Statistical analysis	11
7. Power Simulations.....	13
Fig. S1. Power simulation results..	15
Table S6. Key strengths of the study.....	16
SI RESULTS.....	18
1. Determination of the competition-outcome window.....	18
Fig. S2. Positive affect (composite of happy, excited, powerful, and confident) and negative affect (sadness) across Election Day	18
Table S7. Means and SDs of positive affect (composite of happy, excited, powerful, and confident) and negative affect (sadness) reported on Election Day.....	19
Table S8. MLMs examining changes in positive affect (composite of happy, excited, powerful, and confident) and negative affect (sadness) across Election Day and across political groups.	19
Table S9. Means and SDs of testosterone levels (in pg/mL) across the competition outcome window.....	20
Table S10. MLMs examining changes in testosterone levels across the competition-outcome window and across political groups.	20
Figure S3. Pirate plot of acute changes in testosterone across the competition-outcome window (standardized within sex) predicted by political groups.	21
Table S11. MLMs examining short-term shifts in attitudes toward the elected leader (Obama) moderated by acute testosterone reactivity during the competition-outcome window, across political groups	22
2. Linear regression of acute testosterone reactivity predicting shifts in attitudes toward the elected leader.....	23
Fig. S4. Scatter plot between acute testosterone reactivity during the competition-outcome window and changes in attitudes toward the elected leader from before the election to after the election, across political groups.....	23

3. Acute testosterone reactivity predicting long-term shifts in attitudes toward the elected leader.....	24
Table S12. MLMs examining long-term shifts in attitudes toward the elected leader (Obama) moderated by acute testosterone reactivity during the competition-outcome window	25
Table S13. Means and SDs of diurnal slope values (lower values indicate steeper slopes, whereas higher values indicate flatter slopes) across the sampling period.....	26
Table S14. MLMs examining changes in diurnal slopes across the naturalistic sampling period.....	26
4. Follow-up analyses of changes in diurnal slopes across the sampling period.	27
5. Changes in diurnal slopes on Election Day.....	27
Table S15. Means and SDs of testosterone levels (in pg/mL) across Election Day.	27
Table S16. MLMs examining changes in testosterone levels across Election Day as a function.....	28
Fig. S5. Testosterone levels (in pg/mL) across Election Day	29
Table S17. MLMs examining short-term shifts in attitudes toward the elected leader (Obama) moderated by diurnal slope change from ED-1 to ED, across political groups.	30
Fig. S6. Attitudes toward the elected leader (Obama) moderated by changes in diurnal slopes from ED-1 to ED	31
6. Linear regression analysis of diurnal slope changes predicting shifts in attitudes toward the elected leader.....	32
Fig. S7. Scatterplot between diurnal slope change from ED-1 to ED (standardized within sex) and changes in attitudes toward the elected leader (Obama) from before the election to after the election.....	32
7. Diurnal slope changes predicting long-term shifts in attitudes toward the elected leader.	32
Table S18. MLMs examining long-term shifts in attitudes toward the elected leader (Obama) moderated by diurnal slope change	34
8. Sex differences and similarities.....	35
9. ROBUSTNESS CHECKS.....	37
9.1. Model comparisons between acute testosterone reactivity and diurnal slope changes predicting shifts in attitudes toward the elected leader.	37
Table S19. MLMs examining changes in attitudes toward Obama) moderated by acute testosterone reactivity and diurnal slope changes from ED-1 to ED, across political groups.....	38
9.2. Attitudes toward elected leader expressed via behaviors on Facebook.	39
Fig. S8. Attitudes toward the elected leader via online behaviors on Facebook from before the election (Pre-Election: from ED-2 days through 8 pm on ED) to after (Post-Election: from 8 pm on ED through ED+2 days).	40
Table S20. Correlations between elected leader (Obama) via self-reports and online behaviors on Facebook.	41
Table S21. Means and SDs of attitudes toward the elected leader reported	41
Table S22. MLMs examining changes in attitudes toward the elected leader (Obama) using a composite measure of self-reported attitudes and online social media behavior moderated by acute testosterone reactivity during the competition-outcome window	43
Fig. S9. Composite attitudes toward the elected leader (Obama) as a function of acute testosterone reactivity during the competition-outcome window.....	44

Table S23. MLMs examining changes in attitudes toward the elected leader (Obama) using a composite measure of self-reported attitudes and online social media behavior	45
9.3. Affective change as a covariate	46
9.4. Main results after the exclusion of the Gary Johnson supporter	47
9.5. Robustness checks based on when participants learned about the outcome of the election	48
Table S24. Robustness check based on when participants learned about the election outcome.	48
9.6. Robustness checks based on when participants completed the Pre- and Post-Lab Session	49
Table S25. Robustness check based on when participants completed the Pre- and Post-Election Lab sessions for analysis examining shifts in attitudes toward the elected leader (Obama), moderated by acute testosterone reactivity and diurnal slope changes from ED-1 to ED.....	49
9.7. False Discovery Rate Analysis.....	50
Table S26. Correction for multiple comparisons across our main research questions using the Benjamani & Hochberg (1995) false discovery rate (FDR) method	50
9.8. Multiverse specification curve analysis.....	51
Table S27. Original specifications and alternative reasonable specifications used to test our main research questions.....	51
Table S28. Results from the specification curve analyses across the various data analytic choices	52
Fig. S10. Results of the specification curve analysis examining acute changes in testosterone across the competition-outcome window.	53
Fig. S11. Results of the specification curve analysis examining acute change in testosterone predicting shifts in attitudes toward the elected leader.	54
Fig. S12. Results of the specification curve analysis examining changes in diurnal slope across the naturalistic sampling period.	55
Fig. S13. Results of the specification curve analysis examining diurnal slope changes predicting shifts in attitudes toward the elected leader.	56
9.9. Robustness checks for Supplementary Results	57
Table S29. Results from robust MLMs for the following supplementary results.....	57
11. Comparison between the current study and a study conducted during the 2008 US Presidential Election (Stanton et al., 2009)	59
Table S30. Differences and similarities between the research questions, methods, and results in the current study (conducted during the 2012 election) and a study during the 2008 election (Stanton et al., 2009).	59
SI REFERENCES	64

SI METHODS

1. Political group classification

Approximately two weeks prior to the 2012 US presidential election, we asked participants who they intended to vote for in the upcoming election: *Definitely Obama*, *Probably Obama*, *Probably Romney*, *Definitely Romney*, *Not sure*, and *Other candidates* (Jost et al., 2009). Participants who indicated that they would *Definitely* or *Probably* vote for Obama were categorized as supporters of the winning candidate and referred to as *Obama supporters*. Participants who indicated that they would *Definitely* or *Probably* vote for Romney or supported other losing candidates (e.g., Gary Johnson, $n=1$) were classified as supporters of the losing candidate and referred to as *Romney supporters*. We included the one Gary Johnson supporter within the losing group because we were broadly interested in comparisons between supporters of winners and losers and the inclusion of this individual boosted statistical power of the losing group. For brevity and ease, we refer to the losing group collectively as “*Romney supporters*”. Our main results remain unchanged after the exclusion of the individual supporting Gary Johnson (see SI Results). Finally, participants who were unsure of their support for either leader were categorized as *undecided*. A total of 62 Obama supporters (56.45% Female), 33 Romney supporters (48.48% Female)¹, and 18 undecided voters (77.78% Female) were recruited to participate in the study.²

To the extent possible, participants were matched across political groups across several variables. These were: (a) *demographic characteristics*, such as age and subjective socio-economic status (measured using the MacArthur Scale of Subjective Social Status; Adler et al., 2000), and (b) *political investment*, which was measured using the following items: ‘*How invested do you consider yourself in the upcoming election?*’ (1= Very Invested to 4= Very Uninvested; reverse coded); ‘*How often do you consume media about the upcoming election?*’ (1= Never to 5= Very often) ‘*How often do you talk about the upcoming election in your free time?*’ (1= Never to 5= Very often). The latter two political investment items that correspond to talking about and consuming media about the election were rescaled to a 4-point Likert scale (by subtracting 1) because no participant selected the “Never” option for those two questions. All three items related to political investment were then averaged to create a composite measure of political investment in the election ($\alpha=.76$). See Table S1 for means and *SDs* of the demographic characteristics (i.e., age and SES) and political investment levels across political groups. Obama supporters and Romney supporters did not differ from each other on either demographic characteristic ($ps>.10$). However, undecided voters reported significantly lower SES than Romney supporters ($B=-1.42$, 95%*CI*[-2.51, -0.34], $p=.011$) and marginally lower SES than Obama supporters ($B=-0.98$, 95%*CI*[-1.97, 0.01], $p=.053$). Undecided voters were also older than Romney supporters ($B=5.17$, 95%*CI*[0.34, 10.00], $p=.036$) and Obama supporters ($B=5.64$, 95%*CI*[1.23, 10.06], $p=.013$). All three political groups did not differ on their political investment in the election ($ps>.10$). The race and ethnicities of our sample’s participants across the three political groups were representative of the location of the study (i.e., Oregon; see Table

¹ The skewed distribution in political support in our sample was representative of the political ideologies shared by those in the Eugene/Springfield area in Oregon. The majority of those who participated in our surveys were Democrats and therefore most of our sample were Obama supporters.

² After the initial screening process, we recruited 114 participants. However, due to technical difficulties (i.e., not receiving the evening surveys or text message reminders to complete the daily diary), one participant was not able to complete critical portions of the study and therefore was excluded from all analyses.

S2 for frequency of races and ethnicities across political groups). To compare differences in the frequencies of races and ethnicities across political groups in our study, and given lower representation of people of color across racial/ethnic groups in our sample, we classified our participants as *White* or *People of Color* for ease of interpretability. A chi-square test of independence revealed no significant differences in race/ethnicity across political groups ($\chi^2=3.29, p=.20$).

Table S1. Means and SDs of Age, SES, and Political Investment across the three political groups.

	<i>Romney Supporters</i>		<i>Obama Supporters</i>		<i>Undecided Voters</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SES	6.09	2.01	5.65	1.70	4.67	2.17
Age (years)	23.94	7.98	23.47	7.38	29.11	11.52
Political Investment	3.65	0.70	3.51	0.65	3.33	0.66

Table S2. Frequencies (in % of total available data) of race within each political group.

<i>Race</i>	<i>Romney Supporters</i>	<i>Obama Supporters</i>	<i>Undecided Voters</i>
White	88%	72%	69%
People of Color	12%	28%	31%
African/African American	0%	5%	6%
Asian/Asian American	0%	5%	13%
Hispanic/Latino	6%	7%	6%
Middle Eastern	0%	2%	0%
Native American	6%	3%	0%
Pacific Islander	0%	3%	0%
Multiracial	0%	3%	6%

1.1. Longitudinal follow-up. A subset of our sample ($n=66$; $M_{age}=24.48$; $SD=8.02$) responded to a longitudinal follow-up survey, administered six months after the election. This follow-up sample consisted of 31 Obama supporters (54.84% Female), 21 Romney supporters (42.86 % Female) and 14 Undecided voters (78.57% Female).

2. Confirmation of voting status and voting details

On average, participants reported voting on October 31, 2012; it was possible for participants to vote prior to Election Day because in the year 2012, Oregon was one of the three states in the US where voters had the option to mail in their ballots. On average, participants reported learning about the results on the night of the election (November 6, 2012) at 8:50 pm PST; networks called the election at approximately at 8:12 pm PST (Ariens, 2016). To confirm if participants voted in the 2012 US presidential election, we obtained public voting information from the local election office. Out of those participants with available voting records ($n=93$), 94.62% ($n=88$) voted in the 2012 election. Some participants did not have voting records in the Oregon database. It is possible that these participants voted in other states or that the names and addresses that they provided did not match those in the Oregon database. Participants also retrospectively reported (a) how and when they cast their ballot and (b) the social context in

which they found out about the electoral outcome: the source of the results, and where and with whom they found out about the election result (see Table S3).

Table S3. Frequencies (in % of available data) for the reported voting process and the context within which participants found out about the election. Note: the frequencies are not mutually exclusive.

VOTING		ELECTION RESULTS					
Voting Process	%	Source of results	%	Social Context	%	Location	%
Mail in	74.29	Media	88.46	Alone	23.81	Apt/Home/Dorm	72.40
Dropbox	20.00	TV	61.54	Family	20.00	Public place	13.21
Absentee	3.81	Internet	30.77	Friends	48.57	Work	8.49
Booth	1.9	Social Media	2.88	Peers	19.05	In-transit	1.89
		Social interaction	17.31	Co-workers	1.90	Other	3.77
				Strangers	9.52		
				Partner	7.62		
				Roommates	2.86		

3. Salivary testosterone

3.1. Collection and storage. Saliva samples were collected via passive drool into polypropylene tubes. To control for the effects of food intake on endocrine functioning, participants were asked not to eat or drink anything besides water 30 minutes before they provided each sample and were asked to rinse their mouths prior to sample collection. Participants were asked to not brush their teeth in the 30 minutes that separated the two morning samples. After collection, participants were asked to immediately store all their samples in a freezer at home in order to preserve the integrity of the sample and prevent hormone degradation and mucin precipitation. If participants did not have access to a freezer, they were asked to store their samples temporarily in a box with icepacks (which were provided by the experimenters) during the day, before they transferred the samples to a freezer at home. All sample collection and storage procedures adhered to standard practices in the literature (Schultheiss & Stanton, 2009; Van Anders et al., 2014; Adam & Kumari, 2009).

3.2. Compliance. All take-home saliva samples were self-administered, and to ensure compliance with standard saliva collection procedures, participants were provided with detailed instructions for at-home saliva collection (Adam & Kumari, 2009; Kudielka et al., 2012; see <https://osf.io/xkbuc/> for instructions provided to the participants). Consistent with prior research, participants were also sent automated text messages using a bulk text message service—Red Oxygen—prior to each instance they provided a saliva sample, as both a reminder and means to increase likelihood of compliance (Fernandes et al., 2013). Additionally, after providing each sample, participants time-stamped their daily diaries using an electronic time-stamper (Dymo Corporation, Stamford, Connecticut) (see Table S4 for average hours since awakening for each sample). On average we were able to obtain time-stamp data for 97.97% of the samples that were collected. Further, we did not find any differences on the average time when each sample was provided across political groups ($ps > .10$).

Table S4. Average time (in hours) after awakening for each sample across the naturalistic sampling period. ED - 2: Two days before the election; ED - 1: One day before the election; ED: Election Day, ED + 1: One day after the election; ED + 2: Two days after the election.

	ED - 2		ED - 1		ED		ED + 1		ED + 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Wakeup+30	0.60	0.28	0.51	0.16	0.52	0.12	0.56	0.23	0.59	0.38
Afternoon	6.48	1.66	7.45	1.59	7.13	1.78	7.09	1.82	7.05	2.05
5 pm					9.11	1.82				
7 pm					11.16	1.65				
9 pm					13.11	1.82				
Bedtime	14.17	1.92	15.12	1.89	15.74	1.84	15.14	1.93	14.97	2.18

3.3. Testosterone assays. At the end of the study, when participants dropped off their frozen samples, the samples were immediately stored at -80 °C in the Social Psychoneuroendocrinology Laboratory at the University of Oregon until they were assayed. Before assaying, all samples were thawed to room temperature for 45 minutes, vortexed, centrifuged for 10 minutes at 3500 rpm, and then aliquoted into micro-centrifuge tubes and re-stored at -80 °C. Individual aliquots were thawed, vortexed, and centrifuged a second time before the assay process. Samples were assayed in duplicates using commercially available enzyme-linked immunoassay (EIA) kits (Salimetrics LLC; State College, PA) following standard protocols and practices in the literature (Schultheiss & Stanton, 2009).

3.4. Validity information of testosterone immunoassays. In a random selection of 96 samples from our study³, we assessed the validity of testosterone concentrations obtained from EIAs to those obtained from mass spectrometry (LC-MS/MS) — the gold standard of hormone measurement (Prasad et al., 2019). To comprehensively assess the validity of testosterone levels obtained from EIAs, we examined: (i) correlations between testosterone levels from EIAs and those from LC-MS/MS, (ii) the degree of convergence of testosterone levels from EIAs with those from the reference method (LC-MS/MS) using Deming regression analysis (Martin, 2000), (iii) mean differences in testosterone levels obtained from EIAs relative to those from LC-MS/MS (fixed bias analysis), and (iv) degree of inflation, if any, in testosterone levels from EIAs compared to those from LC-MS/MS across the sample distribution (proportional bias).

As noted in the main manuscript, testosterone concentrations obtained from EIAs correlated well with those obtained from LC-MS/MS across the entire sample, and at higher and lower concentrations, with narrower confidence intervals at higher concentrations (see Table S5). We used Deming regression analysis to approximate how closely the relationship between EIAs and LC-MS/MS conformed to an identity line which assumes equality between methods (intercept= 0, slope= 1). Although the EIA—LC-MS/MS slope did not show direct one-on-one correspondence with the line of identity (95% CI for the slope did not include 1), it did closely approximate the line of identity: the slope was close to 1 and 95% CIs were narrow (see Table S5). Males, and samples in the upper 50% of distribution demonstrated good convergence with the line of identity (slope was close to 1 and 95% CIs were narrow). However, females and samples in the lower 50% of the distribution did demonstrate some deviations from the line of

³ Testosterone concentrations from EIAs reported in Welker et al. (2016) were obtained after re-assaying the random subset of samples for which we report methodological validity in the current paper.

identity (see Table S5), but these deviations were smaller in size and the CIs were narrower compared with those reported in Welker et al. (2016).

Fixed bias analysis revealed that testosterone values obtained from EIAs were significantly inflated compared to those obtained from MS ($p < .001$)—a pattern of results consistent with when immunoassays are used to estimate testosterone (Schultheiss et al., 2019). Fixed bias was found across the entire sample, in both sexes, and in the upper and lower 50%. We did not find evidence of proportional bias in our sample subset. We note that whereas proportional bias was non-significant or marginal in males and in the upper 50% of the sample, there was evidence of proportional bias in females and in the lower 50% of the sample (greater inflation of testosterone was found at higher concentrations).

In summary, we found that EIAs that were conducted in a random subset of our sample demonstrated good accuracy in estimating testosterone concentrations. However, the performance of EIAs was better in samples with higher concentrations of testosterone (i.e., in males and in the upper 50% of the sampling distribution) compared to the performance in samples with lower concentrations of testosterone (i.e., in females and in the lower 50% of the sampling distribution). In future work, we recommend that researchers use larger mixed-sex samples to conduct validity assessments of EIAs, especially given concerns with testosterone measurement at lower concentrations.

Table S5. Correlations, Deming Regression (DR) Analysis, fixed and proportional bias estimates of Salimetrics EIAs compared to mass spectrometry (LC-MS/MS; also reported in Prasad et al., 2019). Analyses were conducted using the overall random sample subset ($n=96$), in males, females, in the upper 50% of the sampling distribution, and in the lower 50% of the sampling distribution.

	Correlations	DR: Intercept	DR: Slope	Fixed Bias	Proportional Bias
Overall Sample ($n=96$)	.80** [.72, .86]; $p < .001$	40.59 [26.86, 54.32]	1.26 [1.07, 1.45]	53.29 [45.55, 61.03]; $F_{(1, 95)}=186.90$, $p < .001$, $\eta^2=.66$	$r = -.04$, $p = .679$
Males ($n=42$)	.54** [.28, .72]; $p < .001$	3.47 [-80.55, 87.49]	1.56 [0.78, 2.35]	57.96 [41.44, 74.47]; $F_{(1, 41)}=50.22$, $p < .001$, $\eta^2=.55$	$r = -.28$, $p = .075$
Females ($n=54$)	.66** [.47, .79]; $p < .001$	18.85 [1.29, 36.41]	3.77 [2.56, 4.97]	49.66 [44.13, 55.19]; $F_{(1, 53)}=324.20$, $p < .001$, $\eta^2=.86$	$r = .35$, $p = .01$
Upper 50% ($n=48$)	.65** [.44, .79]; $p < .001$	15.03 [-36.75, 66.81]	1.46 [0.95, 1.97]	55.65 [41.43, 69.86]; $F_{(1, 47)}=62.01$, $p < .001$, $\eta^2=.57$	$r = -.18$, $p = .234$
Lower 50% ($n=48$)	.62** [.41, .77]; $p < .001$	-13.07 [-44.09, 17.95]	8.48 [5.31, 11.65]	50.93 [44.18, 57.68]; $F_{(1, 47)}=230.60$, $p < .001$, $\eta^2=.83$	$r = .49$, $p < .001$

4. Attitudes toward the elected leader

4.1. Attitudes toward the elected leader via self-report. Leadership capabilities of the elected leader and the leader's capability to handle social, economic, and political issues were assessed using scales adapted from polling surveys that resemble presidential approval ratings (Gallup, 2012a; Gallup, 2012b). *Leadership capabilities* were assessed using the following items: 1) Is a strong and decisive leader; 2) Can manage the government effectively; 3)

Understands the problems Americans face in their daily lives; 4) Shares your values; 5) Is more likely to keep his campaign promises; 6) Cares about the needs of people like you; 7) Is honest and trustworthy; 8) Is likeable; 9) Would stand up to special interests, including those aligned with his own party; 10) Would work well with both parties to get things done in Washington; 11) Can get things done; and 12) Agrees with you on issues you care about. These 12 items were measured using a 5-point Likert scale (1=Strongly Disagree to 5= Strongly Agree). ***The leader's ability to handle social, economic, and political issues*** was assessed by asking participant how well the leader would handle: 1) the economy; 2) the federal budget deficits; 3) employment; 4) healthcare; 5) foreign policy; 6) taxes; 7) Medicare; 8) energy; 9) social issues such as gay marriage and abortion; and 10) terrorism and international threats. These 10 items were measured using a 5-point Likert scale (1=Poor to 5= Excellent). The attitude-measure was administered via evening surveys at the following five time points: *Pre-election Lab*, *ED - 1 day*, *ED + 1 day*, *Post-election Lab*, and at *the six-month follow-up*. Across all time points, each of the 22 items demonstrated high reliability ($\alpha=.99$), and the two subscales were correlated (all $r_s > .69$). Therefore, these subscales were averaged to form a composite measure of self-reported attitudes toward the elected leader (Obama) at each of the five time points.

4.2. Attitudes toward the elected leader via behaviors on Facebook. Online behaviors that reflected attitudes toward the elected leader (Obama) were coded in a subset of participants ($n=92$). These participants consented to having their social media activity recorded via their activity logs on Facebook after the study's completion. To maintain consistency with the timeline of our self-report measures, Facebook activity was coded during the naturalistic sampling period: one *Pre-Election* rating was obtained before the election (from ED-2 days until 8 pm on Election Day), and one *Post-Election* rating was obtained after the election (from 8 pm on ED through ED+2 days). Two additional control files were created from the week prior to the election periods that were matched to these time windows resulting in four sets of data per participant.

Three independent coders assessed the degree to which participants displayed behaviors that corresponded to attitudes about the perceived abilities of Obama as a presidential candidate before, and elected leader after, the election using a 5-point Likert scale. These three research assistants read through the Facebook activity logs and then rated the degree to which they perceived the participant to have expressed views, liked posts, commented on others' posts, and shared materials that reflected their attitudes toward Obama as a leader. Prior to the actual coding process, a random selection of files from 10 participants was used to calculate a reliability score, resulting in a good average inter-rater reliability ($\alpha=.88$) across the measure of attitudes toward Obama. Because we had good reliability across our three coders, we divided the participant files amongst the three coders. The files within each participant (pre-election, post-election, and the two control time windows) were then randomized in order to prevent dependencies within each coder.

4.3. Composite attitudes toward the elected leader across self-report and behaviors on Facebook. Despite a substantial portion of participants consenting access to their Facebook activity logs (81.42% of our sample), after coding all the data only 50.44% ($n=57$)⁴ of

⁴ Out of these 57 individuals only 40 had complete data across both Pre-Election and Post-Election time windows.

participants had available and relevant data for us to code. Therefore, to maximize statistical power, composite scores of attitudes toward the elected leader were calculated by averaging *Pre-Election* Facebook ratings with self-reported attitudes obtained during the *Pre-Election Lab* session and *ED-1 day*, and averaging *Post-Election* Facebook ratings with self-reported attitudes obtained at *ED+1 day* and during the *Post-Election Lab* session.

5. Affective states

We measured in-the-moment affective states in our participants using daily diaries. Participants reported their naturally occurring affective states on each day of the sampling period while providing the following saliva samples on Election Day: *Wake-up+30*, in the *Afternoon*, at *5pm*, *7pm*, *9pm*, and at *Bedtime*. To increase the likelihood of compliance, we sent automated text messages using a bulk text message service—Red Oxygen—prior to each instance they provided a saliva sample and completed a diary entry. Positive affect was measured using a composite of the following items: “happy,” “excited,” “powerful,” and “confident” ($\alpha=.93$). Negative affect was measured using a single-item measure of sadness. All items were measured using 5-point Likert scales (1= Not at all to 5= Extremely).

6. Statistical analysis

6.1. Advantages of MLMs over traditional methods of analysis. Longitudinal MLMs provide several advantages over traditional methods of analysis (e.g. ANOVAs) (see Raudenbush & Bryk, 2002; Hoffman & Rovine, 2007; Steele, 2008; Locascio & Atri, 2011; Kwok et al., 2008). MLMs not only estimate average regression parameters (i.e., fixed effects), but also estimate individual-level variance that allows the intercepts and slopes for each participant to vary randomly. MLMs therefore provide greater flexibility in the modelling approach, compared to traditional methods of analysis that do not estimate random individual-level variance, make assumptions about errors being independent, require that correlations among observations across all levels of time be homogeneous (i.e., the sphericity assumption) in within-subjects designs, and require the relationship between the covariate and the outcome to be the same across the different groups (homogeneity of regression slopes assumption) in models with covariates. Unlike traditional models, MLMs do not use listwise deletion of data. Individuals with incomplete data are retained in the analysis, thereby boosting statistical power. Finally, using MLMs also allows a focused approach when analyzing both the main effects and the interactions, via the use of planned contrasts (Schad et al., 2020) as opposed to non-focused omnibus tests that require follow up *post hoc* tests. In this study, the use of planned contrasts helped us test differences: 1) between specific political groups, and 2) in linear and non-linear patterns of change across time, as opposed to non-specific omnibus tests.

6.2. Equations for the models. Below are the equations for the multilevel models conducted in the study.

Acute fluctuations in testosterone across the competition-outcome window

Level 1- Within-individual variability:

$$\text{Testosterone}_{ti} = \beta_{0i} + \beta_{1i}(\text{Linear Time}_{1i}) + \beta_{2i}(\text{Quadratic Time}_{2i}) + e_{ti}$$

Level 2- Between-individual variability⁵:

$$\begin{aligned}
\beta_{0i} &= \gamma_{00} + \gamma_{01}(\text{Romney versus Obama supporters}_{01}) \\
&\quad + \gamma_{02}(\text{Romney supporters versus Undecided voters}_{02}) + \gamma_{03}(\text{Sex}_{03}) \\
&\quad + u_{01} \\
\beta_{1i} &= \gamma_{10} + \gamma_{11}(\text{Romney versus Obama supporters}_{11}) \\
&\quad + \gamma_{12}(\text{Romney supporters versus Undecided voters}_{12}) + u_{11} \\
\beta_{2i} &= \gamma_{20} + \gamma_{21}(\text{Romney versus Obama supporters}_{21}) \\
&\quad + \gamma_{22}(\text{Romney supporters versus Undecided voters}_{22})
\end{aligned}$$

Changes in diurnal slopes across the sampling period**Level 1- Within-individual variability:**

$$\text{Testosterone Diurnal Slopes}_{ti} = \beta_{0i} + \beta_{1i}(\text{Linear Day}_{1i}) + \beta_{2i}(\text{Quadratic Day}_{2i}) + e_{ti}$$

Level 2- Between-individual variability:

$$\begin{aligned}
\beta_{0i} &= \gamma_{00} + \gamma_{01}(\text{Romney versus Obama supporters}_{01}) \\
&\quad + \gamma_{02}(\text{Romney supporters versus Undecided voters}_{02}) + \gamma_{03}(\text{Sex}_{03}) \\
&\quad + u_{01} \\
\beta_{1i} &= \gamma_{10} + \gamma_{11}(\text{Romney versus Obama supporters}_{11}) \\
&\quad + \gamma_{12}(\text{Romney supporters versus Undecided voters}_{12}) + u_{11} \\
\beta_{2i} &= \gamma_{20} + \gamma_{21}(\text{Romney versus Obama supporters}_{21}) \\
&\quad + \gamma_{22}(\text{Romney supporters versus Undecided voters}_{22}) + u_{21}
\end{aligned}$$

Model used to extract slopes on each day**Level 1- Within-individual variability:**

$$\text{Testosterone}_{ti} = \beta_{0i} + \beta_{1i}(\text{Linear Time}_{1i}) + e_{ti}$$

Level 2- Between-individual variability:

$$\begin{aligned}
\beta_{0i} &= \gamma_{00} + u_{01} \\
\beta_{1i} &= \gamma_{10} + u_{11}
\end{aligned}$$

Changes in diurnal rhythms on Election Day (Supplementary analyses)**Level 1- Within-individual variability:**

$$\text{Testosterone}_{ti} = \beta_{0i} + \beta_{1i}(\text{Linear Time}_{1i}) + \beta_{2i}(\text{Quadratic Time}_{2i}) + \beta_{3i}(\text{Cubic Time}_{2i}) + e_{ti}$$

Level 2- Between-individual variability:

$$\begin{aligned}
\beta_{0i} &= \gamma_{00} + \gamma_{01}(\text{Romney versus Obama supporters}_{01}) \\
&\quad + \gamma_{02}(\text{Romney supporters versus Undecided voters}_{02}) + \gamma_{03}(\text{Sex}_{03}) \\
&\quad + u_{01} \\
\beta_{1-3i} &= \gamma_{10-30} + \gamma_{11-31}(\text{Romney versus Obama supporters}_{11-31}) \\
&\quad + \gamma_{12-32}(\text{Romney supporters versus Undecided voters}_{12-32}) + u_{11}
\end{aligned}$$

⁵ Quadratic slopes could not randomly vary in this model because the model did not have the necessary degrees of freedom due to overidentification.

Attitudes toward the elected leader moderated by testosterone-change predictors

Level 1- Within-individual variability:

$$\text{Attitudes toward Obama}_{ti} = \beta_{0i} + \beta_{1i}(\text{Linear Time}_{1i}) + \beta_{2i}(\text{Quadratic Time}_{1i}) + e_{ti}$$

Level 2- Between-individual variability:

$$\begin{aligned}\beta_{0i} &= \gamma_{00} + \gamma_{01}(\text{Romney versus Obama supporters}_{01}) \\ &\quad + \gamma_{02}(\text{Romney supporters versus Undecided voters}_{02}) \\ &\quad + \gamma_{03}(\Delta \text{Testosterone}_{03}) \\ &\quad + \gamma_{04}(\text{Romney supporters versus Undecided voters}_{04}) \\ &\quad * (\Delta \text{Testosterone}_{04}) \\ &\quad + \gamma_{05}(\text{Romney supporters versus Undecided voters}_{05}) \\ &\quad * (\Delta \text{Testosterone}_{05}) + \gamma_{06}(\text{Sex}_{06}) + u_{01} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}(\text{Romney versus Obama supporters}_{11}) \\ &\quad + \gamma_{12}(\text{Romney supporters versus Undecided voters}_{12}) \\ &\quad + \gamma_{13}(\Delta \text{Testosterone}_{13}) \\ &\quad + \gamma_{14}(\text{Romney versus Obama supporters}_{14}) * (\Delta \text{Testosterone}_{14}) \\ &\quad + \gamma_{15}(\text{Romney versus Obama supporters}_{15}) * (\Delta \text{Testosterone}_{15}) + u_{11} \\ \beta_{2i} &= \gamma_{20} + \gamma_{21}(\text{Romney versus Obama supporters}_{21}) \\ &\quad + \gamma_{22}(\text{Romney supporters versus Undecided voters}_{22}) \\ &\quad + \gamma_{23}(\Delta \text{Testosterone}_{23}) \\ &\quad + \gamma_{24}(\text{Romney versus Obama supporters}_{24}) * (\Delta \text{Testosterone}_{24}) \\ &\quad + \gamma_{25}(\text{Romney versus Obama supporters}_{25}) * (\Delta \text{Testosterone}_{25}) + u_{21}\end{aligned}$$

6.3. Statistical analysis packages in R. The lme4 package (version 1.1-18-1; Bates et al., 2015) was used to conduct all multilevel modelling. R's default polynomial function was used to create linear and quadratic contrasts for the time parameters. Effect sizes for linear regressions were estimated using the *heplots* package (Fox et al., 2018) and converted to Cohen's *d* estimates using an online converter (Lenhard & Lenhard, 2016). Given the lack of a consistent approach in estimating effect sizes for MLMs, we relied on standardized coefficients as metrics of effect sizes. These standardized coefficients for the MLMs were estimated using the *sjplot* package, which refits models on standardized data—an approach recommended for fitting complex models with interactions and polynomial terms (Gelman, 2008; Lüdtke, 2020). 95% CIs are reported for the unstandardized regression coefficients. Simple slopes were estimated using the *reghelper* package (Hughes, 2020). Robust MLMs were conducted using the *robustlmm* package (Koller, 2016). Graphs were created via the *ggplot2* package (Wickham, 2016) using estimated marginal means from models that estimated all possible time-variant contrasts (via the *emmeans* package; Lenth et al. 2019). This approach allowed us to estimate marginal means while accounting most accurately for covariates in the model.

7. Power Simulations

We generated simulations of our principal analyses (i.e., acute testosterone-change in the competition window and acute testosterone change predicting shifts in attitudes across four measurements) using the *simr* package (Green & MacLeod, 2016) in R. Our approach emulates

previously published multilevel model simulations (Maas & Hox, 2005). Although the analyses were conducted after data collection was complete, these analyses were designed in a manner consistent with an *a priori* approach. That is, we did not calculate *post hoc* power to detect the effect size evident in the data. Rather, we determined what effect size we were 80% powered to detect for each of the primary models at varying sample sizes and intraclass correlations (ICC).

For the acute testosterone change model, we simulated a multilevel model containing all the fixed and random effects necessary to examine testosterone change across three time points and three groups of political supporters (Obama, Romney, and undecided voters), controlling for sex (see model equations below). Each simulation was performed at a total sample size of 60, 100, 120, or 200 participants, within which 15% of the sample was assumed to be undecided voters and the remaining participants were split evenly between Obama and Romney supporters (e.g., for a total sample size of $n = 120$, the undecided group was simulated with 18 participants and the Obama and Romney groups were simulated with 51 participants). 50% of each simulated group was female.

The simulations were repeated with several effect sizes ($\beta = .3, .5, .7$) for the interaction of interest (Time (Linear) \times Group (Obama versus Romney)), and with several ICCs ($r = .3, .5, .7, .9$) to examine power in the context of more or less reliable testosterone assays (at the request of a reviewer). The range of effect sizes were chosen to emulate the effect of an election outcome on testosterone changes in previously published research ($\beta = .45$, Stanton et al., 2009). *Post hoc* ICC calculations on this dataset indicate relatively high correlation among all testosterone samples ($ICC = .737$).

Our simulations contained the following additional assumptions: All variables in the model are standardized; fixed intercept = 1; all fixed slopes = .3 (a medium effect size; cf. Cohen, 1988) with the exception of the fixed effect of interest (see above); residual variance = .5; the random intercept variance was set equal to the random slope variance, which was calculated based off the assumed values set for the ICC and residual variance. We generated 1000 simulations at each grouping of sample size, effect size, and ICC.

With these assumptions, the simulations indicated this study would have greater than 80% power to detect moderate effect sizes ($\beta = .5$) at total sample sizes of 100 or greater, assuming an $ICC \geq .7$ (see Fig. S1A). With lower ICCs ($r = .5$, e.g., due to unreliable testosterone assays), this study would be at least 80% powered to detect moderate effect sizes with sample sizes of 120 or greater.

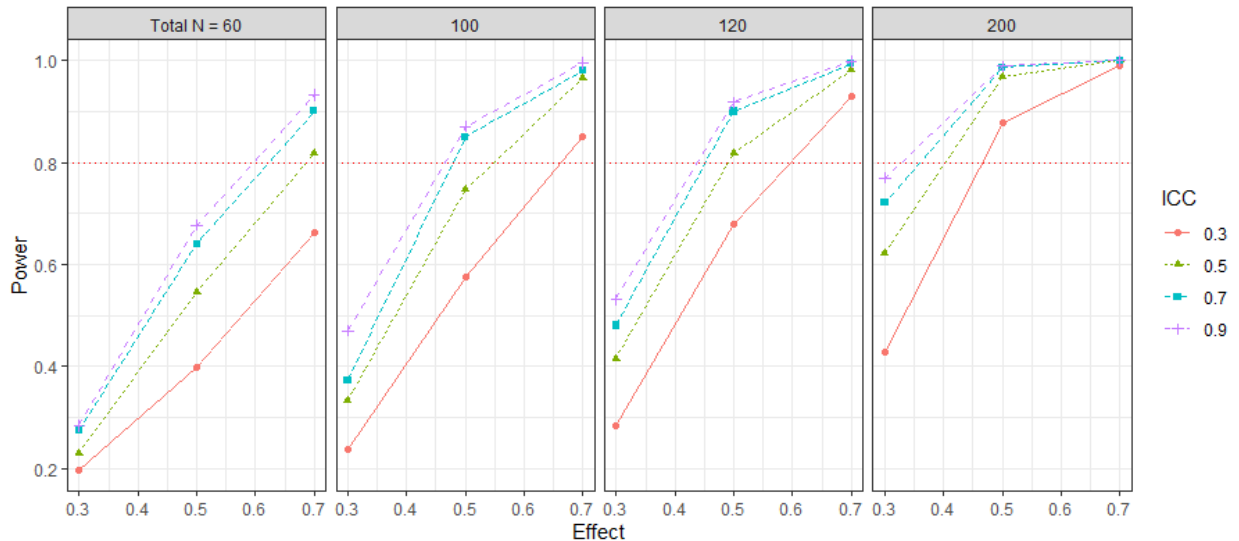
The same assumptions were used for the model of acute testosterone change predicting shifts in attitudes across four time points. For the fixed effect of interest—Time (Linear) \times Group (Obama versus Romney) \times Acute testosterone change—we again examined a set of effect sizes based on prior work. Specifically, Carré and colleagues (2009) and Mehta and Josephs (2006) each show effect sizes equivalent to approximately $\beta = .50$ and $\beta = .40$ (respectively) for the association of testosterone reactivity with behavior after losing in a laboratory competition. The set of sample sizes and effect sizes were simulated at the range of ICCs described above.

We generated a random normal distribution to represent acute testosterone change across the full sample. We created this random distribution ten times for each of the sample sizes, then generated 100 simulations to determine power for each of these random normal distributions representing acute testosterone reactivity. We examined mean power across the ten instances of 100 simulations.

Given these assumptions, the simulations suggest that the study was approximately 80% powered to detect a moderately sized Time (Linear) \times Group (Obama versus Romney) \times Acute

attitude change effect with a total sample size of 100 or greater, assuming a moderately high ICC ($ICC \geq .7$; see Fig. S1B). With a total sample size of 120, moderate effect sizes were approximately 80% powered in all but the lowest ICC examined ($r = .3$).

A. Power simulations: Group (Obama versus Romney) predicting testosterone change



B. Power simulations: Group (Obama versus Romney) \times testosterone change predicting changes in attitudes.

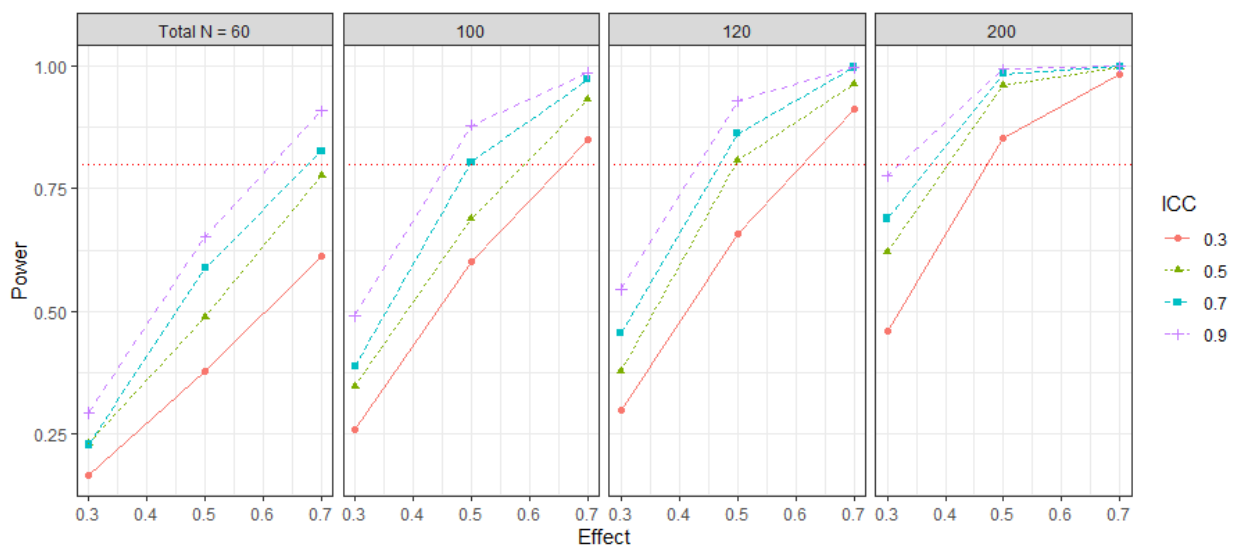


Fig. S1. Power simulation results. Each dot represents power calculated from 1000 simulations for a given effect size (horizontal axis), degree of intraclass correlation (ICC; dot and line color and type), and total sample size (facets of graph). Dotted red horizontal line indicates 80% power. See supplemental text for more details.

Table S6. Key strengths of the study.

Strengths of the present study ¹
<p>Attitude shifts as a novel outcome related to testosterone change: The study tests the extent to which testosterone responses to competition are related to shifts in attitudes toward an elected leader. In doing so, the study provides an important bridge between research on biological responses to competition and research on the psychology of attitude change.</p>
<p>Ecological validity: Previous studies that have examined associations between testosterone responses to competition and psychological outcome measures have focused primarily on laboratory models, which have limited ecological validity (e.g. Carré et al., 2009; Mehta & Josephs, 2006). The present study extends this work to a naturalistic setting of significant societal importance.</p>
<p>Temporal changes beyond minutes:</p> <ul style="list-style-type: none"> • <i>Psychological changes:</i> Previous work has focused primarily on testosterone-related psychological changes within minutes after a competition (Carré et al., 2009; Mehta & Josephs, 2006). The present study expands this temporal window by measuring attitudes at two instances before the election, as well as the day after, several days after, and 6 months after the election. This design permits characterization of the time course of testosterone-moderated shifts in attitudes. • <i>Acute hormone response and diurnal rhythms:</i> Previous research has largely assumed that fluctuations in testosterone levels occur within minutes after a competition and are short-lived. But the societal importance of democratic elections suggests that their outcomes may produce biological changes that last hours or even days. To test this hypothesis, the present study measured acute testosterone responses to the election outcome over several hours as well as shifts in testosterone's daily diurnal rhythm over a five-day period.
<p>Statistical power: Our final sample was within the range of the targeted sample size obtained from a conservative power analysis. Specifically, simulations indicated that the study was approximately 80% powered to detect a moderate effect size assuming a total sample size of $n = 100$ or greater and assuming moderately high ICCs ($r \geq .7$). At lower ICCs ($r = .5$; e.g., due to less reliable testosterone assays), these simulations indicate that a total sample size of 120 or greater was necessary to achieve approximately 80% power for a moderate effect size.</p>
<p>Validity of hormone measurement: Most studies in the field have measured testosterone levels using immunoassays but have not directly evaluated the measurement validity of immunoassays. The present study examined correspondence between testosterone measured with immunoassays and testosterone measured with a highly accurate reference method (mass spectrometry) in a random selection of samples. EIAs conducted on a random subset of our samples demonstrated good accuracy in estimating testosterone concentrations. However, performance of EIAs was better in samples with higher concentrations of testosterone compared to samples with lower concentrations of testosterone.</p>
<p>Multi-faceted measurement of attitudes: The present study measured shifts in attitudes toward the elected leader using a multi-method approach: (1) surveys similar to presidential approval scales from polling organizations, such as Gallup, served as our primary attitude measure; and (2) social media activity (via Facebook) was coded for attitudes in a subset of participants with usable data.</p>
<p>Sample composition:</p> <ul style="list-style-type: none"> • <i>Sex/gender:</i> Research on testosterone responses to competition has disproportionately focused on males. The present study included both males and females and explored sex/gender similarities and differences. • <i>Undecided voters:</i> Previous research primarily examined differences between winners and losers. With this design it remains unclear whether any differences are driven by winners, losers, or both. The present study included undecided voters as an exploratory reference group to facilitate interpretation of differences between supporters of the winning and losing presidential candidates.
<p>Open dataset and publicly available code: The present study's unique dataset and code necessary for analyses included in this report are publicly available on the Open Science Framework.</p>

MLMs as the tool for statistical analysis: The present study adopted multi-level models that have advantages over traditional approaches (e.g., ANOVA), including modelling fixed and random effects, avoiding listwise deletion of data thereby boosting statistical power, modelling linear and non-linear change to evaluate the extent to which effects persist or weaken over time, and allowing greater flexibility in the modelling approach (e.g., not requiring assumptions of sphericity, or homogeneity of regression slopes).

Use of compliance protocols for the at-home portion of the study: Electronic time stampers and automated text messages were used to ensure fidelity of naturalistic data.

Robustness checks: We verified the robustness of our findings via: (i) False Discovery Rate (FDR) corrections for multiple comparisons; (ii) inclusion of the following covariates across our main models: a) age and SES, b) changes in affect during the competition-outcome window, c) when participants learned about the election outcome, and d) when participants completed the Pre- and Post-Election Lab sessions; and (iii) specification curve analyses across alternative data analytic approaches, such as the type of model (MLM versus robust MLM versus GLM), whether slopes and intercepts were allowed to randomly vary, treatment of outliers (inclusion, exclusion, or winsorization), and the type of model used to estimate diurnal slopes (MLM versus GLM).

1. Note: Details about these strengths can be found in the primary methods section and supplementary material.

SI RESULTS

1. Determination of the competition-outcome window

In the 2012 US presidential election, several media sources and social media platforms unofficially declared the electoral results around 8:12 pm PST (Ariens, 2016). In our sample, 95.58% of our participants reported following the election between 7 pm and *Bedtime*, and on average reported finding out about the results at 8:50 pm. As shown in Fig. S2, we also found robust shifts in affective responses to the election on Election Day, across political groups (see Table S7 and Table S8). Visual examination of changes in positive and negative affect on Election Day highlights that these shifts in affect were especially salient only after 7 pm. Thus, we concluded that: (i) our participants were tracking and following the election, (ii) on average they reported finding out about the outcome at the same time as it was declared in the media (i.e., after 8 pm), and (iii) they demonstrated robust shifts in affect in response to the electoral outcome after it was declared (see Fig. S2). Based on these analyses, 7 pm to *Bedtime* was treated as the theorized competition-outcome window to measure endocrine shifts in response to the election.

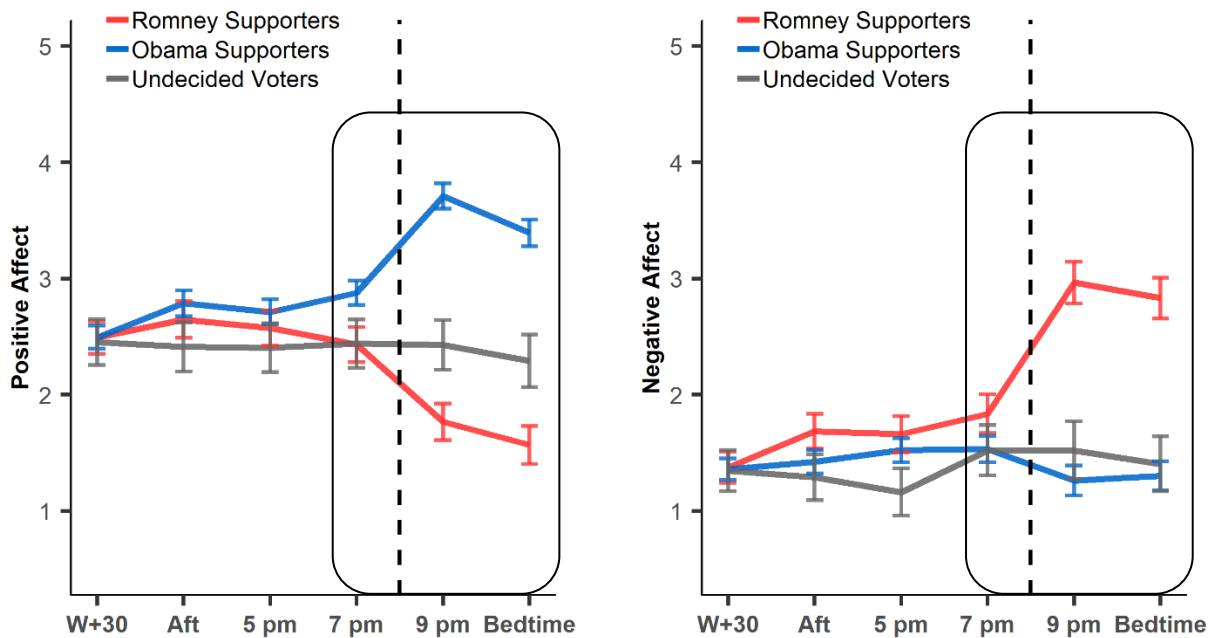


Fig. S2. Positive affect (composite of happy, excited, powerful, and confident) and negative affect (sadness) across Election Day from 30 minutes after participants woke up (*W+30*), in the afternoon (*Aft*), at 5 pm, 7 pm, 9 pm, and when they went to bed (*Bedtime*). Solid lines represent estimated marginal means. Error bars = \pm SEs. The vertical dashed line represents the time at which the electoral results were declared. The box represents the competition outcome window.

Table S7. Means and *SDs* of positive affect (composite of happy, excited, powerful, and confident) and negative affect (sadness) reported on Election Day (measured on a 5-point Likert scales).

	POSITIVE AFFECT						NEGATIVE AFFECT					
	<i>Romney Supporters</i>		<i>Obama Supporters</i>		<i>Undecided voters</i>		<i>Romney Supporters</i>		<i>Obama Supporters</i>		<i>Undecided voters</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Wakeup+30	2.49	0.79	2.50	0.73	2.47	0.91	1.38	0.82	1.36	0.63	1.35	0.70
Afternoon	2.62	0.89	2.79	0.83	2.41	0.80	1.70	1.03	1.43	0.67	1.25	0.45
5 pm	2.56	0.88	2.71	0.82	2.42	0.73	1.67	0.96	1.52	0.70	1.12	0.34
7 pm	2.43	0.92	2.88	0.76	2.45	0.86	1.81	1.04	1.53	1.03	1.53	0.80
9 pm	1.76	1.00	3.70	0.88	2.44	1.10	2.97	1.52	1.27	0.61	1.50	0.89
Bedtime	1.57	0.75	3.38	0.88	2.30	0.91	2.83	1.53	1.31	0.56	1.41	0.71

Table S8. MLMs examining changes in positive affect (composite of happy, excited, powerful, and confident) and negative affect (sadness) across Election Day and across political groups.

	POSITIVE AFFECT			NEGATIVE AFFECT		
	<i>B</i>	<i>95 % CI</i>	<i>p</i>	<i>B</i>	<i>95 % CI</i>	<i>p</i>
(Intercept)	2.25	1.99 – 2.50	<0.001	2.10	1.86 – 2.34	<0.001
Time- Linear	-0.88	-1.19 – -0.57	<0.001	1.35	1.01 – 1.68	<0.001
Time- Quadratic	-0.45	-0.67 – -0.23	<0.001	0.23	-0.01 – 0.47	0.062
Romney supporters versus Obama supporters	0.75	0.48 – 1.02	<0.001	-0.68	-0.94 – -0.42	<0.001
Romney supporters versus Undecideds	0.16	-0.22 – 0.54	0.417	-0.71	-1.07 – -0.35	<0.001
Time- Linear × Romney versus Obama	1.76	1.39 – 2.14	<0.001	-1.43	-1.84 – -1.03	<0.001
Time- Quadratic × Romney versus Obama	0.52	0.25 – 0.79	<0.001	-0.4	-0.70 – -0.11	0.007
Time- Linear × Romney versus Undecideds	0.79	0.26 – 1.32	0.004	-1.18	-1.74 – -0.63	<0.001
Time- Quadratic × Romney versus Undecideds	0.40	0.02 – 0.78	0.041	-0.21	-0.61 – 0.18	0.298
Sex (0= Female, 1= Male)	-0.00	-0.24 – 0.24	0.998	-0.03	-0.25 – 0.19	0.796
Random Effects						
σ^2	0.30			0.41		
τ_{00}	0.34 pid			0.28 pid		
τ_{11}	0.46 pid.timepoly2.L			0.45 pid.timepoly2.L		
	0.09 pid.timepoly2.Q			0.02 pid.timepoly2.Q		
ρ_{01}	0.00			0.48		
	-0.38			-0.40		
N	111 pid			110 pid		
Observations	646			632		

Table S9. Means and SDs of testosterone levels (in pg/mL) across the competition outcome window.

	<i>Romney Supporters</i>		<i>Obama Supporters</i>		<i>Undecided voters</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
7 pm	78.72	46.26	78.72	53.17	74.12	43.04
9 pm	72.75	47.63	66.73	41.00	60.20	27.26
Bedtime	85.41	57.68	64.57	36.76	61.30	23.96

Table S10. MLMs examining changes in testosterone levels across the competition-outcome window (including 7 pm⁶, 9 pm, and Bedtime) and across political groups. Model B includes age and socioeconomic status as covariates.**ACUTE TESTOSTERONE REACTIVITY DURING THE COMPETITION-OUTCOME WINDOW**

	MODEL A			MODEL B		
	<i>B</i>	<i>95% CI</i>	<i>p</i>	<i>B</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	49.22	38.36 – 60.07	<0.001	56.53	29.60 – 83.45	<0.001
Time- Linear	3.93	-5.39 – 13.24	0.41	4.02	-5.31 – 13.34	0.401
Time- Quadratic	3.78	-2.59 – 10.15	0.247	3.81	-2.57 – 10.19	0.244
Romney supporters versus Obama supporters	-5.31	-17.06 – 6.44	0.377	-5.54	-17.40 – 6.32	0.362
Romney supporters versus Undecideds	1.69	-14.51 – 17.88	0.839	2.19	-14.79 – 19.18	0.801
Time- Linear × Romney versus Obama	-14.60	-26.00 – -3.21	0.013	-14.69	-26.10 – -3.28	0.013
Time- Quadratic × Romney versus Obama	0.62	-7.12 – 8.37	0.875	0.59	-7.15 – 8.34	0.881
Time- Linear × Romney versus Undecideds	-12.61	-28.20 – 2.97	0.116	-12.66	-28.27 – 2.94	0.115
Time- Quadratic × Romney versus Undecideds	1.35	-9.13 – 11.83	0.801	1.35	-9.14 – 11.83	0.802
Sex (0= Female, 1= Male)	60.67	50.38 – 70.97	<0.001	60.29	49.81 – 70.77	<0.001
Age				-0.20	-0.83 – 0.42	0.524
Socioeconomic Status				-0.37	-3.14 – 2.39	0.791
Random Effects						
σ^2	310.68			310.9		
τ_{00}	652.28 _{pid}			657.95 _{pid}		
τ_{11}	379.82 _{pid.timepoly11}			381.23 _{pid.timepoly11}		
ρ_{01}	-0.29 _{pid}			-0.27 _{pid}		
N	113 _{pid}			113 _{pid}		
Observations	329			329		

⁶ Supporters of the two main candidates (i.e., Obama and Romney) did not differ in testosterone concentrations at 7 pm ($B=6.10$, $95\%CI[-10.20, 22.41]$, $p=.487$).

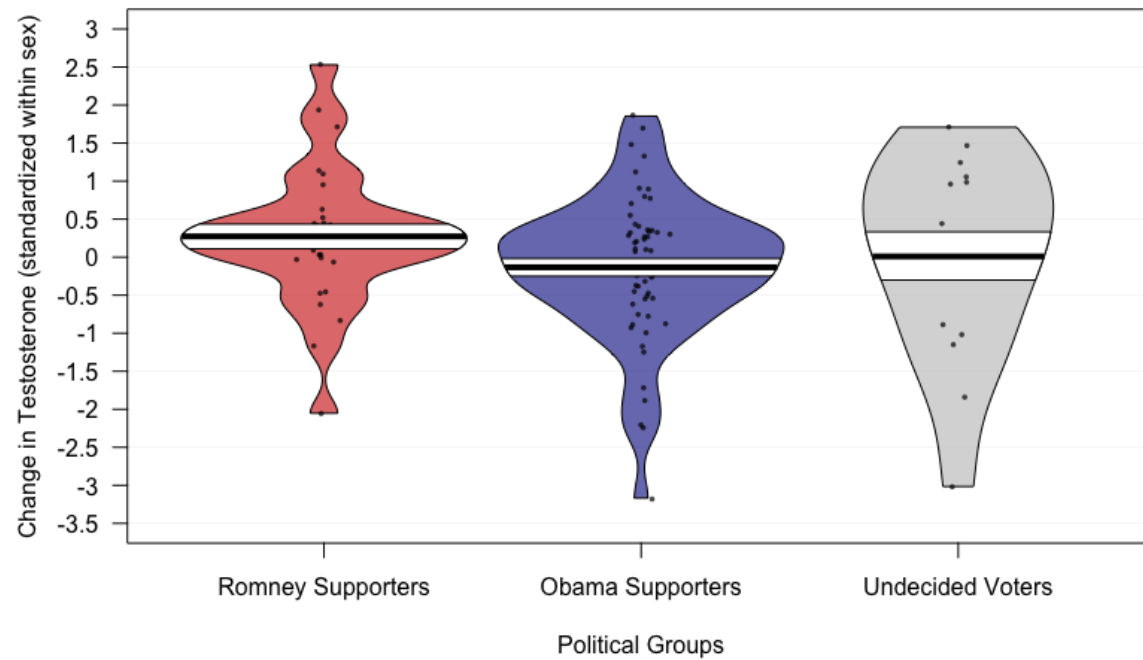


Figure S3. Pirate plot of acute changes in testosterone across the competition-outcome window (standardized within sex) predicted by political groups. Dots represent raw testosterone change scores (standardized within sex). Beans represent the smoothed density curve showing the full data distribution of change scores. Bands represent SEs.

Table S11. MLMs examining short-term shifts in attitudes toward the elected leader (Obama) moderated by acute testosterone reactivity during the competition-outcome window, across political groups. Model B includes age and socioeconomic status as covariates.

ACUTE TESTOSTERONE REACTIVITY PREDICTING SHORT-TERM SHIFTS IN SELF-REPORTED ATTITUDES TOWARD THE ELECTED LEADER

	MODEL A			MODEL B		
	<i>B</i>	<i>95% CI</i>	<i>p</i>	<i>B</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	2.29	2.03 – 2.54	<0.001	2.32	1.71 – 2.92	<0.001
Day- Linear	0.07	-0.02 – 0.16	0.152	0.07	-0.03 – 0.16	0.153
Day- Quadratic	0.09	-0.00 – 0.19	0.065	0.09	-0.00 – 0.19	0.065
Romney supporters versus Obama supporters	1.46	1.19 – 1.73	<0.001	1.48	1.22 – 1.74	<0.001
Romney supporters versus Undecideds	0.78	0.41 – 1.15	<0.001	0.97	0.60 – 1.35	<0.001
Acute Change in Testosterone (Δ Testosterone)	-0.17	-0.41 – 0.07	0.171	-0.18	-0.42 – 0.06	0.138
Day- Linear \times Romney versus Obama	0.00	-0.12 – 0.11	0.955	0.00	-0.11 – 0.11	0.964
Day- Quadratic \times Romney versus Obama	-0.12	-0.24 – -0.01	0.042	-0.12	-0.24 – -0.01	0.041
Day- Linear \times Romney versus Undecideds	-0.07	-0.22 – 0.09	0.398	-0.07	-0.22 – 0.09	0.404
Day- Quadratic \times Romney versus Undecideds	0.00	-0.16 – 0.16	0.996	0.00	-0.16 – 0.16	0.997
Day- Linear \times Δ Testosterone	-0.27	-0.37 – -0.17	<0.001	-0.27	-0.37 – -0.17	<0.001
Day- Quadratic \times Δ Testosterone	-0.08	-0.18 – 0.02	0.113	-0.08	-0.18 – 0.02	0.112
Δ Testosterone \times Romney versus Obama	0.08	-0.22 – 0.37	0.613	0.09	-0.20 – 0.37	0.551
Δ Testosterone \times Romney versus Undecideds	0.29	-0.04 – 0.63	0.092	0.32	-0.01 – 0.64	0.058
Day- Linear \times Δ Testosterone \times Romney versus Obama	0.29	0.17 – 0.41	<0.001	0.29	0.17 – 0.41	<0.001
Day- Quadratic \times Δ Testosterone \times Romney versus Obama	0.10	-0.02 – 0.22	0.12	0.10	-0.02 – 0.22	0.12
Day- Linear \times Δ Testosterone \times Romney versus Undecideds	0.26	0.12 – 0.40	<0.001	0.26	0.12 – 0.40	<0.001
Day- Quadratic \times Δ Testosterone \times Romney versus Undecideds	0.13	-0.01 – 0.27	0.072	0.13	-0.01 – 0.27	0.072
Sex (0= Female, 1= Male)	-0.20	-0.43 – 0.03	0.099	-0.19	-0.42 – 0.04	0.102
Age				-0.02	-0.03 – -0.00	0.031
Socioeconomic Status				0.06	-0.00 – 0.12	0.066
Random Effects						
σ^2	0.06			0.06		
τ_{00}	0.33 _{pid}			0.31 _{pid}		
τ_{11}	0.00 _{pid.daypoly2.L}			0.00 _{pid.daypoly2.L}		
	0.00 _{pid.daypoly2.Q}			0.00 _{pid.daypoly2.Q}		
ρ_{01}	0.77			0.94		
	-0.53			-0.71		
N	106 _{pid}			106 _{pid}		
Observations	418			418		

2. Linear regression analysis of acute testosterone reactivity predicting shifts in attitudes toward the elected leader.

We conducted a simplified linear regression analysis to bolster the findings reported in the main text that described shifts in attitudes toward the elected leader (Obama) from before to after the election (the *Pre-Election Lab* session, *ED-1*, *ED+1*, and the *Post-Election Lab* session; see Fig. 3). We examined the extent to which acute testosterone reactivity predicted changes in attitudes toward the elected leader (Obama) using a difference score of attitudes reported before the election (average of attitudes from *Pre-Election Lab* and *ED-1*) to after the election (average of attitudes from *ED+1* and *Post-Election Lab*). This analysis revealed a significant Acute Testosterone Reactivity \times Group (Romney versus Obama) interaction ($B=0.25$, 95%CI[0.13, 0.37], $\beta=.91$, $p<.001$, $d=0.98$; see Fig. S4 for scatterplot), supporting the pattern of results described in the main text (Fig. 3). Simple slopes analysis indicated that supporters of the losing candidate (Romney) who demonstrated greater acute testosterone reactivity on the night of the election reported less positive attitudes toward the elected leader from before to after the election ($B=-.23$, $SE=.05$, $t_{(99)}=-4.54$, $p<.001$; $r=-.60$). We did not observe any relationship between acute testosterone reactivity and shifts in attitudes amongst supporters of the winner ($p=.547$).

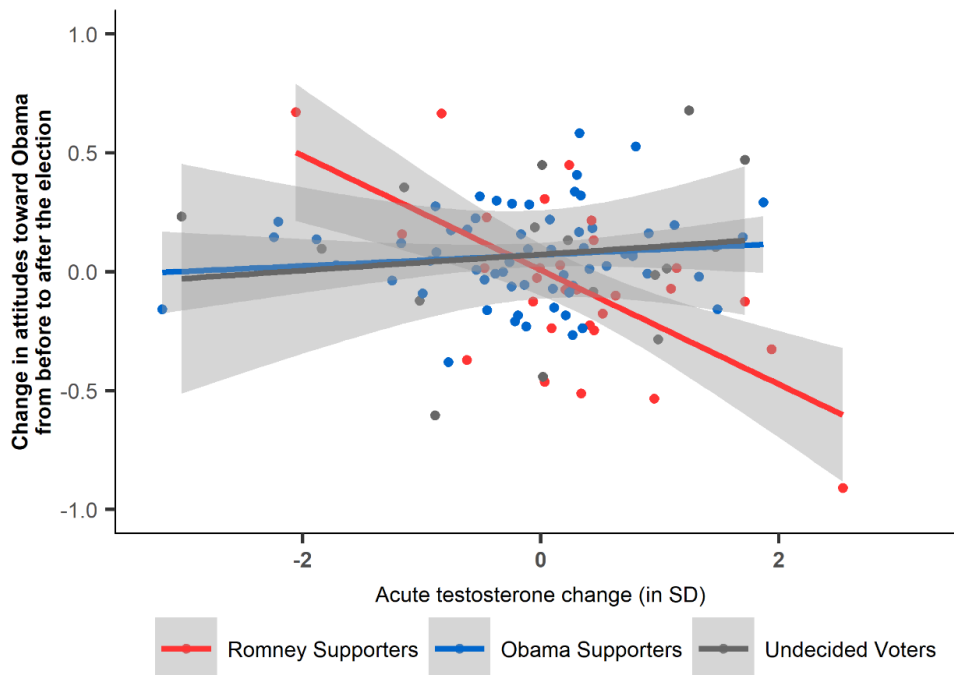


Fig. S4. Scatter plot between acute testosterone reactivity (standardized within sex) during the competition-outcome window and changes in attitudes toward the elected leader (Obama) from before the election (average of attitudes from the *Pre-Election Lab* and *ED-1*) to after the election (average of attitudes from *ED+1* and *Post-Election Lab*), across political groups. Shaded regions indicate *SEs*. *SD*= Standard deviation.

3. Acute testosterone reactivity predicting long-term shifts in attitudes toward the elected leader.

A subset of our sample ($n=66$) responded to the longitudinal survey six months after the election. In this sample, we tested whether acute testosterone reactivity predicted longer-term shifts in attitudes from before the election (*Pre-Election Lab* and *ED-1*) to six months after it. We included only the linear time parameter in this model because we were assessing shifts in attitudes from pre-election attitude measures to only one post-election attitude measure. The MLM revealed a weak Time (Linear) \times Acute Testosterone Reactivity \times Group (Romney versus Obama) interaction ($B=0.17$, 95% $CI[-0.04, 0.37]$, $\beta=.19$, $p=.124$; see Table S12; Fig. 3). This analysis suggests that testosterone-moderated shifts in attitudes six months after the election were in the same direction but not as robust as short-term shifts in attitudes.

We further confirmed that the attenuated association between acute testosterone reactivity and long-term shifts in attitudes was not explained solely by the loss of statistical power due to attrition. Indeed, consistent with our findings in the full sample, participants who completed the six-month follow-up still robustly showed the Time (Linear) \times Acute Testosterone Reactivity \times Group interaction predicting short-term shifts in attitudes (Romney versus Obama: $B=0.36$, 95% $CI[0.23, 0.49]$, $\beta=.40$, $p<.001$). Collectively, this pattern suggests that acute testosterone responses on election night robustly predicted shifts in attitudes toward the elected leader in the days immediately after the election, and the long-term effect of acute testosterone responses on attitudinal changes was likely in the same direction but smaller in magnitude.

Table S12. MLMs examining long-term shifts in attitudes toward the elected leader (Obama) moderated by acute testosterone reactivity during the competition-outcome window, across political groups, in a subset of participants who consented to a six-month longitudinal follow-up ($n=66$). Model B includes age and socioeconomic status as covariates.

ACUTE TESTOSTERONE REACTIVITY PREDICTING LONG-TERM SHIFTS IN SELF-REPORTED ATTITUDES TOWARD THE ELECTED LEADER						
	MODEL A			MODEL B		
	<i>B</i>	95% <i>CI</i>	<i>p</i>	<i>B</i>	95% <i>CI</i>	<i>p</i>
(Intercept)	2.22	1.87 – 2.56	<0.001	1.81	0.88 – 2.74	<0.001
Day- Linear	-0.18	-0.34 – -0.02	0.03	-0.18	-0.34 – -0.02	0.03
Romney supporters versus Obama supporters	1.29	0.92 – 1.67	<0.001	1.32	0.94 – 1.69	<0.001
Romney supporters versus Undecideds	0.78	0.32 – 1.24	0.002	0.88	0.37 – 1.38	0.001
Acute Change in Testosterone (Δ Testosterone)	-0.07	-0.37 – 0.22	0.625	-0.12	-0.43 – 0.18	0.429
Day- Linear \times Romney versus Obama	-0.16	-0.36 – 0.05	0.143	-0.16	-0.36 – 0.05	0.143
Day- Linear \times Romney versus Undecideds	-0.11	-0.36 – 0.14	0.391	-0.11	-0.36 – 0.14	0.391
Day- Linear \times Δ Testosterone	-0.12	-0.28 – 0.04	0.132	-0.12	-0.28 – 0.04	0.132
Δ Testosterone \times Romney versus Obama	-0.02	-0.41 – 0.37	0.935	0.01	-0.38 – 0.41	0.945
Δ Testosterone \times Romney versus Undecideds	0.09	-0.31 – 0.50	0.651	0.13	-0.28 – 0.54	0.535
Day- Linear \times Δ Testosterone \times Romney versus Obama	0.17	-0.04 – 0.37	0.124	0.17	-0.04 – 0.37	0.123
Day- Linear \times Δ Testosterone \times Romney versus Undecideds	0.08	-0.14 – 0.30	0.483	0.08	-0.14 – 0.30	0.483
Sex (0= Female, 1= Male)	-0.23	-0.57 – 0.11	0.185	-0.2	-0.54 – 0.14	0.257
Age				0.00	-0.02 – 0.02	0.973
Socioeconomic Status				0.06	-0.03 – 0.16	0.176
Random Effects						
σ^2	0.1			0.1		
τ_{00}	0.34 _{pid}			0.34 _{pid}		
τ_{11}	0.02 _{pid.daypoly11}			0.02 _{pid.daypoly11}		
ρ_{01}	-0.11 _{pid}			-0.07 _{pid}		
N	60 _{pid}			60 _{pid}		
Observations	178			178		

Table S13. Means and *SDs* of diurnal slope values (lower values indicate steeper slopes, whereas higher values indicate flatter slopes) across the sampling period.

	<i>Romney Supporters</i>		<i>Obama Supporters</i>		<i>Undecided voters</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ED - 2 days	-35.14	27.10	-30.51	17.26	-26.59	18.20
ED - 1 day	-32.96	18.04	-34.51	19.23	-28.69	14.91
ED	-26.32	17.60	-32.55	18.44	-24.39	15.32
ED + 1 day	-27.62	17.80	-31.14	20.54	-26.07	20.21
ED + 2 days	-25.75	16.78	-31.67	20.36	-24.91	18.62

Table S14. MLMs examining changes in diurnal slopes across the naturalistic sampling period and across political groups. Model B includes age and socioeconomic status as covariates.

DIURNAL SLOPE CHANGE ACROSS ALL FIVE DAYS OF THE SAMPLING PERIOD USING EXTRACTED SLOPE VALUES

	MODEL A			MODEL B		
	<i>B</i>	<i>95% CI</i>	<i>p</i>	<i>B</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	-14.57	-18.31 – -10.83	<0.001	-12.48	-21.57 – -3.39	0.008
Time- Linear	7.63	3.69 – 11.57	<0.001	7.63	3.69 – 11.56	<0.001
Time- Quadratic	-2.29	-5.83 – 1.26	0.208	-2.29	-5.83 – 1.25	0.208
Romney supporters versus Obama supporters	-4.84	-8.89 – -0.78	0.021	-5.08	-9.15 – -1.02	0.016
Romney supporters versus Undecideds	-5.00	-10.61 – 0.60	0.083	-6.24	-12.07 – -0.42	0.038
Time- Linear × Romney versus Obama	-7.30	-12.17 – -2.43	0.004	-7.30	-12.17 – -2.43	0.004
Time- Quadratic × Romney versus Obama	4.00	-0.39 – 8.38	0.076	4.00	-0.39 – 8.38	0.076
Time- Linear × Romney versus Undecideds	-5.44	-12.11 – 1.23	0.112	-5.44	-12.10 – 1.22	0.112
Time- Quadratic × Romney versus Undecideds	2.68	-3.31 – 8.68	0.382	2.68	-3.31 – 8.67	0.382
Sex (0= Female, 1= Male)	-29.09	-32.61 – -25.56	<0.001	-29.23	-32.79 – -25.68	<0.001
Age				0.07	-0.14 – 0.28	0.532
Socioeconomic Status				-0.60	-1.53 – 0.34	0.216
Random Effects						
σ^2	76.8			76.99		
τ_{00}	76.49 _{pid}			75.61 _{pid}		
τ_{11}	56.37 _{pid.daypoly2.L}			55.98 _{pid.daypoly2.L}		
	31.03 _{pid.daypoly2.Q}			30.70 _{pid.daypoly2.Q}		
ρ_{01}	0.36			0.34		
	0.13			0.13		
N	113 _{pid}			113 _{pid}		
Observations	564			564		

4. Follow-up analyses of changes in diurnal slopes across the sampling period.

In the main text, we reported changes in diurnal slopes across the sampling period among supporters of the losing candidate (Time (Linear): $B=-7.30$, 95% $CI[-12.17, -2.43]$, $p=.004$). To further decompose the trajectory of diurnal slope change in the supporters of the losing candidate, we conducted t -tests to determine whether slopes on the day before the election ($ED-1$; i.e., the reference measurement instance) differed from all other days in the naturalistic sampling period. Our analyses revealed no significant difference in the diurnal slopes between $ED-1$ and $ED-2$ ($p=.507$), but significantly flatter slopes on ED ($t_{(32)}=2.40$, $p=.022$), $ED+1$ ($t_{(32)}=2.63$, $p=.013$), and $ED+2$ ($t_{(32)}=3.03$, $p=.005$; see Table S13 for mean diurnal slope values at each day of the sampling period). These analyses support our interpretation that testosterone slopes were flatter on ED relative to the days before the election, and these flatter slopes persisted two days after the election.

5. Changes in diurnal slopes on Election Day.

We also examined changes in testosterone levels across all seven samples on Election Day (*Wake-up*, *Wake-up + 30*, *Afternoon*, *5 pm*, *7 pm*, *9 pm*, and *Bedtime*; see Table S15 for Means and *SDs*). In this model, in addition to the linear and quadratic terms to estimate the trajectory of change in testosterone levels across Election Day, a cubic term was introduced to detect disruptions in diurnal rhythms at the end of the day, when the electoral results were announced. The analysis revealed a Time \times Group (Romney versus Obama) interaction (Cubic: $B=-12.35$, 95% $CI[-22.68, -2.02]$, $\beta=-.19$, $p=.021$, Linear: $B=-20.82$, 95% $CI[-42.02, 0.38]$, $\beta=-.32$, $p=.057$; see Table S16 for full model). To interpret this interaction, we plotted testosterone levels across time in each group (Fig. S5). The marginally significant linear interaction term supported the overall findings that supporters of the candidate who lost the election demonstrated flatter diurnal slopes compared to supporters of the winning candidate. Further, the presence of a significant cubic interaction term tracked the reverse winner-loser effect pattern (Zilioli et al., 2014): Supporters of the losing candidate had higher testosterone levels on the night of the election compared to supporters of the winning candidate.

Table S15. Means and *SDs* of testosterone levels (in pg/mL) across Election Day.

	<i>Romney Supporters</i>		<i>Obama Supporters</i>		<i>Undecided voters</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Wake-up	152.65	85.52	165.11	86.45	125.94	65.92
Wake-up+30	129.07	66.55	131.79	70.18	123.68	61.01
Afternoon	101.78	67.51	91.00	52.14	85.95	42.23
5 pm	99.89	60.82	87.59	48.90	76.64	38.24
7 pm	78.72	46.26	78.72	53.17	74.12	43.04
9 pm	75.43	59.01	66.73	41.00	60.69	28.14
Bedtime	85.41	57.68	65.10	37.56	61.30	23.96

Table S16. MLMs examining changes in testosterone levels across Election Day as a function of political groups. Model B includes age and socioeconomic status as covariates.

CHANGE IN TESTOSTERONE LEVELS ACROSS ELECTION DAY						
	MODEL A			MODEL B		
	<i>B</i>	<i>95% CI</i>	<i>p</i>	<i>B</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	67.67	56.37 – 78.96	<0.001	71.84	45.43 – 98.25	<0.001
Time- Linear	-63.02	-80.18 – -45.86	<0.001	-63.02	-80.18 – -45.86	<0.001
Time- Quadratic	27.98	16.38 – 39.59	<0.001	28.00	16.39 – 39.61	<0.001
Time- Cubic	2.84	-5.57 – 11.25	0.51	2.86	-5.55 – 11.27	0.507
Romney supporters versus Obama supporters	-0.16	-12.63 – 12.30	0.979	-0.22	-12.79 – 12.36	0.973
Romney supporters versus Undecideds	2.98	-14.18 – 20.14	0.735	4.11	-13.77 – 21.98	0.653
Time- Linear × Romney versus Obama	-20.82	-42.02 – 0.38	0.057	-20.83	-42.02 – 0.37	0.057
Time- Quadratic × Romney versus Obama	3.42	-10.88 – 17.73	0.64	3.41	-10.90 – 17.71	0.641
Time- Cubic × Romney versus Obama	-12.35	-22.68 – -2.02	0.021	-12.38	-22.71 – -2.04	0.021
Time- Linear × Romney versus Undecideds	-0.79	-29.62 – 28.03	0.957	-0.77	-29.59 – 28.05	0.958
Time- Quadratic × Romney versus Undecideds	-10.89	-30.41 – 8.62	0.276	-10.89	-30.41 – 8.63	0.277
Time- Cubic × Romney versus Undecideds	1.94	-12.09 – 15.96	0.787	1.93	-12.10 – 15.96	0.788
Sex (0= Female, 1= Male)	69.96	59.95 – 79.97	<0.001	69.99	59.80 – 80.17	<0.001
Age				-0.19	-0.80 – 0.41	0.531
Socioeconomic Status				0.07	-2.61 – 2.76	0.957
Random Effects						
σ^2	436.84			437		
τ_{00}	801.04 _{pid}			806.79 _{pid}		
τ_{11}	2061.86 _{pid.timepoly3.L}			2060.53 _{pid.timepoly3.L}		
	690.12 _{pid.timepoly3.Q}			689.79 _{pid.timepoly3.Q}		
	141.05 _{pid.timepoly3.C}			141.15 _{pid.timepoly3.C}		
ρ_{01}	-0.44			-0.44		
	0.00			0.00		
	-0.25			-0.25		
N	113 _{pid}			113 _{pid}		
Observations	775			775		

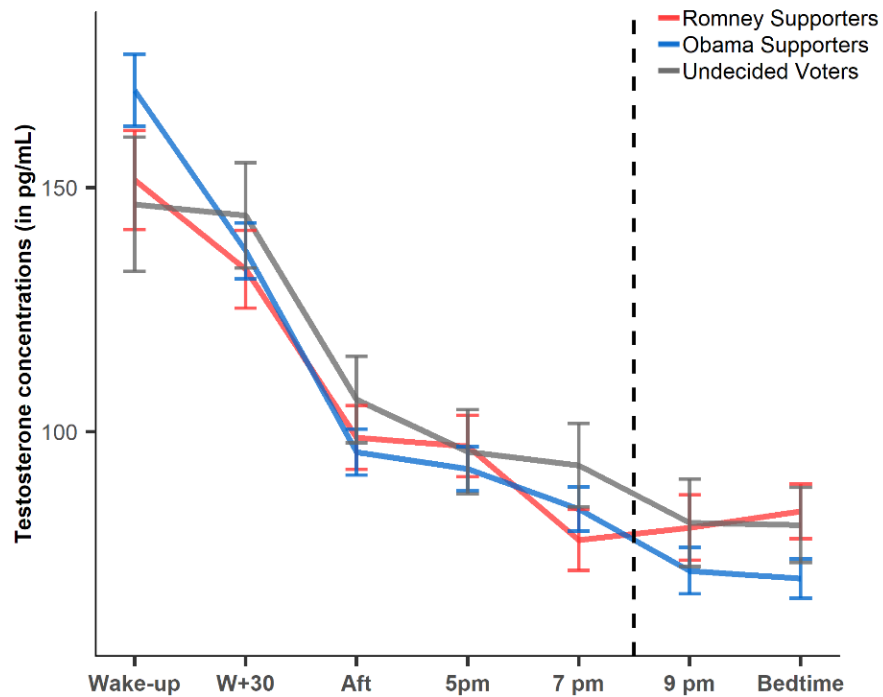


Fig. S5. Testosterone levels (in pg/mL) across Election Day from the time participants woke up (*Wake-up*), 30 minutes after they woke up (*W+30*), in the afternoon (*Aft*), at 5 *pm*, 7 *pm*, 9 *pm*, and when they went to bed (*Bedtime*). Solid lines represent estimated marginal means (controlling for sex). Error bars are SEs. The vertical dotted line represents the time at which the electoral results were declared.

Table S17. MLMs examining short-term shifts in attitudes toward the elected leader (Obama) moderated by diurnal slope change from *ED-I* to *ED*, across political groups. Model B includes the covariates of age and socioeconomic status.

	MODEL A			MODEL B		
	<i>B</i>	95% <i>CI</i>	<i>p</i>	<i>B</i>	95% <i>CI</i>	<i>p</i>
(Intercept)	2.2	1.96 – 2.44	<0.001	2.16	1.56 – 2.75	<0.001
Day- Linear	0	-0.09 – 0.09	0.98	0	-0.09 – 0.09	0.981
Day- Quadratic	0.06	-0.02 – 0.15	0.152	0.06	-0.02 – 0.15	0.153
Romney supporters versus Obama supporters	1.55	1.29 – 1.82	<0.001	1.57	1.30 – 1.83	<0.001
Romney supporters versus Undecideds	0.75	0.38 – 1.12	<0.001	0.87	0.48 – 1.25	<0.001
Change in Diurnal slopes (Δ Diurnal Slope)	-0.07	-0.26 – 0.12	0.47	-0.06	-0.25 – 0.13	0.547
Day- Linear \times Romney versus Obama	0.06	-0.05 – 0.17	0.28	0.06	-0.05 – 0.17	0.278
Day- Quadratic \times Romney versus Obama	-0.1	-0.20 – 0.01	0.086	-0.1	-0.20 – 0.01	0.086
Day- Linear \times Romney versus Undecideds	-0.02	-0.18 – 0.14	0.822	-0.02	-0.18 – 0.14	0.823
Day- Quadratic \times Romney versus Undecideds	0.05	-0.10 – 0.21	0.519	0.05	-0.10 – 0.21	0.515
Day- Linear \times Δ Diurnal Slope	-0.12	-0.20 – -0.04	0.003	-0.12	-0.20 – -0.04	0.003
Day- Quadratic \times Δ Diurnal Slope	0.07	-0.01 – 0.15	0.083	0.07	-0.01 – 0.15	0.083
Δ Testosterone \times Romney versus Obama	0.11	-0.14 – 0.36	0.392	0.08	-0.17 – 0.33	0.537
Δ Testosterone \times Romney versus Undecideds	-0.05	-0.56 – 0.46	0.85	-0.05	-0.57 – 0.46	0.837
Day- Linear \times Δ Diurnal Slope \times Romney versus Obama	0.13	0.02 – 0.23	0.019	0.13	0.02 – 0.23	0.019
Day- Quadratic \times Δ Diurnal Slope \times Romney versus Obama	-0.05	-0.15 – 0.06	0.391	-0.05	-0.15 – 0.06	0.39
Day- Linear \times Δ Diurnal Slope \times Romney versus Undecideds	0.05	-0.18 – 0.27	0.675	0.05	-0.18 – 0.27	0.678
Day- Quadratic \times Δ Diurnal Slope \times Romney versus Undecideds	-0.04	-0.26 – 0.18	0.695	-0.04	-0.26 – 0.18	0.693
Sex (0= Female, 1= Male)	-0.18	-0.41 – 0.05	0.128	-0.17	-0.40 – 0.06	0.158
Age				-0.01	-0.02 – 0.00	0.186
Socioeconomic Status				0.04	-0.02 – 0.10	0.17
Random Effects						
σ^2	0.06			0.06		
τ_{00}	0.36 _{pid}			0.36 _{pid}		
τ_{11}	0.01 _{pid.daypoly2.L}			0.01 _{pid.daypoly2.L}		
	0.00 _{pid.daypoly2.Q}			0.00 _{pid.daypoly2.Q}		
ρ_{01}	0.39			0.46		
	-0.37			-0.44		
N	113 _{pid}			113 _{pid}		
Observations	444			444		

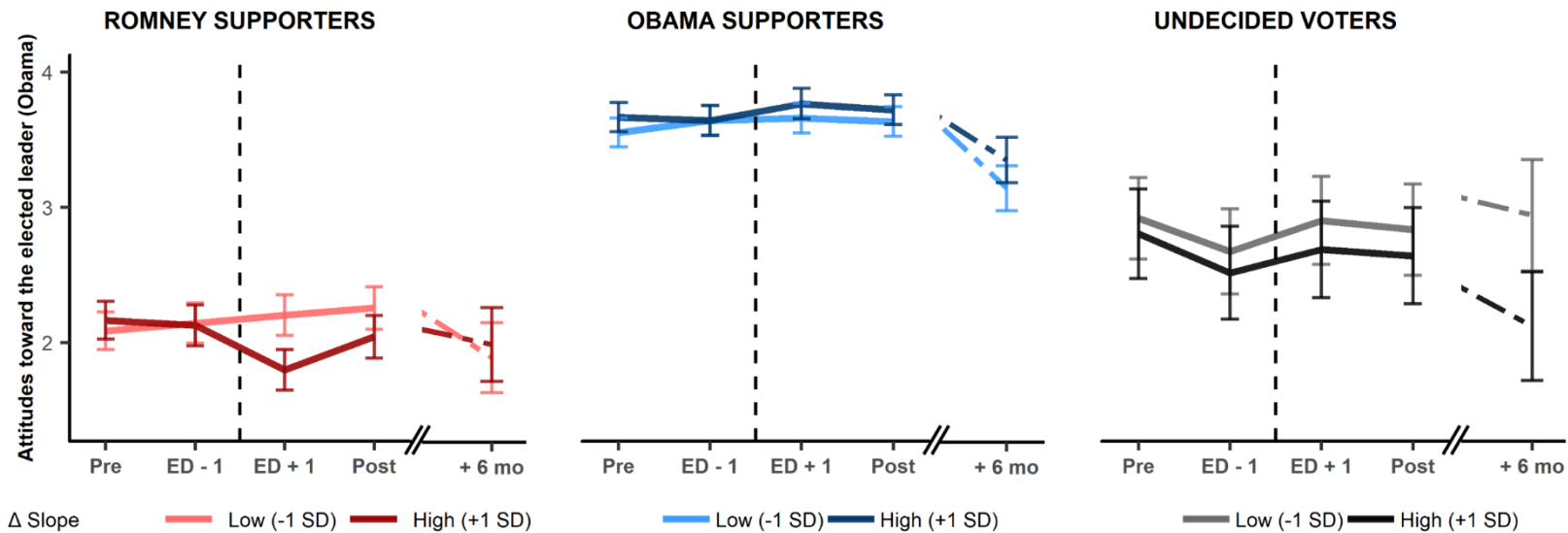


Fig. S6. Attitudes toward the elected leader (Obama) moderated by changes in diurnal slopes from ED-1 to ED, at one *SD* above and below the mean for each political group. Attitudes toward the elected leader were measured via self-report at the pre-election lab session (Pre), one day before the election (ED - 1), one day after the election (ED + 1), during the post-election lab session (Post), and in a subset of our sample at a six-month follow-up (+ 6mo). Colored lines represent estimated marginal means (controlling for sex) and error bars = ± 1 SE. The vertical dashed black line between *ED-1* and *ED+1* indicates when the election results were declared.

6. Linear regression analysis of diurnal slope changes predicting shifts in attitudes toward the elected leader.

We again conducted a linear regression to examine the extent to which diurnal slope changes (from *ED-1* to *ED*) predicted shifts in attitudes toward Obama using a difference score of attitudes reported before the election (average of attitudes at *Pre-Election Lab* session and *ED-1*) to after the election (average of attitudes at *ED+1* and the *Post-Election Lab* session). This analysis revealed a significant Diurnal Slope Change \times Group (Romney versus Obama) interaction ($B=0.17$, 95%CI[0.07, 0.27], $\beta=0.61$, $p=.001$, $d=0.78$; see Fig. S7 for scatterplot), corroborating findings described in the main text (and also in Fig. S6). Again, Romney supporters who showed flatter diurnal slopes demonstrated less positive attitudes toward Obama from before to after the election ($B=-.15$, $SE=.04$, $t_{(105)}=-3.91$, $p<.001$; $r=-.53$).

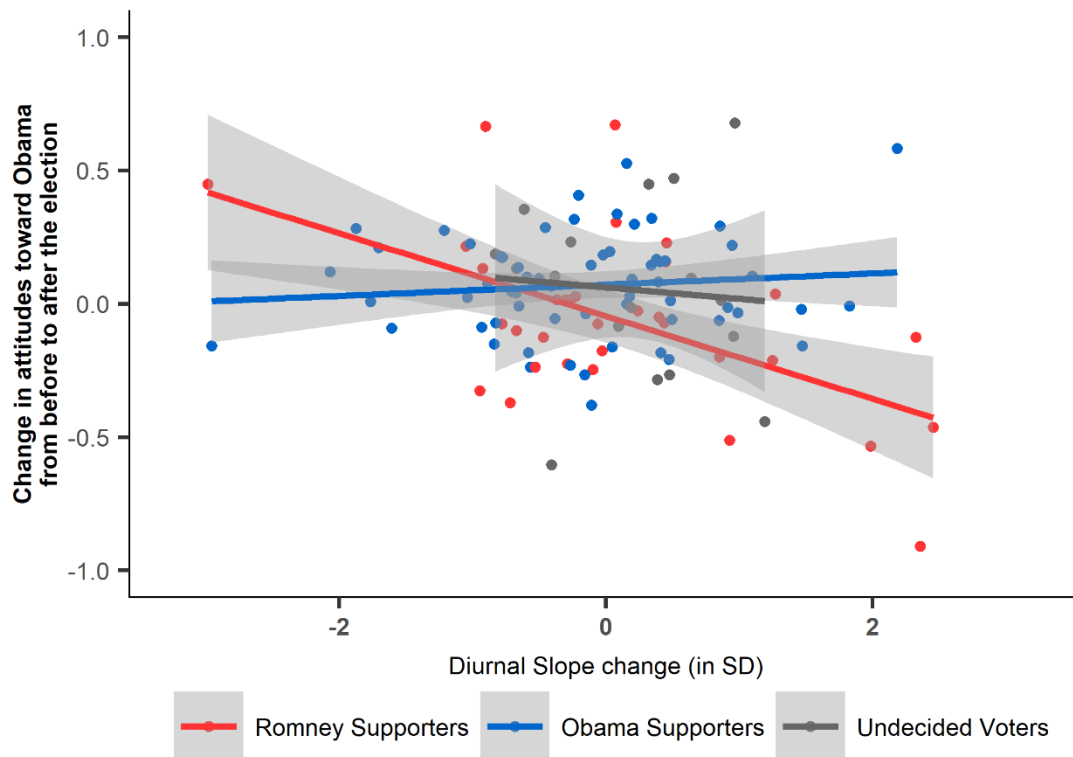


Fig. S7. Scatterplot between diurnal slope change from *ED-1* to *ED* (standardized within sex) and changes in attitudes toward the elected leader (Obama) from before the election (average of attitudes from the *Pre-Election Lab* and *ED-1*) to after the election (average of attitudes from *ED+1* and *Post-Election Lab*), across political groups. Shaded regions indicate SEs. *SD*= Standard deviation.

7. Diurnal slope changes predicting long-term shifts in attitudes toward the elected leader.

We also tested whether diurnal slope change moderated longer-term shifts in attitudes toward the elected leader six months after the election. We did not find a significant Time (Linear) \times Diurnal Slope Change \times Group (Romney versus Obama supporters) interaction ($B=0.11$, 95% CI[-0.08, 0.29], $\beta=.11$, $p=.260$; Table S18), suggesting that diurnal slope changes exerted short-term effects on attitudes but not effects that sustained six months after the election. We again confirmed that the attenuated association between diurnal slope change and long-term shifts in attitudes was not explained solely by the loss of statistical power due to attrition. The

subset of participants who completed the six-month follow-up showed a Time (Linear) \times Diurnal Slope Change \times Group interaction predicting short-term shifts in attitudes (Romney versus Obama: $B=0.15$, 95% $CI[0.00, 0.29]$, $\beta=.15$, $p=.048$). Collectively, these patterns of results suggest that diurnal slope changes predicted shifts in attitudes toward the elected leader in the short-term and that the association of diurnal slope changes with long-term attitudinal changes was weaker, albeit in the same direction.

Table S18. MLMs examining long-term shifts in attitudes toward the elected leader (Obama) moderated by diurnal slope change from *ED-1* to *ED*, across political groups, in a subset of our participants who consented to a six-month longitudinal follow-up ($n=66$). Model B includes the covariates of age and socioeconomic status.

DIURNAL SLOPE CHANGES PREDICTING LONG-TERM SHIFTS IN SELF-REPORTED ATTITUDES TOWARD THE ELECTED LEADER

	MODEL A			MODEL B		
	<i>B</i>	95% <i>CI</i>	<i>p</i>	<i>B</i>	95% <i>CI</i>	<i>p</i>
(Intercept)	2.14	1.81 – 2.47	<0.001	1.8	1.00 – 2.60	<0.001
Day- Linear	-0.19	-0.34 – -0.04	0.016	-0.19	-0.34 – -0.04	0.016
Romney supporters versus Obama supporters	1.37	1.01 – 1.73	<0.001	1.39	1.02 – 1.76	<0.001
Romney supporters versus Undecideds	0.71	0.25 – 1.16	0.003	0.7	0.22 – 1.18	0.006
Change in Diurnal slopes (Δ Diurnal Slope)	0.03	-0.20 – 0.26	0.803	0.02	-0.22 – 0.25	0.888
Day- Linear \times Romney versus Obama	-0.14	-0.33 – 0.06	0.172	-0.14	-0.34 – 0.06	0.17
Day- Linear \times Romney versus Undecideds	-0.15	-0.38 – 0.09	0.24	-0.15	-0.38 – 0.09	0.241
Day- Linear \times Δ Diurnal Slope	-0.01	-0.13 – 0.12	0.901	-0.01	-0.13 – 0.12	0.901
Δ Testosterone \times Romney versus Obama	0.01	-0.33 – 0.35	0.943	0.04	-0.31 – 0.39	0.831
Δ Testosterone \times Romney versus Undecideds	-0.41	-1.01 – 0.19	0.185	-0.4	-1.00 – 0.21	0.206
Day- Linear \times Δ Diurnal Slope \times Romney versus Obama	0.11	-0.08 – 0.29	0.26	0.11	-0.08 – 0.29	0.262
Day- Linear \times Δ Diurnal Slope \times Romney versus Undecideds	-0.22	-0.54 – 0.11	0.201	-0.22	-0.54 – 0.11	0.201
Sex (0= Female, 1= Male)	-0.17	-0.49 – 0.16	0.323	-0.15	-0.48 – 0.18	0.365
Age				0.01	-0.01 – 0.03	0.385
Socioeconomic Status				0.02	-0.06 – 0.10	0.651
Random Effects						
σ^2	0.09			0.09		
τ_{00}	0.37 _{pid}			0.38 _{pid}		
τ_{11}	0.03 _{pid.daypoly11}			0.03 _{pid.daypoly11}		
ρ_{01}	0.04 _{pid}			0.03 _{pid}		
N	66 _{pid}			66 _{pid}		
Observations	196			196		

8. Sex differences and similarities

We included both males and females in all our main analysis to maximize statistical power. To further probe whether the pattern of our results varied across men and women, we added sex as a moderator of the effects of interest. For example, to test for sex differences or similarities in acute testosterone change, we conducted a Time \times Group \times Sex interaction predicting testosterone concentrations. We report follow up analyses separately across males and females for a given effect (e.g., Time \times Group) for interactions with sex that were significant.

Acute testosterone reactivity during the competition-outcome window. Consistent with previous research, we found a significant Sex \times Time (Linear) \times Group (Romney versus Obama supporters) interaction ($B=-26.37$, 95% CI $[-48.69, -4.04]$, $\beta=-0.59$, $p=.023$; Geniole et al., 2017; Stanton et al., 2009). Separate analyses across sexes indicated that these acute changes in testosterone were primarily driven by the males in our sample (Males: Time (Linear) \times Group (Romney versus Obama supporters): $B=-28.63$, 95% CI $[-51.57, -5.69]$, $\beta=-0.63$, $p=.018$; Females: Time (Linear) \times Group (Romney versus Obama supporters): $B=-2.12$, 95% CI $[-9.85, 5.62]$, $\beta=-0.11$, $p=.594$), though, descriptively, we note a similar directional pattern for both sexes.

Acute testosterone reactivity predicting shifts in attitudes toward the elected leader. Although we found a sex difference in acute testosterone reactivity to the election, we did not find evidence for sex differences in the analysis examining acute testosterone-moderated short-term shifts in attitudes (Sex \times Time (Linear) \times Group (Romney versus Obama supporters) \times Acute testosterone change: $B= -0.18$, 95% CI $[-0.43, 0.06]$, $\beta=-0.20$, $p=.141$), long-term shifts in attitudes (Sex \times Time (Linear) \times Group (Romney versus Obama supporters) \times Acute testosterone change: $B= 0.12$, 95% CI $[-0.34, 0.59]$, $\beta=0.15$, $p=.603$), and using a composite measure of attitudes via self-report and Facebook (Sex \times Time (Linear) \times Group (Romney versus Obama supporters) \times Acute testosterone change: $B= -0.02$, 95% CI $[-0.27, 0.24]$, $\beta=-0.02$, $p=.897$).

Changes in diurnal slopes. We examined sex differences in the diurnal slope changes across the naturalistic sampling period. First, we note that both males and females demonstrated diurnal rhythms: testosterone levels were high in the morning and dropped at the end of the day (Males: $B= -104.08$, 95% CI $[-111.59, -96.58]$, $p<.001$; Females: $B=-41.09$, 95% CI $[-44.07, -38.11]$, $p<.001$). On further examining sex differences across the naturalistic sampling period, we found that the trajectory of diurnal slopes across political groups differed significantly in terms of their magnitude of change across both sexes (Sex \times Time (Linear) \times Group (Romney versus Obama supporters): $B=-12.55$, 95% CI $[-22.01, -3.09]$, $\beta=-0.65$, $p=.010$). This pattern was consistent with the sex differences we observed for acute testosterone reactivity. Changes in diurnal slopes across the sampling period were primarily driven by male participants (Males: Time (Linear) \times Group (Romney versus Obama supporters): $B=-13.39$, 95% CI $[-23.21, -3.56]$, $\beta=-0.76$, $p=.010$; Females: Time (Linear) \times Group (Romney versus Obama supporters): $B=-0.83$, 95% CI $[-4.08, 2.41]$, $\beta=-.09$, $p=.615$). Males who supported the losing candidate demonstrated flatter slopes compared to females, although we again note a similar directional pattern across both sexes.

Diurnal slope changes predicting shifts in attitudes toward Obama. We did not find evidence for sex differences in the analysis examining diurnal slope change moderating short-

term shifts in attitudes (Sex \times Time (Linear) \times Group (Romney versus Obama supporters) \times Diurnal slope change: $B=-0.01$, 95% CI[-0.22, 0.21], $\beta=-0.01$, $p=.958$), long-term shifts in attitudes (Sex \times Time (Linear) \times Group (Romney versus Obama supporters) \times Diurnal slope change: $B=-0.25$, 95% CI[-0.62, 0.12], $\beta=-0.26$, $p=.188$), and using a composite measure of attitudes via self-report and Facebook (Sex \times Time (Linear) \times Group (Romney versus Obama supporters) \times Diurnal slope change): $B=-0.06$, 95% CI[-0.27, 0.15], $\beta=-.07$, $p=.559$).

Overall, we note mixed evidence for sex differences in the results. Although we found stronger testosterone changes to the electoral competition in males relative to females, there emerged no sex differences in these testosterone changes predicting shifts in attitudes toward the elected leader.

9. ROBUSTNESS CHECKS

9.1. Model comparisons between acute testosterone reactivity and diurnal slope changes predicting shifts in attitudes toward the elected leader.

In this study, we examined how two indices of testosterone change predicted shifts in attitudes toward the elected leader: acute testosterone reactivity and diurnal slope change. These two metrics were positively correlated ($r=.41$, $95\%CI[0.24, 0.56]$). To compare the robustness of each, we conducted additional analyses wherein we included both indices of testosterone change, and their corresponding interactions in an expanded model (Table S19; for similar model comparisons see Akinola et al., 2018). In this model, we found that acute testosterone reactivity was a stronger predictor of shifts in attitudes (Time (Linear) \times Acute Testosterone Change \times Group (Romney versus Obama supporters): $B=0.26$, $95\%CI[0.13, 0.39]$, $\beta=0.28$, $p<.001$) than diurnal slope change (Time (Linear) \times Diurnal Slope Change \times Group (Romney versus Obama supporters): $B=0.07$, $95\%CI[-0.05, 0.18]$, $\beta=.07$, $p=.243$; Table S19). Next, we compared the model fit indices of this expanded model to two individual models: one model of the effects of acute testosterone reactivity, and one model of the effects of diurnal slope changes as predictors of shifts in attitudes toward the elected leader. To ensure equivalence across both models, we only compared those participants with available data across both testosterone-change predictors (resultant $n=106$).⁷ These comparisons showed that the model which included only diurnal slope change predicting shifts in attitudes (AIC=391.93; BIC=496.85) demonstrated significantly poorer model fit than the expanded model (AIC=375.50; BIC=516.74; $\Delta\chi^2_{(9)}=34.43$, $p<.001$). However, the model that included acute testosterone reactivity (AIC=372.46; BIC=477.38) demonstrated similar model fit compared to the expanded model ($\Delta\chi^2_{(9)}=14.96$, $p=.092$). These results suggest that acute testosterone reactivity may be a more parsimonious and robust predictor of attitude change, compared to diurnal slope change.

⁷ In this smaller sample ($n=106$), models with acute testosterone reactivity and diurnal slope changes as predictors of shifts in attitudes remained significant and similar in magnitude to the effects reported in the main text and Tables S8 and S14 respectively (Time (Linear) \times Acute testosterone reactivity \times Group (Romney versus Obama supporters) interaction: $B=0.29$, $95\%CI[0.17, 0.41]$, $p<.001$; Time (Linear) \times Diurnal slope change \times Group (Romney versus Obama supporters) interaction: $B=0.13$, $95\%CI[0.02, 0.24]$, $p=.018$).

Table S19. MLMs examining changes in attitudes toward Obama) moderated by acute testosterone reactivity and diurnal slope changes from *ED-1* to *ED*, across political groups. Model B includes the covariates of age and socioeconomic status.

ACUTE TESTOSTERONE REACTIVITY AND DIURNAL SLOPE CHANGES PREDICTING SELF-REPORTED ATTITUDES						
	MODEL A			MODEL B		
	<i>B</i>	95% <i>CI</i>	<i>p</i>	<i>B</i>	95% <i>CI</i>	<i>p</i>
(Intercept)	2.28	2.03 – 2.54	<0.001	2.21	1.58 – 2.83	<0.001
Day- Linear	0.06	-0.03 – 0.16	0.186	0.06	-0.03 – 0.16	0.186
Day- Quadratic	0.10	0.01 – 0.20	0.033	0.1	0.01 – 0.20	0.034
Romney supporters versus Obama supporters	1.48	1.21 – 1.75	<0.001	1.49	1.23 – 1.75	<0.001
Romney supporters versus Undecideds	0.82	0.44 – 1.20	<0.001	1.00	0.61 – 1.39	<0.001
Acute Change in Testosterone (Δ Acute Testosterone)	-0.16	-0.42 – 0.09	0.205	-0.19	-0.44 – 0.06	0.14
Diurnal slope change (Δ Diurnal Slope)	-0.01	-0.21 – 0.18	0.886	0.01	-0.18 – 0.20	0.919
Day- Linear \times Romney versus Obama	0.00	-0.11 – 0.11	0.978	0.00	-0.11 – 0.11	0.97
Day- Quadratic \times Romney versus Obama	-0.13	-0.24 – -0.02	0.026	-0.13	-0.24 – -0.02	0.026
Day- Linear \times Romney versus Undecideds	-0.05	-0.21 – 0.11	0.552	-0.05	-0.21 – 0.11	0.565
Day- Quadratic \times Romney versus Undecideds	-0.01	-0.17 – 0.15	0.886	-0.01	-0.17 – 0.15	0.898
Day- Linear \times Δ Acute Testosterone	-0.25	-0.35 – -0.14	<0.001	-0.25	-0.35 – -0.14	<0.001
Day- Quadratic \times Δ Testosterone	-0.13	-0.23 – -0.02	0.020	-0.13	-0.23 – -0.02	0.021
Δ Acute Testosterone \times Romney versus Obama	-0.01	-0.33 – 0.31	0.951	0.03	-0.28 – 0.34	0.842
Δ Acute Testosterone \times Romney versus Undecideds	0.30	-0.04 – 0.65	0.089	0.34	0.00 – 0.67	0.052
Day- Linear \times Δ Diurnal Slope	-0.07	-0.15 – 0.01	0.091	-0.07	-0.15 – 0.01	0.091
Day- Quadratic \times Δ Diurnal Slope	0.11	0.03 – 0.20	0.009	0.11	0.03 – 0.20	0.009
Δ Diurnal Slope \times Romney versus Obama	0.16	-0.11 – 0.43	0.256	0.1	-0.17 – 0.37	0.465
Δ Diurnal Slope \times Romney versus Undecideds	-0.16	-0.67 – 0.35	0.531	-0.19	-0.69 – 0.31	0.456
Day- Linear \times Δ Acute Testosterone \times Romney versus Obama	0.26	0.13 – 0.39	<0.001	0.26	0.13 – 0.39	<0.001
Day- Quadratic \times Δ Acute Testosterone \times Romney versus Obama	0.12	-0.01 – 0.25	0.072	0.12	-0.01 – 0.25	0.072
Day- Linear \times Δ Acute Testosterone \times Romney versus Undecideds	0.24	0.10 – 0.38	0.001	0.24	0.10 – 0.38	0.001
Day- Quadratic \times Δ Acute Testosterone \times Romney versus Undecideds	0.17	0.03 – 0.32	0.019	0.17	0.03 – 0.32	0.019
Day- Linear \times Δ Diurnal Slope \times Romney versus Obama	0.07	-0.05 – 0.18	0.243	0.07	-0.05 – 0.18	0.241
Day- Quadratic \times Δ Diurnal Slope \times Romney versus Obama	-0.08	-0.19 – 0.03	0.173	-0.08	-0.20 – 0.03	0.173
Day- Linear \times Δ Diurnal Slope \times Romney versus Undecideds	0.02	-0.21 – 0.24	0.895	0.01	-0.21 – 0.23	0.915
Day- Quadratic \times Δ Diurnal Slope \times Romney versus Undecideds	-0.11	-0.33 – 0.11	0.342	-0.11	-0.33 – 0.11	0.334
Sex (0= Female, 1= Male)	-0.20	-0.43 – 0.03	0.097	-0.19	-0.42 – 0.04	0.113
Age				-0.01	-0.03 – 0.00	0.088
SES				0.07	0.00 – 0.13	0.043
Random Effects						
σ^2	0.06			0.06		
τ_{00}	0.33 _{pid}			0.31 _{pid}		
τ_{11}	0.00 _{pid.daypoly2.L}			0.00 _{pid.daypoly2.L}		
	0.00 _{pid.daypoly2.Q}			0.00 _{pid.daypoly2.Q}		
ρ_{01}	1			1		
	-1			-1		
N	106 _{pid}			106 _{pid}		
Observations	418			418		

9.2. Attitudes toward elected leader expressed via behaviors on Facebook.

In addition to measuring attitudes toward the elected leader via self-reports, we also coded for attitudes toward Obama that were expressed via online behaviors on Facebook (Gil, 2012; Carlisle & Patton, 2013). We found no evidence of either acute testosterone reactivity or diurnal slopes changes predicting shifts in attitudes toward the elected leader with Facebook posts ($ps > .52$). The relatively small sample size of participants with useable social media data may have limited our statistical power to test and detect three-way interactions. Despite not finding effects of testosterone change metrics predicting shifts in attitudes expressed via behaviors on social media, we tested for behavioral differences toward the elected leader (Obama) across political groups. These analyses revealed a Time (Linear) \times Group (Romney versus Obama supporters) interaction: $B=0.61$, 95% CI[0.22, 1.00], $\beta=0.55$, $p=.003$). Simple slopes analysis indicated that Obama supporters revealed more positive attitudes on social media toward Obama from before to after the election (Time (Linear): $B=0.32$, 95% CI[0.07, 0.56]; see Fig S8). These results were consistent with a pattern of increases in pro-Obama attitudes amongst Obama supporters also found in self-reports (Time (Linear): $B=0.06$, 95% CI[0.00, 0.13]; see Fig. 3, Middle Panel).

Although changes in testosterone did not significantly predict shifts in social media behavior (likely due to lack of statistical power), political attitudes measured via self-report and social media behavior correlated well with one another (see Table S20 for correlations across all time points between self-reports and social media behavior). These correlations are consistent with burgeoning evidence that online social media statements correlate with offline political opinions and political behaviors (O'Connor, et al., 2010; Kim et al., 2016). Therefore, we created composite scores of attitudes toward the elected leader across self-report and behavior on social media – one score before the election and one score after the election.

We also note that that self-reported political attitudes tended to correlate strongly across measurement instances (see Table S20). The broader literature has also reported similar within-person stability in political attitudes over time (Hatemi & Verhulst, 2015; Alabastro et al., 2013; Crano & Prislin, 2011; Krosnick, 1991). Despite this stability, there is also evidence that political attitudes fluctuate within-individuals in response to political events such as wars, changes in the economy, elections (Alabastro et al., 2013; Lebo & Cassino, 2007; Kriner & Schwartz, 2009) increased exposure to media (Young, 2004), and especially toward the out-group leader (Alabastro et al., 2013; Lebo & Cassino, 2007). Therefore, the testosterone-moderated fluctuations in political attitudes that we report in our study support the notion that although political attitudes are relatively stable, they may dynamically fluctuate in response to a large-scale political event (i.e., a democratic election).

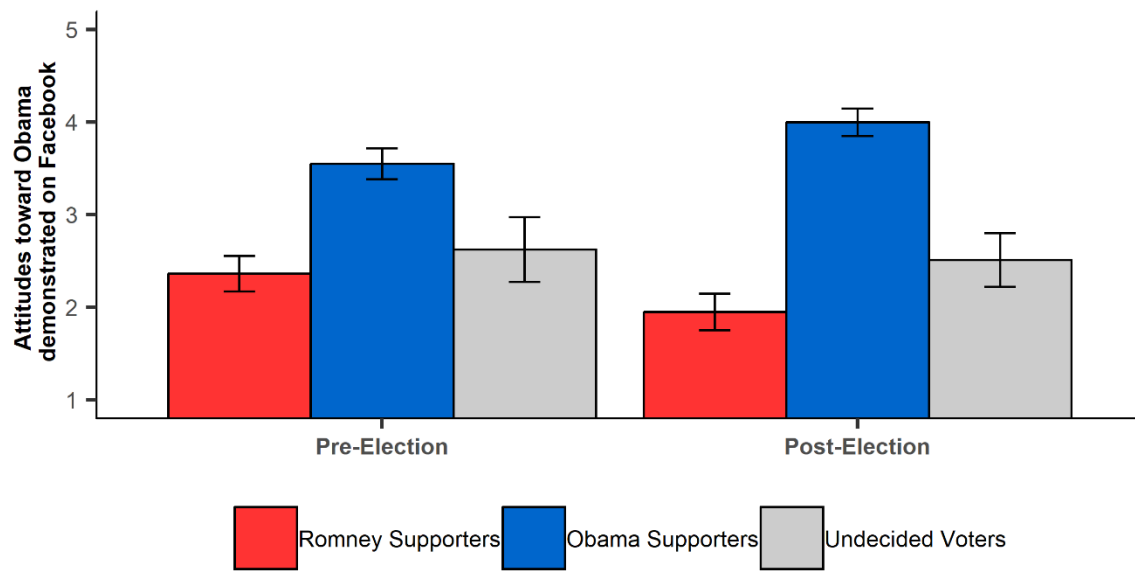


Fig. S8. Attitudes toward the elected leader (Obama) via online behaviors on Facebook from before the election (*Pre-Election*: from *ED*-2 days through 8 pm on *ED*) to after (*Post-Election*: from 8 pm on *ED* through *ED*+2 days). Error bars = ± 1 SE.

Table S20. Correlations between elected leader (Obama) via self-reports and online behaviors on Facebook. *Note.* Values in square brackets indicate the 95% confidence intervals for each pairwise correlation. * $p < .05$. ** $p < .01$.

Variable	1	2	3	4	5	6
1. Pre-Election self-report						
2. ED-1 self-report	.92** [.88, .94]					
3. ED+1 self-report	.94** [.91, .96]	.94** [.91, .96]				
4. Post-Election self-reports	.92** [.88, .94]	.92** [.89, .95]	.94** [.92, .96]			
5. ED + 6 months self-report	.84** [.75, .90]	.85** [.77, .91]	.86** [.77, .91]	.86** [.78, .91]		
6. Pre-Election social media	.68** [.49, .82]	.68** [.48, .81]	.69** [.49, .82]	.62** [.39, .78]	.49* [.11, .75]	
7. Post-Election social media	.76** [.61, .85]	.74** [.59, .84]	.73** [.57, .84]	.70** [.52, .81]	.70** [.43, .85]	.70** [.50, .83]

Table S21. Means and *SDs* of attitudes toward the elected leader reported across the naturalistic sampling period (measured on a 5-point Likert scale).

	<i>Romney Supporters</i>		<i>Obama Supporters</i>		<i>Undecided Voters</i>	
Self-reported attitudes	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pre-Election	2.12	0.56	3.61	0.60	3.00	0.84
ED-1	2.13	0.55	3.65	0.59	2.74	0.80
ED+1	1.99	0.71	3.72	0.64	3.03	0.89
Post-Election	2.14	0.66	3.68	0.59	3.00	0.82
+6 months	1.86	0.73	3.06	0.58	2.48	0.94
Social media behavior						
Pre-Election	2.41	0.87	3.59	0.67	2.80	0.45
Post-Election	2.00	0.82	4.03	0.91	2.62	0.74

Next, we tested how acute testosterone reactivity and change in diurnal slopes predicted shifts in attitudes toward the elected leader as a function of political group using the composite attitude score (created by averaging self-reported attitudes and online social media behavior).

9.2.1. Acute testosterone reactivity predicting shifts in composite attitudes toward the elected leader. Corroborating the primary findings reported in the main text, MLM analysis revealed a Time (Linear) \times Acute testosterone reactivity \times Group (Romney versus Obama supporters) interaction using the composite attitudes measure ($B=0.17$, 95% CI[0.05, 0.29], $\beta=.19$, $p=.008$; see Table S22). Supporters of the losing candidate with greater increases in acute testosterone demonstrated less positive attitudes toward the elected leader (Time (Linear) \times Acute testosterone reactivity: $B=-0.16$, 95% CI[-0.27, -0.06], $\beta=-0.26$, $p=.005$; see Fig. S9).

9.2.2. Diurnal slope changes predicting shifts in composite attitudes toward the elected leader. Finally, we found marginal support for diurnal slope changes moderating shifts in attitudes toward the elected leader using our composite measure of attitudes (Time (Linear) \times Diurnal Slope change \times Group (Romney versus Obama supporters): $B=0.10$, 95% CI[-0.01, 0.20], $\beta=.10$, $p=.066$; see Table S23). Supporters of the losing candidate with flatter slopes on the day of the election demonstrated less positive attitudes toward the elected leader at a trend level (Time (Linear) \times Diurnal Slope change: $B=-0.07$, 95% CI[-0.15, .02], $\beta=-.13$, $p=.135$).

Table S22. MLMs examining changes in attitudes toward the elected leader (Obama) using a composite measure of self-reported attitudes and online social media behavior moderated by acute testosterone reactivity during the competition-outcome window, across political groups. Model B includes the covariates of age and socioeconomic status.

ACUTE TESTOSTERONE REACTIVITY PREDICTING SHIFTS IN ATTITUDES TOWARD OBAMA (Facebook + Self-report Composite)						
	MODEL A			MODEL B		
	<i>B</i>	<i>95% CI</i>	<i>p</i>	<i>B</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	2.31	2.08 – 2.55	<0.001	2.38	1.80 – 2.96	<0.001
Day- Linear	-0.05	-0.15 – 0.05	0.309	-0.05	-0.15 – 0.05	0.309
Romney supporters versus Obama supporters	1.47	1.21 – 1.72	<0.001	1.48	1.23 – 1.72	<0.001
Romney supporters versus Undecideds	0.72	0.37 – 1.06	<0.001	0.87	0.51 – 1.22	<0.001
Acute Change in Testosterone (Δ Testosterone)	-0.09	-0.32 – 0.13	0.414	-0.10	-0.32 – 0.12	0.382
Day- Linear \times Romney versus Obama	0.15	0.04 – 0.27	0.011	0.15	0.04 – 0.27	0.011
Day- Linear \times Romney versus Undecideds	0.05	-0.10 – 0.21	0.519	0.05	-0.10 – 0.21	0.519
Day- Linear \times Δ Testosterone	-0.16	-0.26 – -0.06	0.002	-0.16	-0.26 – -0.06	0.002
Δ Testosterone \times Romney versus Obama	0.01	-0.26 – 0.29	0.915	0.02	-0.25 – 0.29	0.877
Δ Testosterone \times Romney versus Undecideds	0.20	-0.11 – 0.52	0.209	0.22	-0.08 – 0.53	0.156
Day- Linear \times Δ Testosterone \times Romney versus Obama	0.17	0.05 – 0.29	0.008	0.17	0.05 – 0.29	0.008
Day- Linear \times Δ Testosterone \times Romney versus Undecideds	0.14	-0.00 – 0.28	0.053	0.14	-0.00 – 0.28	0.053
Sex (0= Female, 1= Male)	-0.24	-0.45 – -0.02	0.038	-0.24	-0.46 – -0.02	0.033
Age				0.04	-0.02 – 0.10	0.149
Socioeconomic Status				-0.01	-0.03 – 0.00	0.053
Random Effects						
σ^2	0.06			0.06		
τ_{00}	0.27 _{pid}			0.26 _{pid}		
N	106 _{pid}			106 _{pid}		
Observations	212			212		

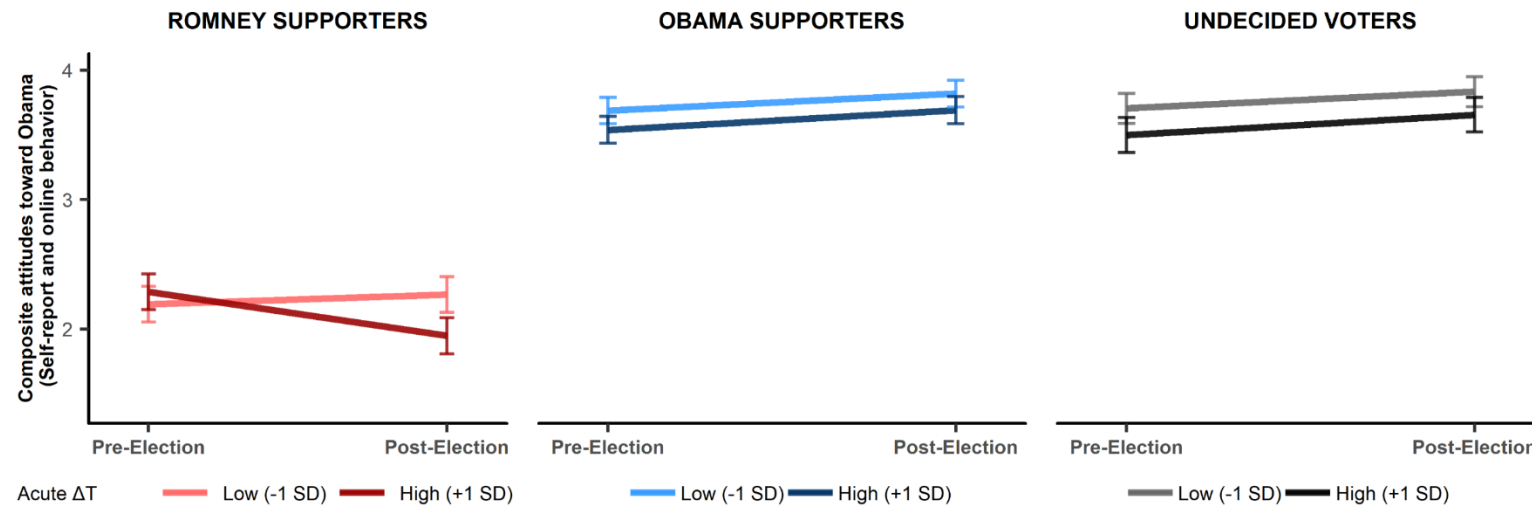


Fig. S9. Composite attitudes toward the elected leader (Obama) as a function of acute testosterone reactivity during the competition-outcome window one SD above and below the mean for each political group. Composite attitudes toward Obama were measured via self-report and online behaviors on Facebook. Solid lines represent estimated marginal means (controlling for sex). Error bars = ± 1 SE.

Table S23. MLMs examining changes in attitudes toward the elected leader (Obama) using a composite measure of self-reported attitudes and online social media behavior, across political groups, moderated by diurnal slope changes from *ED-1* to *ED*. Model B includes the covariates of age and socioeconomic status.

DIURNAL SLOPE CHANGES PREDICTING SHIFTS IN ATTITUDES TOWARD THE ELECTED LEADER
(Facebook + Self-report Composite)

	MODEL A			MODEL B		
	<i>B</i>	95% <i>CI</i>	<i>p</i>	<i>B</i>	95% <i>CI</i>	<i>p</i>
(Intercept)	2.24	2.02 – 2.46	<0.001	2.19	1.64 – 2.74	<0.001
Day- Linear	-0.09	-0.17 – 0.00	0.056	-0.09	-0.17 – 0.00	0.056
Romney supporters versus Obama supporters	1.54	1.30 – 1.78	<0.001	1.55	1.31 – 1.79	<0.001
Romney supporters versus Undecideds	0.71	0.37 – 1.05	<0.001	0.80	0.44 – 1.15	<0.001
Change in Diurnal slopes (Δ Diurnal Slope)	-0.11	-0.28 – 0.06	0.208	-0.10	-0.27 – 0.07	0.246
Day- Linear \times Romney versus Obama	0.19	0.08 – 0.30	0.001	0.19	0.08 – 0.30	0.001
Day- Linear \times Romney versus Undecideds	0.08	-0.07 – 0.23	0.29	0.08	-0.07 – 0.23	0.290
Day- Linear \times Δ Diurnal Slope	-0.07	-0.14 – 0.01	0.097	-0.07	-0.14 – 0.01	0.097
Δ Diurnal Slope \times Romney versus Obama	0.15	-0.08 – 0.38	0.198	0.13	-0.10 – 0.36	0.275
Δ Diurnal Slope \times Romney versus Undecideds	-0.01	-0.48 – 0.46	0.966	-0.01	-0.48 – 0.45	0.954
Day- Linear \times Δ Diurnal Slope \times Romney versus Obama	0.1	-0.01 – 0.20	0.066	0.10	-0.01 – 0.20	0.066
Day- Linear \times Δ Diurnal Slope \times Romney versus Undecideds	-0.01	-0.22 – 0.20	0.914	-0.01	-0.22 – 0.20	0.914
Sex (0= Female, 1= Male)	-0.21	-0.43 – 0.00	0.054	-0.21	-0.42 – 0.01	0.063
Age				0.03	-0.02 – 0.09	0.249
Socioeconomic Status				-0.01	-0.02 – 0.01	0.311
Random Effects						
σ^2	0.06			0.06		
τ_{00}	0.28 _{pid}			0.28 _{pid}		
N	113 _{pid}			113 _{pid}		
Observations	226			226		

9.3. Affective change as a covariate

As an additional robustness check, we examined whether the main findings occurred independent of changes in participant affect. Therefore, we entered affective changes that occurred during the competition-outcome window (i.e., difference scores of positive affect and negative affect from 7pm to Bedtime on ED) as covariates into all the main models.

Acute testosterone reactivity during the competition-outcome window. The model examining changes in testosterone across the competition-outcome window revealed a robust Time (Linear) \times Group (Romney versus Obama supporters) interaction after entering changes in positive affect ($B=-14.30$, 95% CI[-26.64, -1.96], $\beta=-0.32$, $p=.025$) and negative affect ($B=-15.03$, 95% CI[-27.82, -2.24], $\beta=-0.33$, $p=.023$) as covariates.

Acute testosterone reactivity predicting shifts in attitudes toward the elected leader. Further, acute testosterone reactivity on election night also robustly predicted short-term shifts in attitudes toward the elected leader over and above changes in positive affect ($B=0.30$, 95% CI[0.18, 0.42], $\beta=0.34$, $p<.001$) and negative affect ($B=0.30$, 95% CI[0.18, 0.42], $\beta=.34$, $p<.001$).

Changes in diurnal slopes. The model examining changes in diurnal slopes across the naturalistic sampling period revealed a robust Time (Linear) \times Group (Romney versus Obama supporters) interaction after entering changes in positive affect ($B=-7.38$, 95% CI[-12.65, -2.11], $\beta=-0.38$, $p=.007$) and negative affect ($B=-7.99$, 95% CI[-13.42, -2.56], $\beta=-0.42$, $p=.005$) as covariates.

Diurnal slope changes predicting shifts in attitudes toward Obama. Finally, diurnal slope changes on election night predicted short-term shifts in attitudes toward the elected leader independent of changes in positive affect ($B=0.13$, 95% CI[0.02, 0.23], $\beta=0.13$, $p=.020$) and negative affect ($B=0.13$, 95% CI[0.02, 0.24], $\beta=0.14$, $p=.027$).

Overall, the main results reported in the paper were largely robust to changes in affect that occurred during the competition-outcome window.

9.4. Main results after the exclusion of the Gary Johnson supporter

In our sample, supporters of the losing candidate primarily consisted of Romney supporters. This group however also included one Gary Johnson supporter (a Libertarian candidate who also lost the election). Given that we were primarily interested in comparing supporters of winning and losing candidates, we included this participant within the group of Romney supporters to boost statistical power of the group. Here we report all our main results after excluding this individual from the analyses.

Acute testosterone reactivity during the competition-outcome window. The model examining acute changes in testosterone across the competition-outcome window revealed a significant Time (Linear) \times Group (Romney versus Obama supporters) interaction after excluding the Gary Johnson supporter ($B=-15.65$, 95% CI[-27.15, -4.15], $\beta=-0.35$, $p=.009$). Further, Romney supporters continued to resist the typical circadian decline (i.e., had flatter diurnal testosterone slopes; Time (Linear): $B=4.80$, 95% CI[-4.34, 13.93], $\beta=.09$, $p=.312$).

Acute testosterone reactivity predicting shifts in attitudes toward the elected leader. After excluding the Gary Johnson supporter, our analysis examining acute testosterone changes moderating shifts in attitudes toward the elected leader, continued to reveal a significant three-way Time (Linear) \times Acute Testosterone Reactivity \times Group interaction (Romney versus Obama: $B=0.29$, 95% CI[0.17, 0.41], $\beta=0.32$, $p<.001$). These testosterone-moderated attitude shifts were driven primarily by Romney supporters (Time (Linear) \times Acute Testosterone Reactivity: $B=-0.28$, 95% CI[-0.39, -0.16], $\beta=-0.39$, $p<.001$).

Changes in diurnal slopes. The model examining changes in diurnal slopes across the naturalistic sampling period still revealed a robust Time (Linear) \times Group (Romney versus Obama supporters) interaction upon excluding the Gary Johnson supporter ($B=-7.06$, 95% CI[-12.00, -2.12], $\beta=-0.37$, $p=.006$). Romney supporters continued to exhibit flatter slopes (i.e., decreased magnitude of the diurnal slopes) across the sampling period ($B=7.39$, 95% CI[1.86, 12.92], $\beta=0.37$, $p=.013$).

Diurnal slope changes predicting shifts in attitudes toward Obama. After excluding the Gary Johnson supporter, our analysis examining diurnal slope changes moderating shifts in attitudes toward the elected leader continued to reveal a significant three-way Time (Linear) \times Diurnal slope change \times Group interaction (Romney versus Obama: $B=0.13$, 95% CI[0.02, 0.23], $\beta=0.13$, $p=.020$). Further, these diurnal slope change moderated short-term shifts in attitudes were driven primarily by Romney supporters ($B=-0.1295$, 95% CI[-0.22, -0.02], $\beta=-0.22$, $p=.023$).

Overall, we note that even if we exclude the Gary Johnson supporter from our sample, our results and interpretations remain unchanged.

9.5. Robustness checks based on when participants learned about the outcome of the election

Media networks called the 2012 election at 8:12 pm PST (Ariens, 2016) and on average participants learned about the election outcome at 8:50 pm PST. To assess whether *when* participants learned about the results altered our main results, we conducted robustness checks wherein we controlled for the time difference between when participants reported they learned about the results and when the results were declared. We note that 8 participants did not report when they learned about the election results but reported the times of their 7 pm, 9 pm, and *Bedtime* samples and we therefore imputed 8:12 pm as the time they learned about the outcome of the election. We report all our main findings with and without the imputed values. Also, two participants in our sample reported learning about the results after they provided their *Bedtime* sample; therefore, we also report analyses while excluding these two participants from our main analyses. The pattern of results remained largely unchanged with these additional robustness checks.

Table S24. Robustness check for all our main results based on when participants learned about the outcome of the election.

	Controlling for the difference between when the results were declared and when participants learned about the outcome		Excluding two participants who learned about the election after providing their <i>Bedtime</i> sample
	Imputed	No Imputation	
Acute testosterone reactivity during the competition-outcome window.	$B=-14.63$, 95%CI[-26.05, -3.20], $\beta=-0.33$, $p=.014$	$B=-13.85$, 95%CI[-25.88, -1.82], $\beta=-0.33$, $p=.026$	$B=-14.54$, 95%CI[-26.15, -2.94], $\beta=-0.32$, $p=.016$
Acute testosterone reactivity predicting shifts in attitudes toward the elected leader.	$B=0.29$, 95% CI[0.17, 0.41], $\beta=0.31$, $p<.001$	$B=0.25$, 95% CI[0.12, 0.39], $\beta=0.28$, $p<.001$	$B=0.29$, 95% CI[0.16, 0.41], $\beta=0.31$, $p<.001$
Changes in diurnal slopes	$B=-7.30$, 95% CI[-12.19, -2.41], $\beta=-0.38$, $p=.004$	$B=-5.63$, 95%CI[-10.39, -0.88], $\beta=-0.31$, $p=.022$	$B=-7.47$, 95% CI[-12.43, -2.51], $\beta=-0.39$, $p=.004$
Diurnal slope changes predicting shifts in attitudes toward Obama	$B=0.13$, 95% CI[0.02, 0.23], $\beta=0.13$, $p=.019$	$B=0.10$, 95% CI[-0.02, 0.22], $\beta=0.10$, $p=.091$	$B=0.13$, 95% CI[0.02, 0.23], $\beta=0.13$, $p=.017$

9.6. Robustness checks based on when participants completed the Pre- and Post-Lab Session

Participants were scheduled to come in for the laboratory sessions based on their availability in the days before and after the election. Therefore, there was variability in when participants completed the Pre and Post-Election Lab session attitude surveys: On average, the Pre-election Lab session was conducted 4.37 ($SD=1.14$) days before the election and Post-election Lab session 4.46 ($SD=1.33$) days after the election. Upon examining whether testosterone-moderated shifts in attitudes occurred independent of when participants completed their attitudes survey, we note that our results remain robust and unchanged.

Table S25. Robustness check based on when participants completed the Pre- and Post-Election Lab sessions for analysis examining shifts in attitudes toward the elected leader (Obama), moderated by acute testosterone reactivity and diurnal slope changes from ED-1 to ED.

	Controlling for number of days separating election day and the Pre- Election Lab session	Controlling for number of days separating election day and the Post- Election Lab session	Controlling for number of days between election day and the Pre- and Post-Election Lab sessions
Acute testosterone reactivity predicting shifts in attitudes toward the elected leader.	$B=0.29$, 95% $CI[0.17, 0.41]$, $\beta=0.31$, $p<.001$	$B=0.29$, 95% $CI[0.17, 0.41]$, $\beta=0.31$, $p<.001$	$B=0.29$, 95% $CI[0.17, 0.41]$, $\beta=0.31$, $p<.001$
Diurnal slope changes predicting shifts in attitudes toward the elected leader.	$B=0.13$, 95% $CI[0.02, 0.23]$, $\beta=0.13$, $p=.019$	$B=0.13$, 95% $CI[0.02, 0.23]$, $\beta=0.13$, $p=.019$	$B=0.13$, 95% $CI[0.02, 0.23]$, $\beta=0.13$, $p=.019$

9.7. False Discovery Rate Analysis

To correct for multiple comparisons across our main research questions, we used Benjamini & Hochberg (1995)'s false discovery rate (FDR) method to obtain adjusted *p*-values. The *Q* (or false discovery rate) was set at .05. The corrections were applied to the eight results from our main models. We found that all our statistically significant results remained statistically significant after correcting for FDR (see Table S26).

Table S26. Correction for multiple comparisons across our main research questions using the Benjamini & Hochberg (1995) false discovery rate (FDR) method. Both the *p*-value and adjusted *p*-values are reported below.

Analyses	<i>p</i> -value	Adjusted <i>p</i> -value
Time (Linear) × Group (Win/Loss) predicting acute testosterone change	.013	.035
Time (Quadratic) × Group (Win/Loss) predicting acute testosterone change	.875	.875
Time (Linear) × Group × Acute testosterone change predicting shifts in attitudes toward the elected leader	5.01E-06	4.01E-05
Time (Quadratic) × Group × Acute testosterone change predicting shifts in self-reported attitudes toward the elected leader	.12	.16
Time (Linear) × Group (Win/Loss) predicting diurnal slope change	.004	.016
Time (Quadratic) × Group (Win/Loss) predicting diurnal slope change	.076	.122
Time (Linear) × Group × Diurnal slope change predicting shifts in self-reported attitudes toward the elected leader	.019	.038
Time (Quadratic) × Group × Diurnal slope change predicting shifts in self-reported attitudes toward the elected leader	.391	.447

9.8. Multiverse specification curve analysis

We made the following modeling decisions and assumptions in our statistical analysis strategy: First, we used MLMs to conduct our analyses, in which we modelled change over time with the inclusion of both linear and curvilinear (i.e., quadratic) parameters. Second, in our treatment of outliers, we winsorized values of testosterone that were 3 SDs above or below the mean (within sex and measurement instance) to the value at 3 SDs of the mean (within sex and measurement instance). Third, we made assumptions about heteroskedasticity and normality of residuals in our models. As part of the review process, we were encouraged to conduct a specification curve (or multiverse) analysis to examine the extent to which the robustness of our findings was dependent on the choices made in our analytic approach (Simonsohn et al., 2020).

To conduct a specification curve analysis we identified the following reasonable alternative approaches to our original data analytic approach: (i) the choice of the regression models, such as robust MLMs that correct for heteroskedasticity and non-normal errors, generalized linear models (GLM) that allow for non-normal (non-Gaussian) dependent variables, and variations in how slopes and intercepts were modelled (random versus non-random variation), (ii) the inclusion and exclusion of outliers (winsorized values are instead excluded), (iii) the inclusion of the main covariates, and (iv) alternative approaches to estimating diurnal slopes using GLMs (see Table S27). The specification curve analysis included all of the original specifications reported in the main paper and the alternative specifications, yielding a total of 72 specifications for the models examining acute change in testosterone, 48 specifications for the models examining acute testosterone changes predicting shifts in attitudes, 128 specifications for models examining changes in diurnal slopes, and 64 specifications for models examining changes in diurnal slopes predicting shifts in attitudes. An effect size was calculated by dividing each estimated coefficient by its own standard error. The median effect size from each of these four specification curves was used as an estimate of the given effect within the multiverse.

Table S27. Original specifications and alternative reasonable specifications used to test our main research questions. Note: Not all specifications applied to all models.

Decision	Original specifications for the main results reported in the paper	Alternative specifications
Analytic model	Multilevel regression models	Robust multilevel models GLM models to estimate testosterone change using a gamma distribution and identity link function
	Linear and Quadratic parameters of change	Linear parameters of change (only)
	Intercepts and slopes were both allowed to randomly vary*	Only intercepts were allowed to randomly vary
Treatment of outliers	Outliers were winsorized to 3SDs of the mean (within sex and time)	Outliers included in the analyses Outliers excluded from the analyses
Covariates	Covariates were not included	Models include covariates (Age, SES, and hours of sleep for the diurnal slope change model)
Diurnal slope calculation	Multilevel regression models	GLM using gamma distribution and identity link function

*The slopes were not allowed to randomly vary for those models that were overidentified (e.g., acute changes in testosterone across the competition outcome window were measured across three time points and so were limited to linear random effects).

9.8.1. Inferential statistics for the specification curve analysis. To draw inferences from our specification curves, we used the “under-the-null” resampling method for non-experimental data (Simonsohn et al., 2020). Estimating all our models with the observed data for all possible specifications resulted in K parameters of interest (b_k). The b_k for each specification were then used to generate K different null values of the dependent variable (y^* ; $y_k^* = y_k - b_k x_k$). Using this new matrix of data, $N = 113$ participants’ data (to match the original sample size) were drawn at random and with replacement. In other words, resampling was conducted at the level of the participant. The corresponding K specifications for each y^* model were estimated on the randomly generated data. These steps (resampling and estimation) were repeated 500 times for each specification curve. Finally, to draw inferences, we calculated a p -value as the proportion of the median effect sizes from the resampled null specification curves that were more extreme than that observed in the real data.

9.8.2. Results from the specification curve analysis. Inferential statistics for the specification curve analyses revealed that our main results were robust to decisions made as part of the analytic approach ($ps \leq .006$; see Table S28 and Figures S10-S13).⁸

Table S28. Results from the specification curve analyses across the various data analytic choices

	Number of Specifications	Median effect size*	Proportion of null models from resampled data with effect sizes that are as extreme as the median observed effect size	p
Acute change in testosterone	72	-2.26	0/500	<.001
Acute testosterone change predicting shifts in attitudes toward the elected leader	48	3.89	0/500	<.001
Changes in diurnal slopes	128	-2.66	3/500	.006
Changes in diurnal slopes predicting shifts in attitudes toward the elected leader	64	3.00	1/500	.002

* Effect size was determined by dividing each effect by its standard error.

⁸ Quality control checks of the null specification curves revealed two issues when relying on GLMs to analyze acute changes in testosterone. First, the choice of a gamma distribution and identity link function is known to cause estimation issues. Specifically, although the data were appropriate for a gamma distribution (i.e., data were distributed between 0 and infinity), the model can estimate values outside those bounds, causing the model to fail. Failed models ($n = 124$ of 36,000) were ignored for the purposes of determining inferential statistics. In follow-up analyses, we replaced these missing models with an extreme value (effect size = ± 5 , with sign of the replacement value determined by the sign of the median effect of the null specification curve) and re-ran the inferential statistics. Results were entirely unchanged. Second, convergence issues in the GLMs resulted in a set of models ($n = 60$ of 36,000) with extreme effect sizes (e.g., more extreme than ± 100). Poor convergence in these models produced normal estimates with exceedingly small standard errors. Relying on the median effect from each resampling effectively controls for these outliers. Further, an extreme value would only penalize our specification analyses, in terms of producing values from the null specification curves that may be more extreme than our observed effect sizes. Hence, these outliers were left in the null specification curves and the results of the inferential statistics are considered conservative estimates of the robustness of our results.

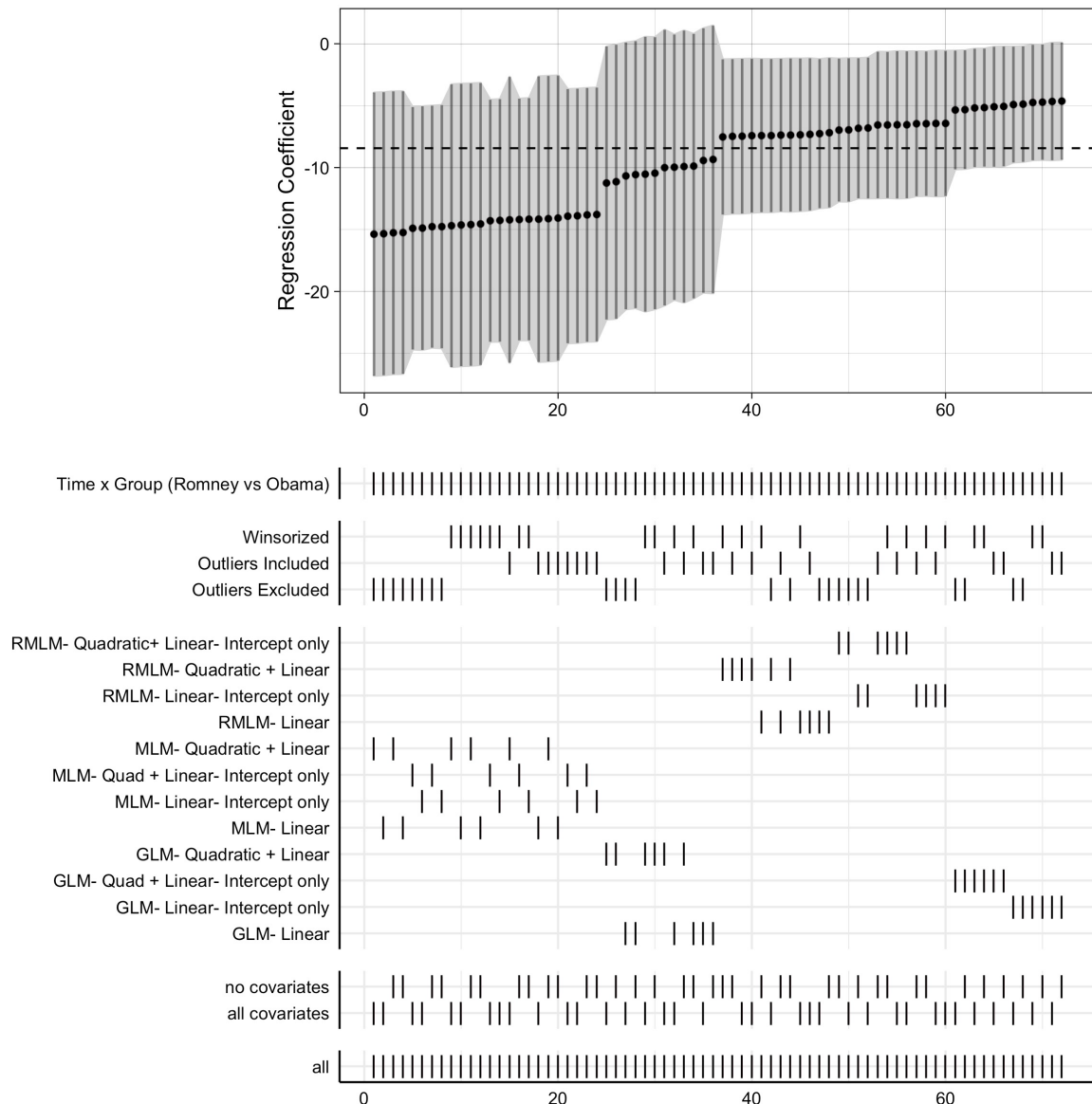


Fig. S10. Results of the specification curve analysis examining acute changes in testosterone across the competition-outcome window. **Upper Panel:** Black dots indicate regression coefficients for the Time (Linear) \times Political Group (Romney versus Obama) effect across various specifications. The grey ribbon represents 95% CIs. The dashed line indicates the median effect across all specifications. **Lower panel:** Tickers indicate the specifications corresponding to a particular model. **Key:** Testosterone- Winsorized: Testosterone concentrations that were winsorized to 3SDs from the mean (within sex and measurement instance); Testosterone- Outliers Included: Testosterone concentrations with outliers included; Testosterone- Outliers Excluded: Testosterone concentrations with outliers excluded; MLM: Multilevel model; RMLM: Robust Multilevel model; GLM: Generalized Multilevel Model; Quadratic + Linear: Quadratic and Linear slopes were included in the model and were allowed to randomly vary; Linear: Only linear slopes were included and were allowed to randomly vary; Quadratic + Linear- Intercept only: Linear and quadratic slopes were included in the model but only the intercept was allowed to randomly vary; Linear- Intercept only: Linear slopes were included in the model but only the intercept was allowed to randomly vary; All covariates: Age and SES were included as covariates.

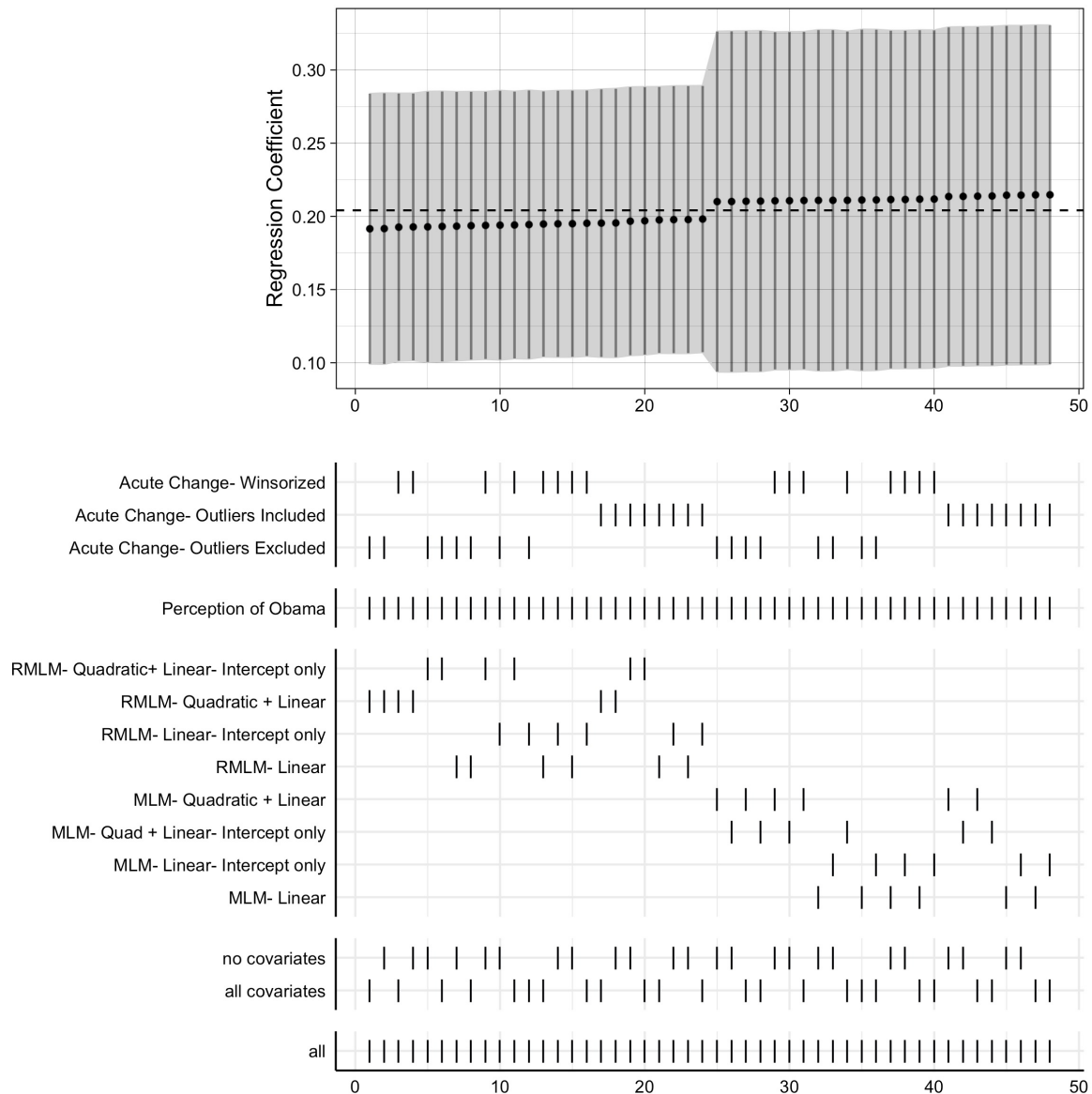


Fig. S11. Results of the specification curve analysis examining acute change in testosterone predicting shifts in attitudes toward the elected leader. **Upper Panel:** Black dots indicate regression coefficients for the Time (Linear) \times Political Group (Romney versus Obama) \times Acute testosterone change effect across various specifications. The grey ribbon represents 95% CIs. The dashed line indicates the median effect across all specifications. **Lower panel:** Tickers indicate the specifications corresponding to a particular model. **Key:** Acute Change-Winsorized: Acute change was calculated using testosterone concentrations that were winsorized to 3SDs from the mean (within sex and measurement instance); Acute Change-Outliers Included: Acute change was calculated using testosterone concentrations with outliers included; Acute Change- Outliers Excluded: Acute change was calculated using testosterone concentrations with outliers excluded; MLM: Multilevel model; RMLM: Robust Multilevel model; GLM: Generalized Multilevel Model; Quadratic + Linear: Quadratic and Linear slopes were included in the model and were allowed to randomly vary; Linear: Only linear slopes were included and were allowed to randomly vary; Quadratic + Linear- Intercept only: Linear and quadratic slopes were included in the model but only the intercept was allowed to randomly vary; Linear- Intercept only: Linear slopes were included in the model but only the intercept was allowed to randomly vary.

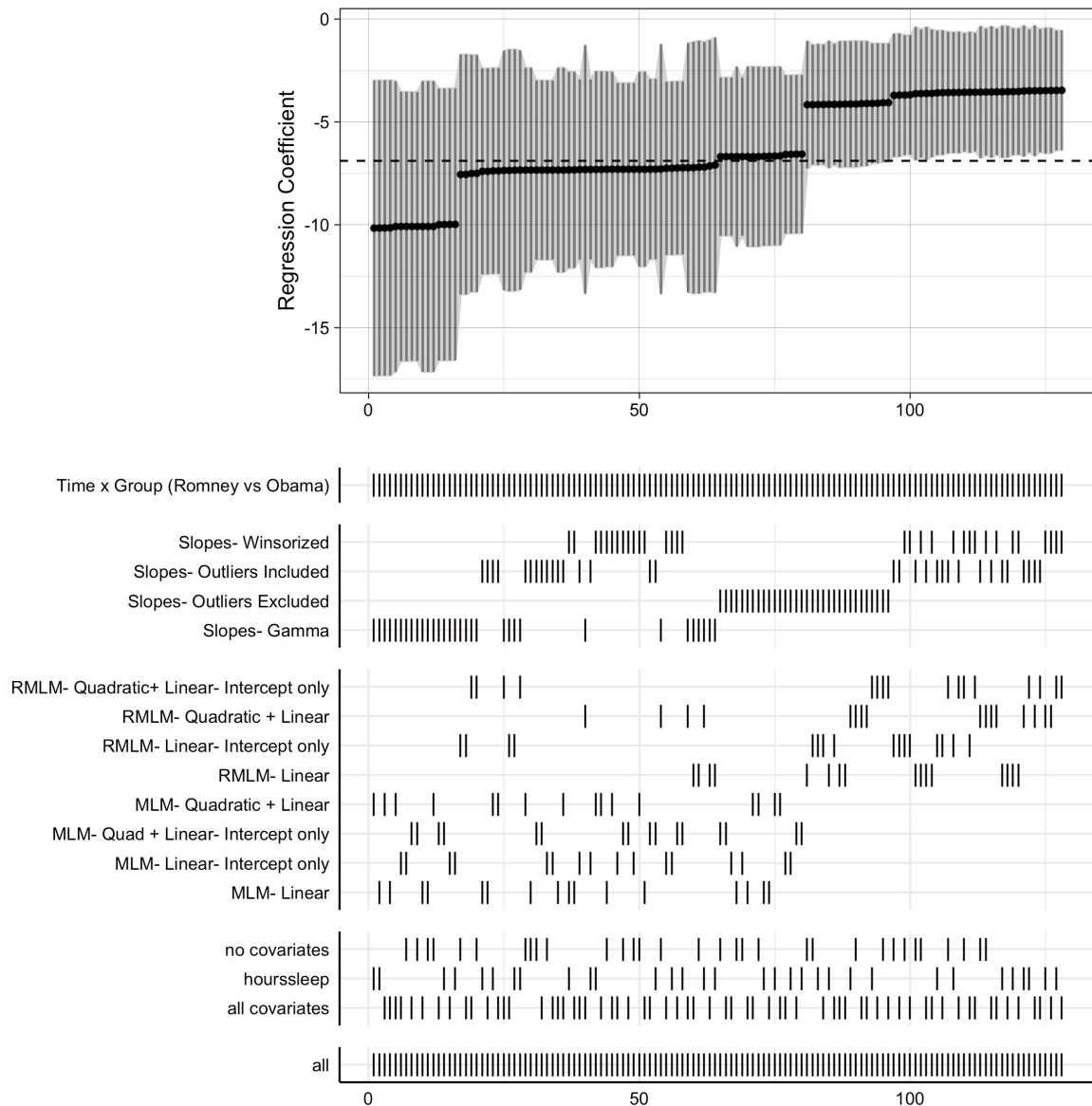


Fig. S12. Results of the specification curve analysis examining changes in diurnal slope across the naturalistic sampling period. **Upper Panel:** Black dots indicate regression coefficients for the Time (Linear) \times Political Group (Romney versus Obama) effect across various specifications. The grey ribbon represents 95% CIs. The dashed line indicates the median effect across all specifications. **Lower panel:** Tickers indicate the specifications corresponding to a particular model. **Key:** Slopes- Winsorized: Slopes were calculated using testosterone concentrations that were winsorized to 3SDs from the mean (within sex and measurement instance); Slopes- Outliers Included: Slopes were calculated using testosterone concentrations with outliers included; Slopes- Outliers Excluded: Slopes were calculated using testosterone concentrations with outliers excluded; Slopes- Gamma: Slopes were calculated using GLMs with a gamma distributions and the link=identity function; MLM: Multilevel model; RMLM: Robust Multilevel model; GLM: Generalized Multilevel Model; Quadratic + Linear: Quadratic and Linear slopes were included in the model and were allowed to randomly vary; Linear: Only linear slopes were included and were allowed to randomly vary; Quadratic + Linear- Intercept only: Linear and quadratic slopes were included in the model but only the intercept was allowed to randomly vary; Linear- Intercept only: Linear slopes were included in the model but only the intercept was allowed to randomly vary; Hourssleep: Hours of sleep as a covariate; All covariates: Age, SES, and hours of sleep were included as covariates.

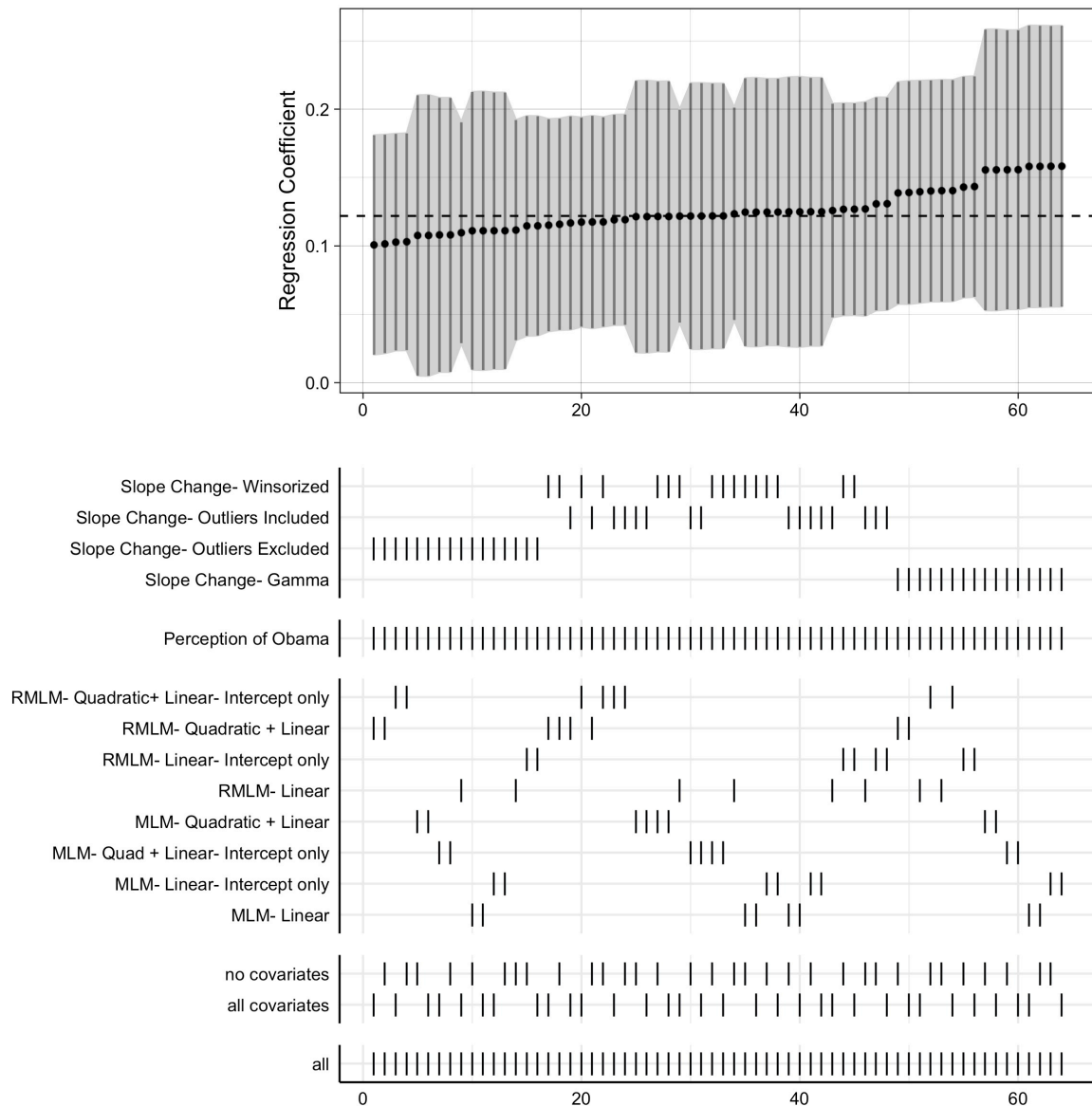


Fig. S13. Results of the specification curve analysis examining diurnal slope changes predicting shifts in attitudes toward the elected leader. **Upper Panel:** Black dots indicate regression coefficients for the Time (Linear) \times Political Group (Romney versus Obama) \times Diurnal slope change effect across various specifications. The grey ribbon represents 95% CIs. The dashed line indicates the median effect across all specifications. **Lower panel:** Tickers indicate the specifications corresponding to a particular model. **Key:** Slope Change-Winsorized: Slope change was calculated using slopes created with testosterone concentrations that were winsorized to 3SDs from the mean (within sex and measurement instance); Slope Change- Outliers Included: Slope change was calculated using slopes created from testosterone concentrations with outliers included; Slope Change- Outliers Excluded: Slope change was calculated using slopes created from testosterone concentrations with outliers excluded; Slope Change- Gamma: Slope change was calculated using slopes created via GLMs with a gamma distributions and link=identity function; MLM: Multilevel model; RMLM: Robust Multilevel model; GLM: Generalized Multilevel Model; Quadratic + Linear: Quadratic and Linear slopes were included in the model and were allowed to randomly vary; Linear: Only linear slopes were included and were allowed to randomly vary; Quadratic + Linear- Intercept only: Linear and quadratic slopes were included in the model but only the intercept was allowed to randomly vary; Linear- Intercept only: Linear slopes were included in the model but only the intercept was allowed to randomly vary.

9.9. Robustness checks for Supplementary Results

For our main analyses, we included comprehensive robustness checks with specification curve analysis. We also tested the robustness of our main supplemental results that we reported in the main manuscript by re-running all of these analyses with robust MLMs. Robust MLMs account for: (i) extreme datapoints by downweighting their influence and (ii) possible violations of assumptions of multilevel regressions (e.g. heteroskedasticity, Field & Wilcox, 2017). Our supplemental results were robust to the use of an alternative data analytic method.

Table S29. Results from robust MLMs for the following supplementary results: (i) testosterone change predicting long-term shifts in attitudes; (ii) testosterone change predicting shifts in composite attitudes (measured via self-reports and social media), and (iii) sex differences and similarities across all our main results.

	MLM (reported earlier in the supplement)	Robust MLM
Testosterone change predicting long-term shifts in attitudes		
Time (Linear) × <i>Acute Testosterone Reactivity</i> × Group (Romney versus Obama)	$B=0.17, 95\% CI[-0.04, 0.37], \beta=.19, p=.124$	$B=0.17, 95\% CI[-0.02, 0.36], \beta=.20, p=.085$
Time (Linear) × <i>Diurnal Slope Change</i> × Group (Romney versus Obama)	$B=0.11, 95\% CI[-0.08, 0.29], \beta=.11, p=.260$	$B=0.10, 95\% CI[-0.07, 0.27], \beta=.11, p=.242$
Testosterone change predicting shifts in composite attitudes (self-reports and social media)		
Time (Linear) × <i>Acute Testosterone Reactivity</i> × Group (Romney versus Obama)	$B=0.17, 95\% CI[0.05, 0.29], \beta=.19, p=.008$	$B=0.18, 95\% CI[0.07, 0.28], \beta=.20, p=.001$
Time (Linear) × <i>Diurnal Slope Change</i> × Group (Romney versus Obama)	$B=0.10, 95\% CI[-0.01, 0.20], \beta=.10, p=.066$	$B=0.09, 95\% CI[>0.00, 0.18], \beta=.09, p=.049$
Sex differences and similarities		
<i>Acute testosterone change:</i> Sex × Time (Linear) × Group (Romney versus Obama)	$B=-26.37, 95\% CI[-48.69, -4.04], \beta=-0.59, p=.023$ <i>Males:</i> $B=-28.63, 95\% CI[-51.57, -5.69], \beta=-0.63, p=.018$ <i>Females:</i> $B=-2.12, 95\% CI[-9.85, 5.62], \beta=-0.11, p=.594$	$B=-15.34, 95\% CI[-26.97, -3.72], \beta=-0.34, p=.010$; <i>Males:</i> $B=-20.72, 95\% CI[-36.84, -4.60], \beta=-0.45, p=.012$ <i>Females:</i> $B=-2.98, 95\% CI[-8.56, 2.60], \beta=-0.15, p=.296$.

<i>Acute testosterone reactivity predicting shifts in attitudes:</i> Sex × Time (Linear) × Group (Romney versus Obama) × Acute testosterone change	Short-term shifts in attitudes: $B = -0.18$, 95% CI[-0.43, 0.06], $\beta = -0.20$, $p = .141$ Long-term shifts in attitudes: $B = 0.12$, 95% CI[-0.34, 0.59], $\beta = 0.15$, $p = .603$ Composite measure of attitudes via self-reports and Facebook: $B = -0.02$, 95% CI[-0.27, 0.24], $\beta = -0.02$, $p = .897$	Short-term shifts in attitudes: $B = -0.09$, 95% CI[-0.29, 0.11], $\beta = -0.10$, $p = .382$ Long-term shifts in attitudes: $B = 0.18$, 95% CI[-0.28, 0.64], $\beta = 0.21$, $p = .437$ Composite measure of attitudes via self-reports and Facebook: $B = -0.04$, 95% CI[-0.26, 0.18], $\beta = -0.04$, $p = .745$
<i>Diurnal slope change:</i> Sex × Time (Linear) × Group (Romney versus Obama)	$B = -12.55$, 95% CI[-22.01, -3.09], $\beta = -0.65$, $p = .010$ Males: $B = -13.39$, 95% CI[-23.21, -3.56], $\beta = -0.76$, $p = .010$ Females: $B = -0.83$, 95% CI[-4.08, 2.41], $\beta = -.09$, $p = .615$	$B = -6.59$, 95% CI[-12.84, -0.34], $\beta = -0.34$, $p = .039$ Males: $B = -9.35$, 95% CI[-16.59, -2.10], $\beta = -0.53$, $p = .011$ Females: $B = -0.95$, 95% CI[-3.86, 1.96], $\beta = -.11$, $p = .522$
<i>Diurnal slope change predicting shifts in attitudes:</i> Sex × Time (Linear) × Group (Romney versus Obama) × Diurnal slope change:	Short-term shifts in attitudes: $B = -0.01$, 95% CI[-0.22, 0.21], $\beta = -0.01$, $p = .958$ Long-term shifts in attitudes: $B = -0.25$, 95% CI[-0.62, 0.12], $\beta = -0.26$, $p = .188$ Composite measure of attitudes via self-reports and Facebook: $B = -0.06$, 95% CI[-0.27, 0.15], $\beta = -.07$, $p = .559$	Short-term shifts in attitudes: $B = <0.00$, 95% CI[-0.16, 0.16], $\beta = <0.00$, $p = .986$ Long-term shifts in attitudes: $B = -0.25$, 95% CI[-0.59, 0.10], $\beta = -0.26$, $p = .168$ Composite measure of attitudes via self-reports and Facebook: $B = -0.06$, 95% CI[-0.25, 0.13], $\beta = -.06$, $p = .549$.

11. Comparison between the current study and a study conducted during the 2008 US Presidential Election (Stanton et al., 2009)

Table S30. Differences and similarities between the research questions, methods, and results in the current study (conducted during the 2012 election) and a study during the 2008 election (Stanton et al., 2009).

	Stanton et al., 2009 (2008 US Presidential Election)	Current study (2012 US Presidential Election)
Primary research questions	Does the electoral outcome predict acute testosterone changes in supporters of the competing candidates on the night of the election? Are there sex differences in testosterone responses to the election outcome?	Does the electoral outcome predict acute testosterone changes in supporters of the competing candidates on the night of the election? Do acute changes in testosterone in supporters of the candidates predict shifts in attitudes toward the elected leader?
Secondary research questions	None reported.	Does the electoral outcome predict changes in diurnal rhythms in supporters of the competing candidates, from two days before the election until two days after the election? Do diurnal slope changes in supporters of the candidates facilitate shifts in attitudes toward the elected leader?
Study duration	Study was conducted during the night of the election, spanning from 8 pm to 40 minutes after the election results were declared.	5-day naturalistic sampling period (from 2 days before the election, Election Day, until two days after the election). Included a 6-month follow-up wherein participants reported attitudes toward the elected leader.
Sample size	163 voters (65% Female)	113 registered voters (57.52% female)
Geographical study location	Durham, North Carolina Ann Arbor, Michigan	Eugene, Oregon
Research study location	At-home	At-home
Political groups composition	Obama supporters (winners), McCain supporters (losers), Robert Barr supporters (losers)	Obama supporters (winners), Romney supporters (losers), and Undecided voters
Testosterone: Collection method	Stick of sugar-free chewing gum + Passive drool.	Passive drool without stimulant

Testosterone: Sample storage	Participants were asked to store samples in a refrigerator overnight and bring it into the lab the next day where they were stored in a laboratory freezer.	Participants were asked store all their samples in a freezer at home immediately after collection. If participants did not have access to a freezer, they were asked to store their samples temporarily in a box with icepacks (which were provided by the experimenters), before they transferred the samples to a freezer at home. At the end of the study, when participants dropped off their frozen samples, the samples were immediately stored at -80 °C
Testosterone: Compliance	Participants recorded sample times.	Participants received text messages prior to the times they were required to provide a saliva sample and were also asked to electronically timestamp their sample times.
Testosterone: Assay method	Radioimmunoassay (Diagnostic Products Corporation)	Enzyme-immunoassays (Salimetrics) Testosterone concentrations from a random subset of samples were validated against a highly accurate reference method- LC-MS/MS.
Testosterone: Measurement instances	Participants provided a baseline saliva sample at 8 pm EST, a time at which many election polls were closing on the east coast of the United States. Participants provided three other samples at 0, 20, and 40 minutes after they had learned that Barack Obama had been declared the winner.	Saliva samples were collected on each of the five days in the naturalistic sampling period when participants woke up (<i>Wake-up</i>), 30 minutes after they woke up (<i>Wake-up+30 mins</i>), at 3pm (Afternoon), and when they went to bed (<i>Bedtime</i>). Additional samples were collected at 5pm, 7pm, and 9pm on Election Day (<i>ED</i>). The period from 7pm to <i>Bedtime</i> was determined as the competition-outcome window because, on average, participants reported learning about the final outcome of the election at 8:50 pm PST (networks called the election approximately at 8:12 pm PST). Diurnal rhythms were indexed by the magnitude of the slope connecting the morning peak and awakening response in testosterone (via the wake-up samples), the afternoon plateau (via the 3 pm sample), and the end-of-day nadir (via the Bedtime sample).
Political attitudes	None reported.	Attitudes toward the elected leader were assessed via scales adapted from polling surveys that resemble presidential approval ratings.

		Social media activity (via Facebook) was coded to assess attitudes toward the elected leader, in a subset of participants with usable data (reported in the supplement).
Affect: measurement	Retrospective affective state questionnaire: Self-reported feelings of pleasantness (unpleasant to pleasant; unhappy to happy) and dominance (dominant to submissive; controlled to controlling) at the moment when Barack Obama was declared the winner.	In-the-moment affect was measured using daily diaries on Election Day at the following times: 30 minutes after they woke up, in the Afternoon, at 5pm, 7pm, 9pm, and at Bedtime. Positive affect was measured using a composite of the following items: happy, excited, powerful, and confident. Negative affect was measured using a single-item measure of sadness.
Statistical analysis technique	Analysis of Covariance (ANCOVA)	Multilevel models (MLMs)
Main Results:	Male voters for winning candidate had stable post-outcome levels of testosterone, and male voters for the losing candidate (losers) showed decrements in their testosterone levels.	<p>Supporters of the losing candidate experienced acute elevations in testosterone concentrations on the night of the election compared to supporters of the winner.</p> <p>Supporters of the losing candidate also exhibited flatter diurnal slopes starting on Election Day, and these flatter slopes persisted up to two days after the election.</p> <p>Among the losing candidate's supporters, greater testosterone reactivity on election-night and flatter diurnal slopes predicted less positive attitudes toward the winning candidate after the election.</p>
Supplemental results	None reported.	Testosterone-moderated attitudinal shifts were smaller in magnitude six months after the election, albeit in the same direction as the short-term shifts in attitudes.
Robustness checks	<p>Results were robust to:</p> <ol style="list-style-type: none"> 1) Inclusion of the following covariates in the model: <ol style="list-style-type: none"> a) Sociability variables: alcohol consumption on the night of the election, where participants viewed the election results (e.g., home, bar, campus hall), with how many others they viewed the election results b) Political investment: Participants' candidate support intensity and participants' estimation of their candidate's likelihood of winning c) Right-wing authoritarianism (RWA) scale 	<p>Our results were robust to:</p> <ol style="list-style-type: none"> 1) Use of a composite measure of attitudes toward the elected leader that included self-reports and coded behaviors on social media (i.e., Facebook). 2) Inclusion of covariates: <ol style="list-style-type: none"> a) Age and SES b) Changes in affect during the competition-outcome window c) When participants learned about the election outcome, and when they completed the Pre- and Post-Election

	<p>d) Timing of post-outcome saliva collection on testosterone change</p> <p>2) Exclusion of Robert Barr supporters.</p>	<p>Lab sessions (during which they self-reported attitudes toward the elected leader)</p> <p>3) Exclusion of a Gary Johnson supporter.</p> <p>4) False Discovery Rate (FDR) corrections for multiple comparisons</p> <p>5) Alternative data analytic choices, as assessed via specification curve analysis.</p>
Sex differences	Competition-related acute changes in testosterone were evident in males and not in females.	<p>Competition-related acute and diurnal changes in testosterone were larger in males than females (though the effects were in the same direction for both sexes).</p> <p>We did not find evidence of sex differences in the relationships between testosterone changes (acute and diurnal) and shifts in attitudes toward the elected leader.</p>
Results with affect	Upon the announcement of Obama as the president-elect, McCain and Barr voters retrospectively reported feeling significantly more unhappy, submissive, unpleasant, and controlled than Obama voters.	Obama supporters reported increases in positive affect and no change in negative affect on Election Day. Romney supporters reported decreases in positive affect and increases in negative affect on Election Day.

Additional Acknowledgements: We are indebted to our army of research assistants for their help with this project: Leo Edmunson, Corrine Berryman, Nikki Gonzales, Heidi Hanst, Andrew Keller, Eliza Chetlin, Jason Bagley, Jessica Glaser, Rain Baker, Nick Snyder, Alexander Wagon, Brianna Cline, Daniel Stone, Adam Tratner, Andrew Swift, Kayla Bucolo, Leslie Coray, Manmeet Kals, Matthew Buchholz, Ashlin Roy, Devan Pennington, Oliver Medak, Alexa Jasmer, Breanna Cunningham, James Blizzard, Jennifer Owen, Kaysea Dahlstrom, Kristi Jacobs, Molly Smith, and Qiang Gu.

SI REFERENCES

- Adam, E.K., Kumari, M., 2009. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2009.06.011>
- Adler, N.E., Epel, E.S., Castellazzo, G., Ickovics, J.R., 2000. Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Heal. Psychol.* <https://doi.org/10.1037/0278-6133.19.6.586>
- Alabastro, A., Rast, D.E., Lac, A., Hogg, M.A., Crano, W.D., 2013. Intergroup bias and perceived similarity: Effects of successes and failures on support for in- and outgroup political leaders. *Gr. Process. Intergr. Relations*. <https://doi.org/10.1177/1368430212437212>
- Ariens, C. 2016. Here's What Time the Networks Called the 2012 Election. URL: <http://www.adweek.com/tvnewser/heres-what-time-the-networks-called-the-2012-election/310119>.
- Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v067.i01>
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B*. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Carlisle, J.E., Patton, R.C., 2013. Is Social Media Changing How We Understand Political Engagement? An Analysis of Facebook and the 2008 Presidential Election. *Polit. Res. Q.* <https://doi.org/10.1177/1065912913482758>
- Carré, J.M., Putnam, S.K., McCormick, C.M., 2009. Testosterone responses to competition predict future aggressive behaviour at a cost to reward in men. *Psychoneuroendocrinology*. <https://doi.org/10.1016/j.psyneuen.2008.10.018>
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Erlbaum.
- Crano, W.D., Prislin, R., 2011. Attitudes and attitude change, *Attitudes and Attitude Change*. <https://doi.org/10.4324/9780203838068>
- Fernandes, A., Skinner, M.L., Woelfel, T., Carpenter, T., Haggerty, K.P., 2013. Implementing Self-Collection of Biological Specimens with a Diverse Sample. *Field methods*. <https://doi.org/10.1177/1525822X12453526>
- Field, A.P., Wilcox, R.R., 2017. Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behav. Res. Ther.* <https://doi.org/10.1016/j.brat.2017.05.013>

- Fox, J., Friendly, M., & Monette, G. 2018. heplots: Visualizing Tests in Multivariate Linear Models. R package version 1.3-5. URL <https://CRAN.R-project.org/package=heplots>
- Gallup Inc. 2012a. Presidential Ratings -- Issues Approval. URL: <http://www.gallup.com/poll/1726/Presidential-Ratings-Issues-Approval.aspx>.
- Gallup Inc. 2012b. Presidential Ratings -- Personal Characteristics. Gallup. URL: <http://www.gallup.com/poll/1726/Presidential-Ratings-Issues-Approval.aspx>.
- Gelman, A., 2008. Scaling regression inputs by dividing by two standard deviations. Stat. Med. <https://doi.org/10.1002/sim.3107>
- Geniole, S.N., Bird, B.M., Ruddick, E.L., Carré, J.M., 2017. Effects of competition outcome on testosterone concentrations in humans: An updated meta-analysis. Horm. Behav. <https://doi.org/10.1016/j.yhbeh.2016.10.002>
- Gil de Zúñiga, H., 2012. Social Media Use for News and Individuals' Social Capital, Civic Engagement and Political Participation. J. Comput. Commun. <https://doi.org/10.1111/j.1083-6101.2012.01574.x>
- Green, P., & MacLeod, C. J. 2016. SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution, 7(4), 493-498.
- Hatemi, P.K., Verhulst, B., 2015. Political attitudes develop independently of personality traits. PLoS One. <https://doi.org/10.1371/journal.pone.0118106>
- Hoffman, L., Rovine, M.J., 2007. Multilevel models for the experimental psychologist: Foundations and illustrative examples. Behav. Res. Methods. <https://doi.org/10.3758/BF03192848>
- Hughes, J. 2020. reghelper: Helper Functions for Regression Analysis. R package version 0.3.6. <https://CRAN.R-project.org/package=reghelper>
- Jost, J.T., West, T. V., Gosling, S.D., 2009. Personality and ideology as determinants of candidate preferences and "obama conversion" in the 2008 U.S. presidential election. Du Bois Rev. <https://doi.org/10.1017/S1742058X09090109>
- Kim, T., Atkin, D.J., Lin, C.A., 2016. The Influence of Social Networking Sites on Political Behavior: Modeling Political Involvement via Online and Offline Activity. J. Broadcast. Electron. Media. <https://doi.org/10.1080/08838151.2015.1127242>
- Koller, M. 2016. robustlmm: An R Package for Robust Estimation of Linear Mixed-Effects Models. doi:10.18637/jss.v075.i06
- Kriner, D., Schwartz, L., 2009. Partisan dynamics and the volatility of presidential approval. Br. J. Polit. Sci. <https://doi.org/10.1017/S0007123409000647>

- Krosnick, J.A., 1991. The Stability of Political Preferences: Comparisons of Symbolic and Nonsymbolic Attitudes. *Am. J. Pol. Sci.* <https://doi.org/10.2307/2111553>
- Kudielka, B.M., Gierens, A., Hellhammer, D.H., Wüst, S., Schlotz, W., 2012. Salivary cortisol in ambulatory assessment-some dos, some don'ts, and some open questions. *Psychosom. Med.* <https://doi.org/10.1097/PSY.0b013e31825434c7>
- Kwok, O.M., Underhill, A.T., Berry, J.W., Luo, W., Elliott, T.R., Yoon, M., 2008. Analyzing Longitudinal Data With Multilevel Models: An Example With Individuals Living With Lower Extremity Intra-Articular Fractures. *Rehabil. Psychol.* <https://doi.org/10.1037/a0012765>
- Lebo, M.J., Cassino, D., 2007. The aggregated consequences of motivated reasoning and the dynamics of partisan presidential approval. *Polit. Psychol.* <https://doi.org/10.1111/j.1467-9221.2007.00601.x>
- Lenhard, W. & Lenhard, A. 2016. Calculation of Effect Sizes. Retrieved from: https://www.psychometrica.de/effect_size.html. Dettelbach (Germany): Psychometrica. <https://doi.org/10.13140/RG.2.2.17823.92329>
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. 2018. Package 'emmeans'. Estimated marginal means, aka least-squares means. R Package Version 1.15-15.
- Locascio, J.J., Atri, A., 2011. An Overview of Longitudinal Data Analysis Methods for Neurological Research. *Dement. Geriatr. Cogn. Dis. Extra.* <https://doi.org/10.1159/000330228>
- Lüdtke D. 2020. sjPlot: Data Visualization for Statistics in Social Science. doi: [10.5281/zenodo.1308157](https://doi.org/10.5281/zenodo.1308157), R package version 2.8.3, <https://CRAN.R-project.org/package=sjPlot>.
- Maas, C.J.M., Hox, J.J., 2005. Sufficient sample sizes for multilevel modeling. *Methodology.* <https://doi.org/10.1027/1614-2241.1.3.86>
- Martin, R.F., 2000. General deming regression for estimating systematic bias and its confidence interval in method-comparison studies. *Clin. Chem.* <https://doi.org/10.1093/clinchem/46.1.100>
- Mehta, P.H., Josephs, R.A., 2006. Testosterone change after losing predicts the decision to compete again. *Horm. Behav.* <https://doi.org/10.1016/j.yhbeh.2006.07.001>
- O'Connor, B., Balasubramanyan, R., Routledge, B.R., Smith, N.A., 2010. From tweets to polls: Linking text sentiment to public opinion time series, in: ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media.
- Prasad, S., Lassetter, B., Welker, K.M., Mehta, P.H., 2019. Unstable correspondence between salivary testosterone measured with enzyme immunoassays and tandem mass

- spectrometry. *Psychoneuroendocrinology*.
<https://doi.org/10.1016/j.psyneuen.2019.104373>
- Raudenbush, S. W., & Bryk, A. S. 2002. *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- Schad, D.J., Vasishth, S., Hohenstein, S., Kliegl, R., 2020. How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *J. Mem. Lang.*
<https://doi.org/10.1016/j.jml.2019.104038>
- Schultheiss, O. C., & Stanton, S. J. 2009. Assessment of salivary hormones. In Harmon-Jones, E. & Beer, J. (Eds.) *Methods in Social Neuroscience*. Guilford Publications, New York.
- Schultheiss, O.C., Dlugash, G., & Mehta, P. H. 2019. Hormone measurement in social neuroendocrinology: A comparison of immunoassay and mass spectrometry methods. In O.C. Schultheiss and P.H. Mehta (Eds.). *The International Handbook of Social Neuroendocrinology*. Routledge Press.
- Simonsohn, U., Simmons, J.P., Nelson, L.D., 2020. Specification curve analysis. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-020-0912-z>
- Stanton, S.J., Beehner, J.C., Saini, E.K., Kuhn, C.M., LaBar, K.S., 2009. Dominance, politics, and physiology: Voters' testosterone changes on the night of the 2008 United States presidential election. *PLoS One*. <https://doi.org/10.1371/journal.pone.0007543>
- Steele, F., 2008. Multilevel models for longitudinal data. *J. R. Stat. Soc. Ser. A Stat. Soc.*
<https://doi.org/10.1111/j.1467-985X.2007.00509.x>
- Van Anders, S.M., Goldey, K.L., Bell, S.N., 2014. Measurement of testosterone in human sexuality research: Methodological considerations. *Arch. Sex. Behav.*
<https://doi.org/10.1007/s10508-013-0123-z>
- Welker, K.M., Lassetter, B., Brandes, C.M., Prasad, S., Koop, D.R., Mehta, P.H., 2016. A comparison of salivary testosterone measurement using immunoassays and tandem mass spectrometry. *Psychoneuroendocrinology*.
<https://doi.org/10.1016/j.psyneuen.2016.05.022>
- Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
- Young, D.G., 2004. Late-Night Comedy in Election 2000: Its Influence on Candidate Trait Ratings and the Moderating Effects of Political Knowledge and Partisanship. *J. Broadcast. Electron. Media*. <https://doi.org/10.1207/s15506878jobem4801>
- Zilioli, S., Mehta, P.H., Watson, N. V., 2014. Losing the battle but winning the war: Uncertain outcomes reverse the usual effect of winning on testosterone. *Biol. Psychol.*
<https://doi.org/10.1016/j.biopsycho.2014.07.022>