Dirty Looks: Politicians' facial appearance and unethical behaviour

12,449 Words Including References

Politician Appearance and Behaviour

Abstract

Over half of British Members of Parliament (MPs) were found to have overclaimed on their expenses in the 2009 expenses scandal. We conduct an exploratory analysis of whether the facial appearance of the MPs (*N*=636) is associated with overclaiming, as research has found that facial appearance is correlated with behavioural outcomes. Participants (*N*=4,727) previously unfamiliar with the MPs made trait ratings (physically attractive, charismatic, criminal, competent, financially greedy, honest, likeable, organized, physically dominant, and sincere) of each politician's face. The latent factor structure indicated the traits could be grouped into the three broad factors identified in previous work: criminality (the traits *criminal, financially greedy*, and *physically dominant*), attractiveness (the traits *physically attractive, charismatic, honest, likeable,* and *sincere*), and competence (*competent* and *organized*). We found more attractive MPs tended to overclaim less, as did more criminal-looking MPs. But more competent-appearing politicians tended to overclaim more. We relate these findings to theories of moral licensing and moral consistency and discuss the limitations and context-specific nature of our findings.

Dirty Looks: Politicians' facial appearance and unethical behaviour

Faces are a rich source of socially-relevant information, enabling observers to make relatively accurate judgments about socially important dimensions, such as sex, age, and race (see Bruce & Young, 2012). It is well known that facial appearance is important for selecting leaders, such as politicians and CEOs (Poutvaara, 2014). But does facial appearance signal the ability and behaviour of the leaders who are selected?

Some research suggests that facial appearance predicts leadership performance (e.g., Haselhuhn & Wong, 2012, Rule & Ambady, 2008, 2009). There are several explanations for such associations. The associations may arise from genetically-determined and evolutionarily-desirable physical traits (e.g., higher levels of testosterone) being associated with psychological traits that underpin effective leadership in domains such as business and government (e.g., Bendahan, Zehnder, Pralong, & Antonakis, 2015; van Vugt, Hogan, & Kaiser, 2008). Alternatively, the association between facial appearance and behaviour could arise from the expectations that observers have of individuals based on their facial appearance (Haselhuhn, Wong, & Ormiston, 2013). Still another possible explanation is that evaluations of behaviour made by observers are influenced by facial appearance (e.g., Berry & Zebrowitz-McArthur, 1988).

To rule out such effects, some research has focussed on contexts where objective measures of effective leadership outcomes are available (e.g. Rule & Ambady, 2008). But, associations between these outcomes and facial appearance have turned out to be spurious and not reflective of causal relationships. For instance, while CEOs are better paid than their counterparts if they have more competent-looking faces (Graham, Harvey, & Puri, 2015) and CEOs look different from the population at large (Stoker, Garretsen, and Spreeuwers, 2016), there is no relationship between CEO performance and their facial appearance when firm performance is measured appropriately (Graham, Harvey, & Puri, 2015; Stoker, Garretsen, and Spreeuwers, 2016).

A major issue with the literature on facial appearance and behaviour is that most investigations do not employ objective real-world behavioural measures of leadership performance, focussing instead on outcomes that may not be directly related to the performance of the CEO, which severely limits causal inference. For instance, although more dominant-looking individuals in the military are more likely to be

promoted than their counterparts (Mazur et al., 1984; Mueller et al. 1996, 1997, Loehr et al. 2013), there is no objective evidence that dominant individuals behave more effectively as leaders, and therefore, are more deserving of promotion. The apparent association between facial appearance and behaviour may be driven by the effect of appearance on the perceivers who make judgements about leadership performance. In the context of CEOs, we do not know whether the lack of evidence for an effect of appearance on outcomes (Graham, Harvey, & Puri, 2015; Stoker, Garretsen, and Spreeuwers, 2016) reflects the lack of any relationship between leaders' facial appearance and their behaviour, or instead reflects a lack of measurement validity in capturing leaders' behaviour. In the present study, we overcome the limitations of previous research by studying a context in which we observe an objective measure of behaviour. In particular, we ask whether the facial appearance of politicians is associated with an objective measure of unethical behaviour.

Only one previous study has sought to examine the relationship between politicians' facial appearance and unethical behaviour (Lin, Adolphs, & Alvarez, 2018). The research found that the facial appearance of officials convicted of political corruption, and those found to violate campaign finance laws, was perceived by naïve raters to look more corruptible, dishonest, selfish, and aggressive, as well as more competent, ambitious, and masculine. Further, the facial width of the politicians was positively associated with naïve raters' evaluations of how corruptible the officials appeared. As tentative explanations for their findings, the researchers proposed that officials with more corruptible-looking faces might be approached more often by those seeking to corrupt public officials, or perhaps facial appearance exerts its influence on behaviour through self-fulfilling prophecy effects. However, it is also possible that the officials' facial appearance is unrelated to their corruptibility; rather, fact finders might be more likely to find officials guilty if their faces appear more corruptible-looking, an explanation that the study could not rule out. The present study addresses this limitation by analysing overclaiming, a measure of unethical behaviour that is not influenced by outside observers' perception of the politician's appearance.

We conducted our study in the context of the British parliamentary expenses scandal of 2009. The scandal involved the unveiling of the previously secret expense claims of all British parliamentarians, which revealed both petty corruption and grand larceny. The scandal affords a unique opportunity to examine whether facial appearance suggests a 'kernel of truth' about the parliamentarians' behaviour. We

operationalize unethical behaviour by using an objective measure of how much politicians overclaimed. Our aim is to test whether the facial appearance of the politicians is associated with whether and how much they overclaimed on their expenses. The sample of politicians is relatively large, which enabled us to analyse the relationship between unethical behaviour and several facial appearance factors that have been previously studied in the literature. Unethical behaviour by elected officials is important not only as a matter of morality, but also because political corruption is an important long-term determinant of living standards (Aidt, 2009). In the next section, we review the literature on politicians' facial appearance, selection, and behaviour, and especially highlight research suggesting that there might be a link between facial appearance and unethical behaviour in government officials (Bendahan et al., 2015. Haselhuhn & Wong, 2012, Lin, Adolphs, & Alvaraz, 2018; Wang et al., 2019).

Politicians' Facial Appearance and Selection

In the literature, the effect of facial appearance on outcomes operates either in terms of selection or behaviour (Antonakis & Eubanks, 2017; Todorov et al., 2015). Loosely, a selection effect is when people are treated differently by others based on their facial appearance, whereas a behaviour effect is when people's facial appearance is related to their behaviour. Most research on politicians conducted to date has focused on selection rather than behaviour effects.

Selection effects encompass the idea that voters might make choices based on facial appearance, even if it is uninformative about behaviour. Olivola and Todorov (2010a) reanalysed previously published data from several papers about how voters respond to politicians' facial appearance, highlighting traits that predict election outcomes. They found voters' preferences were accounted for by several facial appearance factors, particularly how physically attractive, likeable, and competent the politicians' faces appeared to naïve raters. Competence was found to be a key factor associated with electoral outcomes, even after controlling for attractiveness (Olivola & Todorov, 2010a). Competence assessments were also associated with face-based evaluations of trustworthiness, and how organised, dependable and emotionally stable the politician appeared. Politicians who appear more competent also tend to be rated as more attractive, more mature looking (Olivola & Todorov, 2010a) and more dominant (Chiao et al., 2008). There is some evidence nevertheless that baby-faced politicians are more likely to be elected, perhaps because the electorate views them as more approachable (Poutvaara, Jordahl, & Berggren, 2009).

Politician Appearance and Behaviour

The association between facial appearance and the selection of politicians may be accounted for by the *attractiveness-halo effect*, which is the notion that people believe 'what is beautiful is good' (Dion, Berscheid, & Walster, 1972; Langlois et al., 2000). Scholars have invoked this notion to explain why more attractive-looking politicians typically win more votes (Banducci, Karp, Thrasher, & Rallings 2008, Berggren, Jordahl, & Poutvaara, 2010; Berggren, Jordahl, & Poutvaara, 2017; King & Leigh, 2009; Lawson, Lenz, Baker, & Myers, 2010; Rosar, Klein & Beckers, 2008).

Selection effects occur in a range of other contexts besides politics. For example, in the realm of dating, success is more likely for men who appear outgoing and fun, while it is less likely for women who appear serious and smart (Olivola, Eastwick, Finkel, Ariely, & Todorov, 2011). In business, facial appearance is related to evaluations of leadership traits, including charisma (Awamleh & Gardner, 1999; Shea & Howell, 1999), trustworthiness (Krumhuber, Manstead, Cosker, Marshall, & Rosin, 2007), and criminality (e.g., Keating, Mazur, & Segall, 1977). Olivola, Eubanks and Lovelace's (2014) findings underline the importance of context in understanding selection effects. They found that having a physically attractive face conveying competence is associated with being a leader in the business domain, but not in a military or sports domain.

As several scholars have noted (c.f. Zebrowtiz & Montepare, 2008), there can also be a darker side to selection effects and face-based social attributions. For example, in the criminal justice context, baby-faced defendants receive less severe punishments (Berry & Zebrowitz-McArthur, 1988), and obtain better outcomes in small-claims litigation (Zebrowitz & McDonald, 1991). Individuals with more Afrocentric features receive harsher sentences (Blair, Judd, & Chapleau, 2004) and are more likely to receive the death penalty (Eberhardt, Davies, & Purdie-Vaughns, 2006). People are more likely to be identified from criminal lineups if they appear more criminal looking (Flowe & Humphries, 2011), and women with more masculine faces are more likely to be perceived as guilty of a crime compared to their counterparts (Flowe, Klatt, & Colloff, 2014; Ward, Flowe, & Humphries, 2012).

Taken together, it is well established in politics and other domains that social outcomes are associated with face-based appearance judgements. The findings raise the question of whether the facial appearance factors that are associated with the selection of politicians are also related to politicians' behaviour after they obtain office.

Politicians' Facial Appearance and Behaviour

Research on politician's facial appearance and behaviour is scarce compared to research on selection effects. As stated above, only one study has investigated facial appearance and behaviour in the political realm (Lin, Adolphs, & Alvarez, 2018). As noted by the study's authors, the causal mechanisms underpinning the associations they found are unclear. For example, the facial appearance of the officials might differentially attract opportunities to engage in corruption (e.g., other people's willingness to bribe the official), or affect the likelihood that officials are charged and convicted. Given the limited knowledge about facial appearance and the behaviour of politicians, we review face-based social attribution research that has been conducted in other settings. According to this research, how might the facial appearance of politicians be associated with their behaviour while in office?

Recent research shows that men who look less trustworthy are less likely to be sexually faithful (Foo, Loncarevic, Simmons, Sutherland, & Rhodes, 2019), and in experiments, those with objectively wider faces are more likely to cheat in a negotiation game (Haselhuhn, & Wong, 2012). Further, endogenous levels of testosterone are associated with corruption, which is not mediated by individual differences in honesty (Bendahan et al, 2015). Facial appearance has been associated with socially desirable behaviours as well. Rule and Ambady (2008) argued that naïve (subjective) trait ratings of CEOs' faces predict financial success in terms of company profits; but, they did not control for the firms' prior financial success or scale profits by firm size (Graham, Harvey, & Puri, 2017). Wong, Ormiston and Haselhuhn (2011) found that the facial structure of male CEOs is related to financial performance: those with wider faces were more successful. By controlling for the firms' previous performance in their statistical analyses, Wong et al. (2008) argue this rules out selection effects, or the possibility that more successful firms simply appoint CEOs with certain facial characteristics.

However, the interpretation of these findings has been called into question by Graham et al. (2017), who more carefully controlled for prior performance. They found that although facial appearance affected CEO selection and compensation, there was no evidence of facial appearance predicting financial success. Similarly, Stoker, Garretsen, and Spreeuwers (2016) show that, while CEOs might look different from other citizens and professors, when comparing top performing CEOs to other CEOs, facial appearance is not associated with success.

Crucially, questions about the internal validity of the studies challenge the conclusion that facial appearance is a valid indicator of behaviour. Various studies show that even though people might feel confident in making facial appearance-based judgments, the accuracy of these judgments is limited (Hassin & Trope, 2000; Olivola & Todorov, 2010b). According to Olivola and Todorov (2010b), people tend to put too much emphasis on facial appearances when making social judgments about others; knowledge about base-rate frequencies of social categories are underutilized but would be more informative. Perhaps facial appearance is drawn upon because people rely on the representativeness heuristic (Kahneman & Tversky, 1973) to make social inferences. Further, as discussed by Todorov and colleagues (2015), stimulus selection effects also compromise the internal validity of the findings. They point out that although some studies have found that high width-to-height ratio predicts aggressive behaviour (Carré & McCormick 2008, Carré, McCormick & Mondloch, 2009), this has not been consistently replicated, perhaps owing to differences in the stimuli used in the studies. Thus, the question about whether facial appearance-based behavioural inferences are valid is yet to be resolved within the literature.

At the very least, it seems that if facial appearance affects behaviour, the mechanism underpinning the association is likely to be far more complicated than might be first assumed (e.g., Rule & Ambady, 2008; Wong et al., 2008). The potential complexity of the underlying relationships between facial appearance, personality, and behaviour becomes obvious when we consider them as a function of the effects of perceived facial appearance on an individual's own behaviour, and on that of observers' judgements as well as their behaviour towards individuals who possess a given facial appearance. For example, Zebrowitz and co-authors (Zebrowitz 2018; Zebrowitz & Montepare, 2008; Zebrowtiz, Collins, & Dutta, 1998; Zebrowitz, Voinescu, & Collins, 1996) show that there is not only a correlation between early facial appearance and later personality, but also between both early personality and later facial appearance. For example, Zebrowitz et al. (1998) show that men who were more attractive in adolescence were more likely in their 50s to have higher levels of sociability and lower hostility. However, Zebrowitz et al. (1998) also find evidence of compensation effects: adolescents who were baby-faced as boys exhibited less babyish behaviour on average, as if compensating for early societal expectations. More recently, patients rated as appearing more trustworthy and distressed by independent judges were found to be triaged

sooner by emergency healthcare providers (Bagnis et al.. 2020). Examples like these in the literature of such dynamic interdependence makes it possible to see there are several ways in which the willingness of politicians to engage in unethical behaviour might be related to their facial appearance.

In summary, the association between facial appearance and behaviour is weak and equivocal, especially in the context of politics, where there has been little research on this topic. Whilst it seems quite clear that politicians' facial appearance influences voting behaviour (e.g., Olivola & Todorov, 2010a; Todorov et al., 2005), very little is known about whether the facial appearance of politicians is related to their behaviour in office, and the limited research on this topic has not used independent and objective measures of behaviour. The literature taken as a whole suggests that associations between facial appearance and behaviour are complex and are likely bidirectional. Therefore, in the present study, we take an explorative approach, and do not make specific predictions about facial appearance and behaviour in politicians. In particular, we focus on unethical behaviour in a sample of British politicians.

Failure to trust politicians is a common and longstanding source of antipathy and voter discontent in the UK (e.g., Henn & Foard, 2012; Norris, 2011; Farrell, McCallister, Studlar, 1998). In 2016, only 16% of the British public responded that they trusted politicians to tell the truth (Ipsos Mori, 2016). Given these results, it is unsurprising that 67.5% of a sample of British voters stated that they would prefer their MPs to be honest than hard-working (Allen & Birch, 2011). In fact, honesty, along with competence, is the attribute most highly prized by the voting public (Besley, 2005). Given the importance of the issue to the electorate, we studied whether the facial appearance of politicians corresponds with unethical behaviour while in office, drawing on data from the British Parliamentary Expenses scandal, which led to public outrage over widespread unethical behaviour among politicians in the UK.

The British Parliamentary Expenses Scandal

The British Parliamentary Expenses Scandal provides us with a useful opportunity to explore the relationship between facial appearance and unethical behaviour. It provides a context in which trait ratings of facial images can be compared to an objective measure. After the media disclosed the full details of claims made by MPs, an independent auditor found that over half of MPs were guilty of overclaiming; the average MP had overclaimed £1,500 (Pattie & Johnston, 2012) between 2004 and 2008. One individual overclaimed by £63,250, and another claimed expenses to clean a moat on their country

estate (Legg Report, 2010). The majority of these other claims were associated with the second-home allowance, particularly, improper mortgage interest or rent claims, and the costs of cleaners and gardeners (Legg Report, 2010). The BPES remains a significant event, which exacerbated longstanding antipathy and voter discontent in the UK (e.g., Norris, 2011, also see Ipsos-Mori, 2016). Thus, it provides an important context for our exploration of politicians' facial appearance and unethical behaviour.

We explore this question by measuring MPs' facial traits and use overclaiming as an objective measure of behaviour. We had participants, who were previously unfamiliar with these politicians, make trait evaluations (e.g., honest, competent, physically attractive) based solely on the physical appearance of the politicians' faces. Given little is known about the relationship between politician appearance and behaviour, participants evaluated a broad range of traits studied in previous studies on appearance and political selection and facial appearance and behaviour in other domains. These include dominance and other agentic traits (Eagly et al., 2019) which are both important for political selection (Berinsky, Chatfield & Lenz, 2019) as well there being some evidence that having a criminal facial appearance is associated with criminality (Valla, Ceci, & Williams, 2011). Likewise, given the well-documented importance of attractiveness, broadly conceived, for political selection, we collected data on several traits capturing different aspects of attractiveness, identified by Olivola & Todorov (2010) and others such as likeability, sincerity, honesty, etc. Finally, we collected data on whether an MP appeared organized or competent given previous findings that facial appearances of competence are important for political selection (Ballew & Todorov, 2007) in the UK (Mattes & Milazzo, 2014).

Method

Participants

4,727 people (51% female, 56% 18-29 years [range: 18-91 years], 54% from the United States, 20% from India and 26% from other countries) were recruited using Amazon Mechanical Turk. They were paid $0.50 for their participation.

Materials

The photographs of all the 636 MPs serving in 2009, who had been in office since at least 2005, were provided by Dods Parliamentary Communications Ltd.[1] These head and shoulder photographs were high quality and relatively uniform in terms of lighting conditions, focal distance, and professional manner of dress and facial appearance. We chose to work with the entire population of MPs rather than choosing a sub-sample based on an *a priori* power analysis. However, we computed the minimum sample size using the meta-analytic average effect size for published social-personality psychology studies (Richard, Bond, & Stokes-Zoota, 2003), which is $r \approx .20$, assuming 5 covariates, a desired power level of 0.95 and a two-tailed probability level of 0.05. This was 105, so we may be confident that our analysis has sufficient power. Note that the fixed number of MPs precludes a conventional power analysis.

Procedure

Each participant evaluated 55 randomly selected MPs, who were presented in a random order, all on the same trait, which was randomly selected for each participant out of a possible 10.[2] The traits they could be assigned to evaluate included: physically attractive, charismatic, criminal, competent, financially greedy, honest, likeable, organized, physically dominant, and sincere. The traits were measured using a 7-point Likert-type scale, anchored at 1, 'not at all', and 7, 'very much'. Each face remained onscreen until the participant completed their evaluation.[3]

Measures

Each politician was rated on every trait, with each trait rated by at least 30 participants, a sample size chosen based on the findings of the previous literature (e.g., Willis & Todorov, 2006), which suggests around 25 participants per trait per politician gives reliable measures. Around 1% of respondents recognized that the faces were of British politicians and these respondents were excluded. Computing Cronbach's alpha using random effects regression suggested the ratings were reliable (all α's $> .76$). We compute the reliabilities in this way because each participant only assessed a subset of the photographs on a subset of traits. Moreover, because, as described below, our interest is in the overall characteristics of

[1] We exclude MPs who died or stepped-down mid-parliament as well as their replacements. We also exclude Tony Blair and Boris Johnson given as Prime Minister and Mayor of London respectively, they may be internationally well known.

[2] An earlier version of this paper reported results of a similar previous study in which participants rated each MP on multiple traits, potentially biasing ratings and thus the correlation structure. The results of this study are presented in Appendix C.

[3] A screenshot of the raters' task is provided in Appendix A.

politicians' faces and thus with latent factors, it is not possible to directly treat the photographs as a random factor, as described by Judd, Westfall, and Kenny (2012) and used in a similar context by Banchefsky, Westfall, Park, and Judd (2016) in our regression analyses. Given the ordinal nature of our data, and the heterogeneous rating scales of respondents, each politician's score on each trait was obtained via a one parameter item response model with respondent specific effects (Agresti, 2010; Bürkner & Vuorre, 2019). See Appendix B for details.

Our measures of unethical behaviour were derived directly from the Legg report, a lengthy report documenting the amount MPs were judged to have overclaimed (www.data.parliament.uk), and including whether the MP was found to have overclaimed, and how much money the MP was required to repay.[4] We used the initial amount recommended to be repaid by the Legg report as our measure of overclaimed expenses. Legg defined overclaiming as claims that were beyond the limits permitted in the so-called *Green Book*, contrary to the broader rules governing *Standards in Public Life,* or otherwise unjustifiable – for example, claiming for expenses not actually incurred.[5] Figure A1 in Appendix A presents histograms showing the distribution of these data.[6] When the BPES broke, all Members of Parliament (MPs) received salaries of £64,766, except government ministers, the Speaker, and chairs of select committees, who receive additional supplements.[7] Crucially for this study, MPs could claim a range of other expenses. They were eligible to claim up to £70,000 in staff costs and £20,000 for the costs of maintaining an office. MPs who did not represent inner London constituencies were also able to claim up to £20,000 a year towards the costs of maintaining a second home, as well as travel costs, and other incidentals. Thus,

[4] Some repayments were changed following appeal. Given that the choice to appeal is not random, and may be correlated with the trait measures, we preferred to use the initial judgments as the measure.

[5] A full description may be found in the Legg Report itself: https://publications.parliament.uk/pa/cm200910/cmselect/cmmemest/348/app1.pdf
The *Green Book* is available at: https://www.parliament.uk/documents/commons-finance-office/greenbook.pdf

[6] As well as all claims being public, the high level of attention the BPES received meant that Legg, the auditor, was under considerable scrutiny. As such, there are good reasons to believe the investigation was as careful as possible. MPs had the right to make their case, provide additional evidence, and dispute interim findings of the auditor. These judgements were made on the basis of explicit criteria as to what constituted overclaiming. These and the reasoning for them were detailed in his report. Thus, while there may be unconscious bias, it is reasonable to believe the report provides face valid data that is as reliable and as objective as possible.

[7] Such an income placed MPs in the upper decile of earners. However, this is not as large as the salaries earned by other senior public servants or public sector employees. MPs were also eligible for a comparatively generous final salary pension scheme. As well as additional income, government ministers are also eligible, depending on seniority and department, for a range of other benefits such as chauffeur-driven cars and free accommodation.

an MP could claim an amount substantially in excess of £100,000 towards their costs, including travel to and from their constituency, travel for their staff, and the costs of an additional home.

There are a range of other facial appearance and career dimensions along which MPs vary that likely correlate with both raters' perceptions of faces and the election process that we measured. For each MP, we coded their age, ethnicity, and gender. Further, whilst the photos are homogenous head-collar shots, they did vary with regard to whether and how the politician was smiling in the photograph. Since there is evidence that this may affect how faces are perceived (Trichas & Schyns, 2012), we coded whether the MP was displaying a reward smile or an affiliative smile, or had a neutral expression (Rychlowska, et al., 2017).[8] To do so, we drew from the Duchenne smile literature to identify the smile types, which are based on the main facial muscle differences in the eyes and mouth (see Frank, Ekman, & Friesen, 1993).

We captured the career of the MPs on several dimensions, including party affiliation, and tenure (i.e., number of years in office). We defined separate binary variables for each group of MPs with the same tenure length in our data. Allowing tenure to enter in our model in such a flexible way means that we are allowing both for systematic differences across entering cohorts as well as potentially non-linear relationships between the amount of overclaiming and time served in parliament. For example, the different incentives of MPs approaching retirement, and the well-documented process of MPs learning (and being taught) how to claim their expenses to maximize their own financial gain as part of the process through which new MPs are socialised (Rush & Giddings, 2011). Controlling for binary variables for each party means we will capture differences in the social network of parliamentarians and thus differences in the socialization process. It also means we will capture differences across parties in class, wealth, and background (Dubois, Rucker, & Galinsky, 2015).

We also coded three further sources of variation that capture differences in MPs' careers. The first, *seniority*, captures the reality that some MPs are much more prominent in the public's attention than others, and that they may also have more to lose. Equally, they are powerful and thus may take more risks (Anderson & Galinsky, 2006). To measure *seniority* we extracted data on the complete parliamentary career histories of MPs from data.parliament.uk. We then calculate *seniority* as the sum of time spent in each position weighted by the seniority of that position. Details are in Appendix A. The second is the size of

[8] We would like to thank an anonymous reviewer for making this suggestion.

each MP's *majority,* measured as the winning margin of the MP (in percent of votes cast). This captures the intuition of the rational choice approach that MPs who are confident of being re-elected face different incentives than those whose chances are lower. On one hand, if they expect to be comfortably re-elected then they may feel they can afford to take more risks than those in marginal constituencies. On the other, it suggests those who believe they are likely to be voted out of office anyway, have a lower cost of being caught (Besley & Case, 1995; Smart & Sturm, 2013). Finally, Besley and Larcinese (2011) analysed patterns in legitimate claims made by MPs in the three years prior to 2004 and showed that claims are higher for those who visit parliament more, as would be expected, and those that represent constituencies that are further from Westminster. Therefore, thirdly, we coded whether an MP represented an inner London constituency, and was ineligible to claim the costs of a second home, thus limiting their opportunity to overclaim.

Results

The Latent Factor Structure of Politicians' Facial Appearance

We were interested in identifying the relationship between the overall characteristics of politicians' faces and their behaviour. Therefore, we began our analyses by identifying the latent factors that describe the variation in the facial appearance of politicians' faces. We used the Treelet Transform (TT) to obtain the latent factor structure of the 10 traits. A recent innovation from the machine learning literature (Lee, Nadler, & Wasserman, 2008), the TT has several key advantages that make it preferable to traditional approaches, such as principal components analysis. First, it provides a straightforward means to visualize the structure of a dataset. Second, it provides a relatively sparse set of loadings such that each factor has a limited number of variables with nonzero weights, facilitating interpretation. The optimal number of factors and the maximum number of traits per factor are recovered using a cross-validation approach. Moreover, bootstrapping provides a natural means with which to test quantitatively the stability of the factor loadings. Figure 2 is a dendrogram, which shows in an intuitive manner the factor structure of the rating data resulting from the TT analysis. Moving upwards along the y-axis, each join represents the collapsing of the two most highly correlated variables into a common factor via principal components analysis. In our data, *honest* and *likeable* have the highest pairwise correlation, followed by their principal

component with *sincere*, and then that principal component with *charismatic*. Next, *financially greedy* and *criminal* have the highest correlation, and so on.[9]

Inspection of the scree plot of variances of each factor (in Appendix A) shows, following Gorst-Rasmussen and colleagues (2011), that there are three significant latent factors with a variance greater than one. Looking at the final join, we can see which variables are loaded on to these factors. These can be readily seen to align with the three factors identified by Todorov et al. (2015) and Sutherland et al. (2013), as well as those of Eagly et al. (2019). The first comprises *charismatic*, *physically attractive*, *likeable*, *sincere*, and *honest*. The second comprises *criminal*, *financially greedy*, and *physically dominant*. Our third, contains *organized* and *competent*. Figure 3 describes the loadings of these three factors.(Numerical results are reported in Appendix A). Following Todorov et al. (2015) we label the first TT component *Attractiveness,* the second as *Criminality* and the third as *Competence.* That is, *Criminality* is the weighted sum of criminal, financially greedy, and physical dominance, and *Attractiveness* the weighted sum of charismatic, physically attractive, likeable, honest, and sincere, and *Competent* is the weighted sum of organized and competent where in each case the weights are the TT factor loadings. The factors are reliable (*Attractiveness* α=0.88; *Criminality* α=0.80, *Competence* α=0.73).[10]

[9] This sequential procedure means that unlike Principal Components Analysis the TT is sensitive to variables being reverse coded. Here variables are coded as they were collected. In Appendix D we show that we obtain the same factor structure with conventional and sparse-PCA estimators as well as using exploratory factor analysis. All results also support the use of three factors.

[10] We performed a Confirmatory Factor Analysis to assess how well the factors identified by the TT matched the data. The CFI criterion was 0.84 below the conventional criteria of 0.9 this suggested that there were differences between the raw data and the results of the TT. The same inference was reached based using the Likelihood Ratio test instead ($\chi^2(32)$, $p<0.05$). This is perhaps to be expected given our emphasis on sparse interpretable factors. We performed a modification index analysis, and this suggested that the largest improvements could be obtained by including physical attractiveness and physical dominance on the other two traits. All results are included in our replication package.

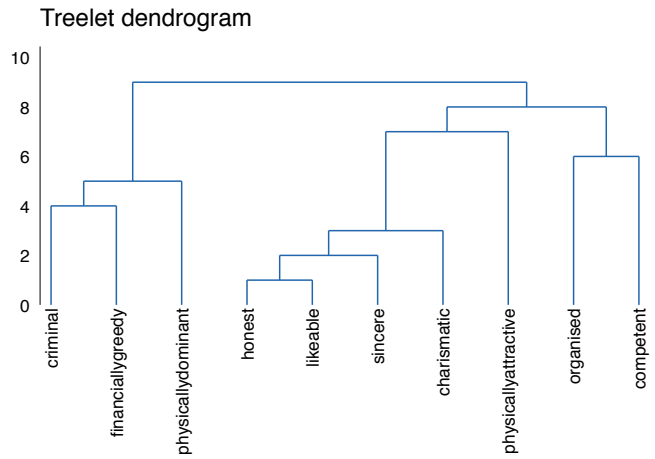

Figure 2. Dendrogram showing the factor structure of the trait ratings. As shown, there are three latent factors, which align with the Attractiveness, Criminality, and Competence factors. See text for further details.

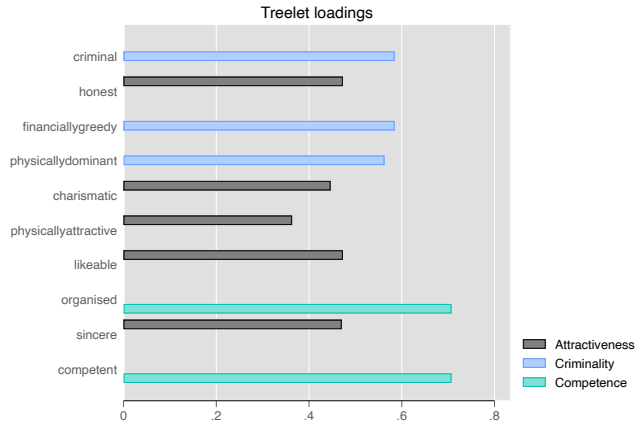

Figure 3. Bar Chart showing the loadings of each trait on the three latent factors: Attractiveness, Criminality, and Competence.

Figure A5 in the Appendix presents scatter plots of the unconditional relationship between each of the latent factors and the dependent variables. Broadly speaking, these suggest that there is no clear unconditional relationship. In the analyses that follow, we correct our standard errors for arbitrary heteroskedasticity. We standardise each of our main effect variables *Attractiveness*, *Criminality*, *Competence* as well as *seniority* and *majority* to have mean 0 and standard deviation 1 for ease of interpretation. Tables 1 and 2 provide summary statistics and correlation matrices respectively.

Facial appearance and overclaiming

To control for confounding factors, we regressed log overclaiming on our control variables, *male, age, white, reward smile, affiliative smile,* and binary variables for each political party and for each length of *tenure.*

We work with the logarithm because it facilitates inference, and ensures our results are not driven by those who overclaimed a great deal.[11] We work with the absolute amount overclaimed, rather than as a percentage of total claims, since the actual amount claimed will reflect a range of confounding factors, such as the location of MPs' constituencies and the amount of time they spend in London. The first three controls are included to capture possible differences in how people perceive the facial appearance of men and women, the facial appearance of older people, and perceptions of ethnicity.

As a first-step, Column (1) of Table 3 reports results including only the control variables capturing differences in demographics and facial appearance only. The results show that of the six control variables, only age and being an inner London MP significantly correlated with overclaiming, *age* (b=0.032,SE=0.16, p=0.04, $r_{partial}$=0.082,CI_{95}=[0.004,0.16]). The R^2 of the regression is 3%, confirming that these observable characteristics have little explanatory power.[12] Column (2) now includes our full set of controls, including *seniority* (b=0.41,SE=0.12,p=0.001, $r_{partial}$=0.14, CI_{95}=[0.06,0.22]) and *majority* (b=-0.18,SE=0.17,p=0.28, $r_{partial}$=-0.05, CI_{95}=[-0.13,0.04]), as well as the party and tenure fixed-effects. The latter have some explanatory power, as the R^2 of the model increases to 0.15, consistent with both increased opportunities for MPs with longer tenure as well as a process of socialisation. Next, we consider, one-by-one in the order identified by the TT, the explanatory power of the three main factors identified above. Column (3) includes *Attractiveness,* which shows no significant relationship with overclaiming (b=-0.07,SE=0.19,p=0.71, $r_{partial}$=-0.015, CI_{95}=[-0.1,0.07]). Column (4) includes *Criminality,* which is significant and negative (b=-0.44,SE=0.19,p=0.023, $r_{partial}$=-0.09, CI_{95}=[-0.18,-0.01]). The R^2 of the model is improved by 0.01. Given our log-linear specification and that we have standardised our main effect variables, we can interpret the coefficient on *Criminality* of an MP's facial appearance as implying that a one standard deviation increase in *Criminality* is associated with a reduction in overclaiming by 44%. While this is a quantitatively large effect, the impact on the overall fit of the model is relatively small, with an increase in R^2 of around one percentage point. Column (5) reports the results including *Competence,* which

[11] In Figures E1 and E2 of the appendix we show that outliers are likely to be a problem working with the untransformed data. Results using robust-regression techniques are similar and are reported in columns 1-3 of Table E2.

[12] One possibility is that there may have been regional differences in the scope for overclaiming. Columns 4 &5 of Table E2 in the Appendix report results for the specification reported in column 6 but additionally including either government office region or county fixed-effects. The results are unaffected.

is not significant, albeit with a reasonably large coefficient, and does not improve the fit of the model *(b=0.21, SE=0.16,p=0.18,* $r_{partial}$=0.06, CI_{95}=[-0.03,0.14]). In sum, the results thus far indicate that politicians who look *less* criminal overclaim more, but that *attractiveness* and *competence* are not significantly related to overclaiming.

Of course, as discussed by Zebrowitz and Montepare (2005), Todorov et al. (2015), and Sutherland et al. (2013), voters will assess a face's features jointly. Thus, in column (6) we report results of a model including all three factors together. *Attractiveness* (b=-0.36,SE=0.21, p=0.09, $r_{partial}$=-0.07, CI_{95}=[-0.15,0.01]) is significant at the 10% level, while *Competence* (b=0.43,SE=0.18, p<0.02, $r_{partial}$=0.10, CI_{95}=[0.02,0.18]), and *Criminality* (b=-0.54, SE=0.2, p<0.01, $r_{partial}$=-0.11, CI_{95}=[-0.19,-0.03]) were significant at the 1% and 5% levels respectively. This suggests that holding all else constant, including the facial appearance of *Criminality* and *Competence*, that *Attractiveness* is now associated with overclaiming less. Similarly, holding *Attractiveness* and *Criminality* constant, *Competence* is associated with overclaiming more. Again, the R^2 of the model only improves marginally. This suggests that, conditional on other factors, that there is insufficient variation in facial-appearance to explain much of the variation in over-claiming. Given that the factors we use are measured with error, one concern is that this may be biasing our results. To allow for this possibility we also estimated a structural equation model in which we treated the three factors as latent variables. The results, contained in our replication package, are very similar, consistent with our results not being biased due to measurement error.

Additional Results

To further investigate the combined impact of the three factors Table 4 Columns (1) to (4) presents a series of models that first allow for the pairwise interactions of traits separately, and then all these interactions and the triple interaction together. That is, we allow for *Criminality* to be moderated by *Attractiveness* or *Competence,* and so on. We find no evidence of any pairwise interactions or for the triple interaction. We also find no evidence of an interaction between either *seniority* or *majority* and *Criminality* suggesting that while more senior politicians overclaimed more, this effect was not significantly moderated by their facial appearance. Column (7) presents results in which the effect of each factor is moderated by age. The results suggest that MPs who were both young and attractive may overclaim more than average, while MPs who were older and attractive overclaimed less. Similarly, it suggests that the

negative relationship between *Criminality* and overclaiming is more pronounced in older MPs, *ceteris paribus*.

Sensitivity Analyses

We undertook a number of additional analyses to check the robustness of our results, including using an alternative measure and models of overclaiming, as well as conducting analyses that restricted the sample to White male politicians.

Alternative overclaiming measures. Panel (c) in Table 5 reports results using the percentage of an MP's annual salary (£64,766) that was overclaimed instead of the log-transform of the amount overclaimed. This alternative dependent variable takes values between 0 and 0.98, and coefficients now describe the increase in percentage points associated with a standard deviation increase in a factor, rather than a percentage increase. The coefficients are again similar, although now only *Competence* is significant. A natural explanation is that this reflects the greater sensitivity of this specification to extreme observations. One advantage of the log specification is that it is less sensitive to outliers as the relative distance of those who overclaimed the most from the mean is lower.

Alternative models of overclaiming. An implicit assumption in our previous analysis is that we should observe the amount of overclaiming varying continuously with facial appearance. But, 340 MPs did not overclaim at all. We approach this issue in two ways. One possibility implied by the rational-choice perspective is that, were it possible, some MPs would have chosen to underclaim. To allow for this possibility we re-estimate the specifications reported in Table 3, but now using a Tobit model in which we allow for censoring from below at £100 (the minimum overclaim Legg considered). These results are presented in Panel (a) of Table 5. We again see that Attractiveness (b=-0.04, SE=0.01), Criminality (b=-0.34, SE=0.01), and Competence (b=0.19, SE=0.01) are now significant by themselves as well as when all three factors are included together.

An alternative view is that the key choice faced by politicians is between overclaiming or not. This would suggest that while how much individuals overclaimed, given they overclaimed, was determined by other factors; the decision to overclaim or not was in part determined by facial appearance. To consider this possibility, Panel (b) in Table 5 reports the results of a logit model with a binary variable *OC*

capturing whether or not an MP overclaimed at all.[13] The results are qualitatively and quantitatively similar to the results in Table 3, although slightly less precise and *Attractiveness* (b=-0.19, SE=0.13) is not significant, as shown in Panel (b) column (4). The logit analysis provides an alternative way in which to interpret the estimated magnitude of the effects of facial-appearance on behaviour. Computing predictive margins suggests that an increase in one standard deviation increase in *Criminality* is associated with an 9pp (M=-0.090, SE=0.029) decrease in the chance of having overclaimed. Likewise a one standard deviation increase in *Competence* is associated with an 6pp (M=0.056, SE=0.027) increase. Thus, again, while the estimated effect sizes are small in the statistical sense, they are substantively important.

A further alternative is to treat those who overclaimed as different from those who did not and estimate a hurdle model which estimates separately a) the traits correlated with overclaiming at all and b) the traits that predict overclaiming more. The results of this model, and two close alternatives (zero-inflated Poisson and negative-binomial models) are reported in Table E1 in the Appendix. The results suggest that the appearance factors that predict whether an MP overclaimed do not predict the extent of overclaiming conditional on having overclaimed. This is consistent with the similarity of our logit and OLS estimates.

Analyses restricted to White males. Given that the effects of race and gender may interact with facial appearance in ways not captured by our specification in column (6) of Table 3, we conducted an additional analysis where we restricted our attention to white male MPs which is displayed in column (7) of Table 3. The sample size is now smaller, but the performance of the model is higher with an R^2 of 0.37 compared to 0.17 in column (6). The magnitude of the coefficient estimates for *Attractiveness Competence,* and *Criminality* are larger and more precisely estimated. Column (5) of each panel in Table 5 suggest the same pattern obtains for the other outcome measures. Taken together, the sensitivity analyses suggest that the results are consistent across measures and models of overclaiming and are not affected by MP gender or race.

In sum, across the results, more attractive and more criminal-looking politicians tended to overclaim less. Further, a more competent facial appearance was positively, albeit less-precisely, related to overclaiming.

[13] An alternative to the Logit is to estimate a Linear Probability Model which allows for all cases to be retained and for heteroskedastic standard errors. We report these results in columns 6-9 of Table E2 in the Appendix.

Although the relationships were statistically significant, the additional explanatory power of facial-appearance was low, only around 2% suggesting the relationship between behaviour and appearance is weak and equivocal. Finally, there was evidence that these effects are moderated by age and whether the politician was from inner London.

Discussion

We studied the British parliamentary expenses scandal of 2009 to explore how facial appearance factors that voters find appealing when selecting politicians are related to unethical behaviour. Introducing an approach from the machine learning literature, the Treelet Transform, we obtained three factors that captured the facial appearance of the politicians: *attractiveness*, *criminality*, and *competence*. This replicates previous findings showing these are key appearance factors in the social perception of faces (Eagly et al., 2019; Chiao, Bowman, & Gill 2008; Todorov et al., 2005; Oosterhof & Todorov 2008; Todorov et al., 2008). Our study is the first, however, to analyse whether these facial appearance factors are associated with unethical behaviour in politicians, which we operationalised by measuring whether, and by how much, politicians overclaimed on their expenses.

We found that while the facial appearance of politicians was associated with overclaiming, the pattern of results was complex. What is more, effect sizes were small, and while the relationships were significant statistically, our control variables (i.e., the politician's age, residing in London, tenure in office, seniority and party affiliation) accounted for relatively more variation (15% versus 2%) in overclaiming. Thus, our findings suggest that facial appearance does not strongly predict behaviour and align more closely with the growing body of work finding that facial appearance is not correlated with behaviour (Graham, Harvey, & Puri, 2015; Stoker, Garretsen, and Spreeuwers, 2016). Insofar as we did find some evidence for limited relationships between facial appearance and behaviour, we now discuss them in more detail and consider theoretical explanations.

When we considered independently the facial appearance factors in models that also included our control variables, criminality was the only factor that was significantly and negatively associated with overclaiming. We also considered interrelationships between our appearance-based factors and overclaiming since there is evidence that voters make social inferences based on holistic evaluations of facial appearance (Zebrowitz & Montepare, 2005; Todorov et al., 2015; Sutherland et al., 2013). When we

modelled all of the appearance-based factors together with the control variables, competence, attractiveness, and criminality were significantly associated with overclaiming, albeit the size of the relationships was small. Competence was positively associated with overclaiming, whereas attractiveness and criminality were negatively associated with overclaiming.

Our finding that politicians who appeared more criminal-looking over-claimed less is somewhat counter to Lin et al.'s (2018) work on facial appearance and corruption in government officials. They found that dominance was positively associated with whether public officials were guilty of corruption. Given these findings, we might have expected to find that criminality would be positively associated with overclaiming. Our discrepant findings might have arisen from differences between the two studies in the type of unethical behaviour studied. Corruption is arguably more of a social enterprise, whereas overclaiming on one's expenses is more of a private activity. Engaging in corruption often means being approached by others to engage in the activity. Perhaps those who have faces that appear more dominant (and criminal-looking) are more likely to be invited by others to engage in corruption, as proposed by Lin et al. (2018)) With respect to our study, the underlying mechanism may be different to Lin et al. because the unethical behaviour (overclaiming) occurs in private. Alternatively, perhaps politicians who appear more criminal-looking attempt to counteract other people's moral presumptions of them by behaving morally when in office. This line of reasoning is in line with Zebrowitz (2018), who found evidence that people may act to counter other people's presumptions about them.

Taken together, perhaps our results may be understood by referring to the literature on moral licensing and moral consistency. Moral licensing and moral consistency emphasise the importance of self-perception in regulating behaviour, but differ in the posited mechanism (Zhong, Liljenquist, & Cain, 2009). Moral licensing posits that individuals seek to engage in moral mental accounting—compensating for self-interested, or immoral actions, with subsequent pro-social actions, and vice versa. Moral consistency emphasizes individuals desire to be consistent in their behaviour over time (Burger, 1999; Gawronski & Strack, 2012). Moreover, consistency is associated with costly (Gneezy, Imas, Brown, Nelson, & Norton, 2012) and private initial behaviour (Kristofferson, White, & Peloza, 2014). It may also reflect differences in self-perception driven by their facial appearance directly. More attractive people are treated better by others (Mobius & Rosenblat, 2006; Rosenblat, 2008) and being treated as 'good' may

lead them to act in a way that is consistent with this, a type of self-fulfilling prophecy (see Lin et al. 2008). Moreover, power tends to validate peoples' innate moral tendencies (Hirsh, Galinsky & Zhong, 2011). The power afforded to individuals by being elected may reinforce politicians' natural tendencies, and thus any tendency for attractive politicians to adopt consistent moral behaviour.

Our finding that MPs who appeared more competent overclaimed seems less compatible with a moral-consistency interpretation. But, one possibility, given the importance of appearing competent for political selection (Olivola & Todorov, 2010a) in the UK (Mattes & Milazzo, 2014) is that competent appearing politicians feel protected by their facial appearance such that small transgressions do not challenge their sense of self (Zhong et al., 2009). Alternatively, it could be that similarly to appearing less-criminal/dominant, that moral license means that more-competent appearing politicians feel able to overclaim more.

Finally, it is noteworthy that the facial characteristics we found to be associated with overclaiming, attractiveness, criminality, and competence, are also those that prior work has identified as predicting political selection (Berggren, Jordahl & Poutvaara, 2010; Rule et al., 2010; Little et al., 2007; Todorov et al. 2005**)**. Moreover, similar to other work finding that voters' evaluate facial characteristics jointly (Zebrowitz & Montepare, 2005; Todorov et al., 2015; Sutherland et al., 2013), we find that *competence* and *attractiveness* only matter for politicians' behaviour if criminality is also considered. Further, our findings that both *competence* and *attractiveness* matter for behaviour is interesting given the differing results of Berggren, Jordahl & Poutvaara (2010) and Todorov et al. (2005) who find no significant effect of competence on selection allowing also for attractiveness, and vice versa, respectively.

Overall, it is clear that our understanding of the mechanisms underpinning facial appearance and ethical behaviour is limited, and that further research that considers the context and type of unethical behaviour is needed.

Limitations

Whilst we find that facial appearance factors are significant predictors of overclaiming it is also the case that much of the variation remains unexplained. The R^2 of the models we estimate is relatively low, at 0.17 in our best specification. This suggests that most of the variation in MPs' behaviour is driven by other factors. Similarly, the estimated effect sizes of the facial appearance factors on whether politicians

overclaimed are between 0.08 and 0.15 which, and while larger than those in other recent studies of politicians' facial appearance, such as Olivola et al. (2018), the effects are smaller than those found in other contexts, such as CEOs studied by Wong et al. (2011). Moreover, among the politicians who overclaimed, facial appearance is not associated with how much politicians overclaimed. The literature on power and unethical behaviour (Hirsh et al., 2011) suggests that studying the interplay between unethical behaviour and power amongst politicians may improve substantially the amount of variation we can explain, and thus be a fruitful avenue for future research. Of particular interest are the dynamics of this relationship, since politicians are often in office for several decades with fluctuating power. Likewise, it would be valuable in future work to draw on the literature studying the relationships between social class and unethical behaviour (Dubois et al., 2015; Korndörfer, Egloff, & Schmukle, 2015). One further limitation is that the results of a confirmatory factor analysis suggest that our sparse factors do not capture all of the information in the individual traits. It is an interesting open statistical question whether it is possible to identify an optimal trade-off between sparsity, and hence interpretability, with CFA fit.

Our findings, of course, are correlational in nature, and the cause(s) of the associations we found is not known. There may be other variables that drive the relationship uncovered between overclaiming and facial appearance. Another consideration is that the participants rated photographs that were selected by the politicians themselves and/or their staffers, and the results may have differed for any given politician had a different photograph been selected (Todorov et al., 2015). Having said that, it is probably the case that the politicians selected photos of themselves that they thought would be most appealing to constituents, and this should have made it more difficult to uncover associations between facial appearance and overclaiming.

There are additional limitations to consider that arise from the demographics of politicians in the UK. The median and mean age of politicians in our dataset is 56 and only 23 MPs were younger than 40. Therefore, we were unable to explore the relationships between behaviour and facial appearance in younger politicians. Further research should incorporate a larger age range to more fully examine the role of aging. It would also be interesting to study the facial appearances of MPs when first elected. Further, while we found similar results when we restricted the sample to the subset of White male MPs, suggesting our results are not being driven by differences in how women or Black or Asian MPs are viewed, around

80% of MPs in the 2005 parliament were White men. It would be valuable in future research to study legislatures or other contexts in which there were more non-White, non-male individuals so that we could examine better the role of race and gender.

In sum, the use of facial appearance as a heuristic is widespread and has been shown in previous work to deliver substantial benefits to people blessed with good looks. The findings of this paper support selection and behaviour effects of politicians' facial appearance and overclaiming. The context we study is a particular one and further research is necessary to understand if our findings generalize beyond the domain of politics, or to other behaviours of politicians besides overclaiming. That being said, the evidence we provide of a relationship between politicians' facial appearances and unethical behaviour is an important step beyond the prior literature on politicians' facial appearance and selection.

Future Directions

An important challenge for future research is to better understand the mechanisms that underlie the relationship we found between facial traits and overclaiming. For instance, the mechanisms suggested through which competence, criminality and attractiveness affect overclaiming behaviour are each different. It will be important in future research to provide direct evidence of these mechanisms and to understand why they differ. It would also be interesting in future research to build upon our findings, analysing the preferences of voters for different facial traits based upon their political affiliation, and the extent to which parties' representatives reflect voters' preferences. A related literature has studied how different facial traits are valued in different contexts (Spisak, Dekker, Krüger, & van Vugt, 2012; Spisak, Homan, Grabo, & Van Vugt, 2012; Spisak, Blaker, Lefevre, Moore, & Krebbers, 2014). Related work has shown people with different political orientations or party allegiances prefer leaders who have different types of faces (Olivola, Sussman, Tsetsos, Kang, & Todorov, 2012; Olivola, Tingley, & Todorov, 2018). Berggren et al. (2017) show that in many countries, right-leaning politicians are consistently more attractive than their leftist counterparts, and benefit from it.

Conclusion

Voters are more likely to elect politicians who look competent (Olivola & Todorov, 2010a), attractive (Berggren et al. (2017), and dominant (Chiao et al., 2008; Rule et al., 2010). We explored the relationship between British MP's facial appearance and the extent to which they behaved unethically. The results

suggest a nuanced relationship between facial appearance and behaviour. We found that more attractive politicians overclaimed less --- a one standard deviation increase in attractiveness is associated with overclaiming 50% less. But, we also found that more criminal-appearing politicians were also likely to overclaim around 75% less and that more competent-appearing politicians were expected to overclaim 40% more. Thus, our results contradict a simple 'what you see is what you get' relationship between appearance and behaviour. Moreover, these effects only improved the explanatory power of the model marginally reflecting that facial appearance was not a primary determinant of overclaiming behaviour. They were also relatively noisy suggesting a weak and equivocal relationship. We discussed our findings through the lens of the literature on moral licensing and consistency, but it is far from unambiguous how the facial characteristics we identify are related to unethical behaviour. In particular, our finding that those who appeared more criminal overclaimed less is in contrast to prior research finding a positive correlation between having a facial appearance that is perceived as criminal-looking and unethical behaviour. This might reflect differences in context, or the population at hand. Further research is necessary to understand how any relationship between facial appearance and unethical behaviour depends on context.

Table 1: Summary Statistics

Variable	N	Mean	Std. Dev.
		Traits	
Criminal	636	3.316	0.536
Honest	636	4.348	0.522
Financially Greedy	636	4.305	0.472
Physically Dominant	636	3.992	0.51
Charismatic	636	4.087	0.493
Physically Attractive	636	3.48	0.627
Likeable	636	4.158	0.55
Organised	636	4.682	0.457
Sincere	636	4.279	0.54
Competent	636	4.589	0.327
		Dependent and Independent Variables	
log (1+ Total Repayment Recommended)	636	3.385	3.733
Repayment Recommended As Percentage of Salary	636	0.024	0.073
Attractiveness	636	0.002	1
Criminality	636	−0.002	0.998
Competence	636	0.001	1.002
Age	636	56.153	9.256
Male	636	0.805	0.396
White	636	0.986	0.118
Affiliative Smile	636	0.392	0.488
Reward Smile	636	0.325	0.469
Tenure	636	0.007	0.998
Seniority	636	0.001	1.002
Size of Majority	636	0	1
Inner London MP	636	0.036	0.187
Party	636	5.388	3.095
Tenure	636	14.127	8.604

Table 2a: Cross-Correlation Matrix: Traits

	Variables	1	2	3	4	5	6	7	8	9	10
1	Criminal	1.00									
2	Honest	−0.55***	1.00								
3	Financially Greedy	0.62***	−0.46***	1.00							
4	Physically Dominant	0.51***	−0.20***	0.59***	1.00						
5	Charismatic	−0.29***	0.63***	−0.27***	−0.06	1.00					
6	Physically Attractive	−0.06	0.44***	−0.05	0.20***	0.52***	1.00				
7	Likeable	−0.44***	0.74***	−0.37***	−0.11**	0.71***	0.56***	1.00			
8	Organised	−0.21***	0.33***	−0.03	0.19***	0.24***	0.34***	0.32***	1.00		
9	Sincere	−0.42***	0.72***	−0.43***	−0.17***	0.64***	0.40***	0.73***	0.29***	1.00	
10	Competent	−0.18***	0.43***	−0.01	0.21***	0.46***	0.53***	0.38***	0.61***	0.41***	1.00

Table 2b: Cross-Correlation Matrix: Dependent and Independent Variables

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	log (1+ Total Repayment Recommended)	1.00														
2	Repayment Recommended As Percentage of Salary	0.51***	1.00													
3	Attractiveness	−0.04	−0.02	1.00												
4	Criminality	−0.04	0.01	−0.37***	1.00											
5	Competence	0.05	0.09*	0.50***	−0.01	1.00										
6	Age	0.07†	0.09*	−0.19***	0.03	−0.10**	1.00									
7	Male	0.02	0.02	−0.39***	0.58***	0.19***	0.06	1.00								
8	White	−0.03	0.00	−0.01	−0.02	−0.03	0.06	−0.03	1.00							
9	Affiliative Smile	−0.03	−0.08†	0.08*	−0.12**	−0.03	0.05	−0.13***	0.00	1.00						
10	Reward Smile	0.03	0.03	0.35***	0.23***	0.08*	−0.09*	−0.07†	0.01	−0.56***	1.00					
11	Tenure	0.13***	0.17***	−0.20***	0.13***	−0.04	0.68***	0.18***	0.06	−0.01	−0.11**	1.00				
12	Seniority	0.09*	0.04	0.10*	0.04	0.12**	−0.12**	−0.01	0.01	−0.03	0.04	−0.01	1.00			
13	Size of Majority	0.02	0.00	0.00	0.11**	0.09*	0.18***	0.09*	−0.04	−0.08†	0.07†	0.28***	0.14***	1.00		
14	Inner London MP	−0.16***	0.06	0.01	−0.06	0.03	0.04	−0.12**	−0.04	0.03	−0.03	0.07†	−0.01	0.00	1.00	
15	Party	−0.11**	−0.14***	0.05	−0.13***	−0.09*	0.02	−0.13***	−0.04	0.00	0.02	−0.14***	0.14***	0.04	0.07†	1.00

Table 3: The relationship between facial appearance and the amount overclaimed.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attractiveness			−0.070			−0.36†	−0.71**
			(0.19)			(0.21)	(0.27)
			⟨0.710⟩			⟨0.090⟩	⟨0.009⟩
			−0.015			−0.070	−0.14
			[0.066 ,-0.097]			[0.011 ,-0.15]	[-0.035 ,-0.24]
Criminality				−0.44*		−0.54**	−0.69**
				(0.19)		(0.20)	(0.26)
				⟨0.023⟩		⟨0.006⟩	⟨0.008⟩
				−0.094		−0.11	−0.14
				[-0.013 ,-0.18]		[-0.033 ,-0.19]	[-0.038 ,-0.25]
Competence					0.21	0.43*	0.63**
					(0.16)	(0.18)	(0.23)
					⟨0.178⟩	⟨0.016⟩	⟨0.007⟩
					0.056	0.10	0.14
					[-0.025 ,0.14]	[0.019 ,0.18]	[0.039 ,0.25]
Age	0.032*	0.0061	0.0050	0.0056	0.0078	0.0036	−0.0039
	(0.016)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.035)
	⟨0.040⟩	⟨0.803⟩	⟨0.837⟩	⟨0.817⟩	⟨0.750⟩	⟨0.884⟩	⟨0.913⟩
	0.082	0.010	0.0086	0.0096	0.013	0.0061	−0.0059
	[0.0038 ,0.16]	[-0.071 ,0.092]	[-0.073 ,0.090]	[-0.072 ,0.091]	[-0.068 ,0.095]	[-0.075 ,0.088]	[0.099 ,-0.11]
Male	−0.090	−0.51	−0.56	0.079	−0.38	0.18	
	(0.37)	(0.40)	(0.43)	(0.47)	(0.41)	(0.49)	
	⟨0.811⟩	⟨0.211⟩	⟨0.188⟩	⟨0.865⟩	⟨0.357⟩	⟨0.705⟩	
	−0.0095	−0.052	−0.055	0.0071	−0.038	0.016	
	[0.069 ,-0.088]	[0.029 ,-0.13]	[0.027 ,-0.14]	[-0.074 ,0.088]	[0.043 ,-0.12]	[-0.066 ,0.097]	
White	−1.79†	−1.99	−1.98	−2.08†	−1.96	−2.03†	
	(1.02)	(1.21)	(1.21)	(1.23)	(1.22)	(1.20)	
	⟨0.080⟩	⟨0.102⟩	⟨0.101⟩	⟨0.091⟩	⟨0.107⟩	⟨0.091⟩	
	−0.070	−0.068	−0.068	−0.070	−0.067	−0.070	
	[0.0083 ,-0.15]	[0.013 ,-0.15]	[0.013 ,-0.15]	[0.011 ,-0.15]	[0.014 ,-0.15]	[0.011 ,-0.15]	
Affiliative	−0.19	−0.11	−0.065	−0.31	−0.10	−0.14	−0.20
Smile	(0.37)	(0.38)	(0.39)	(0.39)	(0.38)	(0.40)	(0.52)
	⟨0.599⟩	⟨0.778⟩	⟨0.870⟩	⟨0.420⟩	⟨0.785⟩	⟨0.730⟩	⟨0.700⟩
	−0.021	−0.012	−0.0068	−0.033	−0.011	−0.014	−0.021
	[0.057 ,-0.099]	[0.070 ,-0.093]	[0.075 ,-0.088]	[0.048 ,-0.11]	[0.070 ,-0.093]	[0.067 ,-0.096]	[0.084 ,-0.13]
Reward Smile	0.12	0.42	0.49	0.11	0.40	0.36	0.22
	(0.39)	(0.40)	(0.44)	(0.42)	(0.40)	(0.45)	(0.61)
	⟨0.759⟩	⟨0.291⟩	⟨0.263⟩	⟨0.784⟩	⟨0.313⟩	⟨0.428⟩	⟨0.714⟩
	0.012	0.044	0.047	0.011	0.042	0.033	0.020
	[-0.066 ,0.090]	[-0.037 ,0.13]	[-0.035 ,0.13]	[-0.070 ,0.093]	[-0.039 ,0.12]	[-0.048 ,0.11]	[-0.085 ,0.12]
Inner London MP	−3.27***	−3.72***	−3.72***	−3.69***	−3.77***	−3.80***	−5.37***
	(0.41)	(0.34)	(0.34)	(0.35)	(0.34)	(0.34)	(1.49)
	⟨0.000⟩	⟨0.000⟩	⟨0.000⟩	⟨0.000⟩	⟨0.000⟩	⟨0.000⟩	⟨0.000⟩
	−0.30	−0.42	−0.42	−0.40	−0.42	−0.42	−0.19
	[-0.23 ,-0.38]	[-0.34 ,-0.49]	[-0.34 ,-0.49]	[-0.33 ,-0.48]	[-0.35 ,-0.49]	[-0.35 ,-0.49]	[-0.086 ,-0.29]
Seniority		0.41***	0.42***	0.43***	0.39**	0.41***	0.30†
		(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.17)
		⟨0.001⟩	⟨0.001⟩	⟨0.000⟩	⟨0.001⟩	⟨0.001⟩	⟨0.085⟩
		0.14	0.14	0.15	0.13	0.14	0.092
		[0.063 ,0.22]	[0.063 ,0.22]	[0.072 ,0.23]	[0.053 ,0.21]	[0.062 ,0.22]	[-0.012 ,0.20]
Size of Majority		−0.18	−0.18	−0.16	−0.20	−0.19	−0.21
		(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.25)
		⟨0.279⟩	⟨0.280⟩	⟨0.352⟩	⟨0.236⟩	⟨0.274⟩	⟨0.417⟩
		−0.045	−0.045	−0.039	−0.049	−0.046	−0.043
		[0.036 ,-0.13]	[0.036 ,-0.13]	[0.043 ,-0.12]	[0.032 ,-0.13]	[0.036 ,-0.13]	[0.061 ,-0.15]
Observations	636	636	636	636	636	636	504
R^2	0.03	0.15	0.15	0.16	0.15	0.17	0.37
Fixed Effects	No	Party/Tenure	Party/Tenure	Party/Tenure	Party/Tenure	Party/Tenure	Party/Tenure
Sample	All	All	All	All	All	All	White Men

Note: The dependent variable is the logarithm of £1 + the total recommended repayment in the Legg report. Thus, MPs for whom no repayment was recommended are treated as having a repayment of £1. The main effects Attractiveness, Criminality, and Competence as well as Seniority, and Majority are standardised to have mean 0 and SD 1. Male, Age, White, are not standardized.† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust Standard Errors in parentheses, p-values in angular brackets. Below which are the Effect Size (Partial Correlation Coefficient) and the 95% Confidence Interval of the Partial Correlation Coefficient in Brackets.

Table 4: The relationship between facial appearance and over-claiming: Interaction Effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attractiveness	−0.30	−0.35†	−0.36†	−0.29	−0.36†	−0.36†	2.68*
	(0.22)	(0.21)	(0.21)	(0.22)	(0.21)	(0.21)	(1.12)
	⟨0.169⟩	⟨0.094⟩	⟨0.087⟩	⟨0.174⟩	⟨0.094⟩	⟨0.091⟩	⟨0.017⟩
	−0.057	−0.070	−0.071	−0.057	−0.070	−0.070	0.099
	[0.024 ,-0.14]	[0.012 ,-0.15]	[0.010 ,-0.15]	[0.025 ,-0.14]	[0.012 ,-0.15]	[0.011 ,-0.15]	[0.018 ,0.18]
Criminality	−0.55**	−0.56**	−0.53**	−0.53*	−0.56**	−0.56**	1.63
	(0.20)	(0.20)	(0.20)	(0.21)	(0.20)	(0.20)	(1.05)
	⟨0.005⟩	⟨0.006⟩	⟨0.008⟩	⟨0.012⟩	⟨0.005⟩	⟨0.005⟩	⟨0.121⟩
	−0.12	−0.11	−0.11	−0.11	−0.12	−0.12	0.065
	[-0.035 ,-0.20]	[-0.034 ,-0.20]	[-0.029 ,-0.19]	[-0.024 ,-0.19]	[-0.036 ,-0.20]	[-0.037 ,-0.20]	[-0.017 ,0.15]
Competence	0.41*	0.42*	0.39*	0.36†	0.42*	0.41*	−0.73
	(0.18)	(0.18)	(0.18)	(0.19)	(0.18)	(0.18)	(1.04)
	⟨0.021⟩	⟨0.018⟩	⟨0.032⟩	⟨0.064⟩	⟨0.018⟩	⟨0.019⟩	⟨0.482⟩
	0.096	0.098	0.089	0.077	0.098	0.097	−0.029
	[0.015 ,0.18]	[0.017 ,0.18]	[0.0078 ,0.17]	[-0.0044 ,0.16]	[0.017 ,0.18]	[0.016 ,0.18]	[0.052 ,-0.11]
Seniority	0.41***	0.41***	0.41***	0.42***	0.45***	0.42***	0.38**
	(0.12)	(0.12)	(0.12)	(0.12)	(0.13)	(0.12)	(0.12)
	⟨0.000⟩	⟨0.001⟩	⟨0.001⟩	⟨0.000⟩	⟨0.001⟩	⟨0.000⟩	⟨0.002⟩
	0.15	0.14	0.14	0.15	0.14	0.14	0.13
	[0.065 ,0.23]	[0.062 ,0.22]	[0.063 ,0.22]	[0.066 ,0.23]	[0.061 ,0.22]	[0.064 ,0.23]	[0.050 ,0.21]
Size of Majority	−0.19	−0.19	−0.18	−0.18	−0.18	−0.18	−0.21
	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
	⟨0.276⟩	⟨0.270⟩	⟨0.293⟩	⟨0.303⟩	⟨0.290⟩	⟨0.294⟩	⟨0.209⟩
	−0.045	−0.046	−0.044	−0.043	−0.044	−0.044	−0.052
	[0.036 ,-0.13]	[0.036 ,-0.13]	[0.038 ,-0.13]	[0.039 ,-0.12]	[0.037 ,-0.13]	[0.038 ,-0.13]	[0.029 ,-0.13]
Criminality	−0.20			−0.23			
×	(0.16)			(0.17)			
Attractiveness	⟨0.190⟩			⟨0.189⟩			
	−0.055			−0.055			
	[0.027 ,-0.14]			[0.027 ,-0.14]			
Criminality		−0.055		−0.015			
×		(0.16)		(0.19)			
Competence		⟨0.725⟩		⟨0.936⟩			
		−0.015		−0.0034			
		[0.067 ,-0.096]		[0.078 ,-0.085]			
Attractiveness			−0.12	−0.16			
×			(0.14)	(0.15)			
Competence			⟨0.388⟩	⟨0.273⟩			
			−0.036	−0.046			
			[0.046 ,-0.12]	[0.036 ,-0.13]			
Attractiveness				−0.012			
×				(0.13)			
Criminality				⟨0.923⟩			
×				−0.0040			
Competence				[0.078 ,-0.086]			
Criminality					−0.18		
×					(0.16)		
Seniority					⟨0.268⟩		
					−0.046		
					[0.035 ,-0.13]		
Criminality						−0.17	
× Size of						(0.17)	
Majority						⟨0.314⟩	
						−0.042	
						[0.040 ,-0.12]	
Attractiveness							−0.055**
× Age							(0.020)
							⟨0.005⟩
							−0.12
							[-0.034 ,-0.20]
Criminality							−0.039*
× Age							(0.018)
							⟨0.034⟩
							−0.088
							[-0.0068 ,-0.17]
Competence							0.021
× Age							(0.018)
							⟨0.255⟩
							0.047
							[-0.034 ,0.13]
Observations	636	636	636	636	636	636	636
R^2	0.17	0.17	0.17	0.17	0.17	0.17	0.18
Fixed Effects	Party/Tenure	Party/Tenure	Party/Tenure	Party/Tenure	Party/Tenure	Party/Tenure	Party/Tenure
Sample	All	All	All	All	All	All	All

Note: See Table 3 for details.

Table 5: The relationship between facial appearance and over-claiming: Sensitivity Analyses.

			Panel A: Tobit Estimates		
	(1)	(2)	(3)	(4)	(5)
Attractiveness	−0.040			−0.29	−0.41*
	(0.16)			(0.18)	(0.20)
Criminality		−0.34*		−0.42**	−0.29
		(0.16)		(0.16)	(0.18)
Competence			0.19	0.36*	0.49**
			(0.14)	(0.15)	(0.17)
Observations	636	636	636	636	504
Estimator	Tobit	Tobit	Tobit	Tobit	Tobit
Dep. Var.	log (1+OClaim)	log (1+OClaim)	log (1+OClaim)	log (1+OClaim)	log (1+OClaim)
Sample	All	All	All	All	White Men
			Panel B: Logit Estimates		
	(1)	(2)	(3)	(4)	(5)
Attractiveness	−0.027			−0.19	−0.26†
	(0.11)			(0.13)	(0.14)
Criminality		−0.31**		−0.36**	−0.31*
		(0.12)		(0.12)	(0.14)
Competence			0.098	0.23*	0.30*
			(0.092)	(0.11)	(0.13)
Observations	581	581	581	581	462
Estimator	Logit	Logit	Logit	Logit	Logit
Dep. Var.	OC	OC	OC	OC	OC
Sample	All	All	All	All	White Men
			Panel C: Percentage Overclaimed		
	(1)	(2)	(3)	(4)	(5)
Attractiveness	0.0021			−0.0043	−0.0054
	(0.0030)			(0.0029)	(0.0034)
Criminality		−0.0017		−0.0040	0.00058
		(0.0036)		(0.0036)	(0.0031)
Competence			0.0082*	0.010*	0.012*
			(0.0038)	(0.0041)	(0.0049)
Observations	636	636	636	636	504
Estimator	OLS	OLS	OLS	OLS	OLS
Dep. Var.	%Salary	%Salary	%Salary	%Salary	%Salary
Sample	All	All	All	All	White Males

Note: **Panel A:** The dependent variable is the logarithm of £1 + the total recommended repayment in the Legg report. Thus, MPs for whom no repayment was recommended are treated as having a repayment of £1. **Panel B:** The dependent variable is a binary variable taking value one for those MPs who had positive repayments recommended in the Legg report, and 0 otherwise. **Panel C:** The dependent variable is the total recommended repayment in the Legg report divided by MPs annual salary (£64,766). All specifications in each panel also include: Male; Age; White; Affilitative Smile; Reward Smile; and Party and Tenure Fixed Effects. Attractiveness, Criminality, and Competence are standardised to have mean 0 and SD 1.† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust Standard Errors in parentheses,

References

Agresti, A. (2010). *Analysis of Ordinal Categorical Data. Analysis of Ordinal Categorical Data: Second Edition.* Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9780470594001

Aidt, T. S. (2009). Corruption, institutions, and economic development. *Oxford Review of Economic Policy*, 25(2), 271–291. https://doi.org/10.1093/oxrep/grp012

Allen, N., & Birch, S. (2011). Political conduct and misconduct: probing public opinion. *Parliamentary Affairs*, *64*(1), 61-81.

Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. *European Journal of Social Psychology, 36*(4), 511-536. https://doi.org/10.1002/ejsp.324

Antonakis, J., & Eubanks, D. L. (2017). Looking leadership in the face. *Current Directions in Psychological Science*, *26*(3), 270–275. https://doi.org/10.1177/0963721417705888

Awamleh, R., & Gardner, W. L. (1999). Perceptions of leader charisma and effectiveness. *The Leadership Quarterly*, *10*(3), 345–373. https://doi.org/10.1016/S1048-9843(99)00022-3

Bailenson, J. N., Iyengar, S., Yee, N., & Collins, N. A. (2008). Facial similarity between voters and candidates causes influence. *Public Opinion Quarterly*, *72*(5), 935–961 https://doi.org/10.1093/poq/nfn064

Bagnis A, Caffo E, Cipolli C, De Palma A, Farina G, Mattarozzi K. Judging health care priority in emergency situations: Patient facial appearance matters. *Social Science & Medicine*, 260, 113180. doi: 10.1016/j.socscimed.2020.113180.

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences, 104*(46), 17948–17953. https://doi.org/10.1073/pnas.0705435104

Banchefsky, S., Westfall, J., Park, B., & Judd, C. M. (2016). But you don't look like a scientist!: Women scientists with feminine appearance are deemed less likely to be scientists. *Sex Roles*, *75*(3–4), 95–109. http://doi.org/10.1007/s11199-016-0586-1

Banducci, S. A., Karp, J. A., Thrasher, M., & Rallings, C. (2008). Ballot photographs as cues in low-information elections. *Political Psychology, 29*(6), 903–917. http://doi.org/10.1111/j.1467-9221.2008.00672.x

Bendahan, S., Zehnder, C., Pralong, F. P., & Antonakis, J. (2015). Leader corruption depends on power and testosterone. *The Leadership Quarterly*, 26(2), 101-122. https://doi.org/10.1016/j.leaqua.2014.07.010

Berggren, N., Jordahl, H., & Poutvaara, P. (2010). The looks of a winner: Beauty and electoral success. *Journal of Public Economics, 94*(1–2), 8–15. http://doi.org/10.1016/j.jpubeco.2009.11.002

Berggren, N., Jordahl, H., & Poutvaara, P. (2017). The right look: Conservative politicians look better and voters reward it. *Journal of Public Economics, 146*(53-83), 79–86 http://doi.org/10.1016/j.jpubeco.2016.12.008

Berinsky, A. J., Chatfield, S., & Lenz, G. (2019). Facial dominance and electoral success in times of war and peace. *The Journal of Politics, 81*(3), 1096–1100. https://doi.org/10.1086/703384

Berry, D. S., & Zebrowitz-McArthur, L. (1988). What's in a face?: Facial maturity and the attribution of legal responsibility. *Personality and Social Psychology Bulletin, 14*(1), 23–33. https://doi.org/10.1177/0146167288141003

Besley, T. (2005). Political Selection. *Journal of Economic Perspectives*, 19 (3), 43-60.

Besley, T., & Case, A. (1995). Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits. *The Quarterly Journal of Economics*, *110*(3), 769-798. https://doi.org/10.2307/2946699

Besley, T., & Larcinese, V. (2011). Working or shirking? Expenses and attendance in the UK Parliament. *Public Choice*, *146*(3–4), 291–317. https://doi.org/10.1007/s11127-009-9591-z

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of afrocentric facial features in criminal sentencing. *Psychological Science, 15*(10), 674–679. http://doi.org/10.1111/j.0956-7976.2004.00739.x

Blanken, I., van de Ven, N., & Zeelenberg, M. (2015). A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin*, *41*(4), 540–558. https://doi.org/10.1177/0146167215572134

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*(1), 1-45. https://doi.org/10.1037/0033-2909.88.1.1

Bruce, V. & Young, A.W. (2012). *Face Perception*. Hove: Psychology Press.

Burger, J. M. (1999). The foot-in-the-door compliance procedure: A multiple-process analysis and review. Personality and Social Psychology Review, 3(4), 303–325. https://doi.org/10.1207/s15327957pspr0304_2

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science*, *2*(1), 77–101. https://doi.org/10.1177/2515245918823199

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1651), 2651-2656. https://doi.org/10.1098/rspb.2008.0873

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behaviour. *Psychological Science*, *20*(10), 1194-1198. https://doi.org/10.1111/j.1467-9280.2009.02423.x

Chiao, J. Y., Adams Jr, R. B., Tse, P. U., Lowenthal, W. T., Richeson, J. A., & Ambady, N. (2008). Knowing who's boss: fMRI and ERP investigations of social dominance perception. *Group Processes & Intergroup Relations*, *11*(2), 201-214. https://doi.org/10.1177/1368430207088038

Chiao J. Y., Bowman N. E., & Gill, H. (2008) The political gender gap: Gender bias in facial inferences that predict voting behaviour. *PLoS ONE, 3*(10): e3666. https://doi.org/10.1371/journal.pone.0003666

Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behaviour. *Personality and Social Psychology Bulletin*, *38*(7), 907-919. https://doi.org/10.1177/0146167212442394

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285–290. http://doi.org/10.1037/h0033731

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2015). Social class, power, and selfishness: When and why upper and lower class individuals behave unethically. *Journal of Personality and Social Psychology*, *108*(3), 436–449. https://doi.org/10.1037/pspi0000008

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*(5), 383–386. http://doi.org/10.1111/j.1467-9280.2006.01716.x

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2019). Gender stereotypes have changed: A cross-temporal meta-analysis of US public opinion polls from 1946 to 2018. *American Psychologist, 75*(3), 301-315. https://doi.org/10.1037/amp0000494

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology*, *6*, 32-35. https://doi.org/10.1016/j.copsyc.2015.03.017

Enelow, J. M., & Hinich, M. J. (1982). Nonspatial candidate characteristics and electoral competition. *The Journal of Politics*, *44*(1), 115-130. https://doi.org/10.2307/2130286

Farrell, D. M., McAllister, I., & Studlar, D. T. (1998). Sex, money and politics: sleaze and the Conservative party in the 1997 election. *British Elections & Parties Review*, *8*(1), 80-94.

Flowe, H. D. & Humphries, J. E. (2011). An examination of criminal face bias in a random sample of police lineups. *Applied Cognitive Psychology, 25*, 265-273. http://doi.org/10.1002/acp.1673.

Flowe, H. D. (2012). Do characteristics of faces that convey trustworthiness and dominance underlie perceptions of criminality? *PLoS ONE*. 7(6): e37253. http://doi.org/10.1371

Flowe, H. D., Klatt, T., & Coloff, M. F. (2014). Selecting fillers on emotional appearance improves lineup identification accuracy. *Law and Human Behavior, 38*, 509-519. http://doi.org/10.1037/lhb0000101.

Foo, Y. Z., Loncarevic, A., Simmons, L. W., Sutherland, C. A. M., & Rhodes, G. (2019). Sexual unfaithfulness can be judged with some accuracy from men's but not women's faces. *Royal Society*

Open Science, *6*(4), 181552. https://doi.org/10.1098/rsos.181552

Frank, M. G., Ekman, P., & Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of Personality and Social Psychology*, *64*(1), 83-93. https://doi.org/10.1037//0022-3514.64.1.83

Gawronski, B., & Strack, F. (Eds.). (2012). *Cognitive Consistency: A Fundamental Principle in Social Cognition*. London: The Guilford Press.

Gneezy, A., Imas, A., Brown, A., Nelson, L. D., & Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behaviour. *Management Science*, 58(1), 179-187.

Gorst-Rasmussen, A., Dahm, C. C., Dethlefsen, C., Scheike, T., & Overvad, K. (2011). Exploring dietary patterns by using the treelet transform. *American Journal of Epidemiology, 173(*10), 1097–104. http://doi.org/10.1093/aje/kwr060

Graham, J. R., Harvey, C. R., & Puri, M. (2017). A corporate beauty contest. *Management Science*, *63*(9), 3044-3056. https://doi.org/10.1287/mnsc.2016.2484

Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science*, 862-886. http://doi.org/10.2307/2669329

Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the labor market. *The American Economic Review*, *84*(5), 1174-1194.

Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, *78*(5), 837-852. https://doi.org/10.1037/0022-3514.78.5.837

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences, 279*(1728), 571-576. https://doi.org/10.1098/rspb.2011.1193

Henn, M., & Foard, N. (2012). Young people, political participation and trust in Britain. *Parliamentary affairs*, *65*(1), 47-67.

Hirsh, J. B., Galinsky, A. D., & Zhong, C. B. (2011). Drunk, powerful, and in the dark: How general processes of disinhibition produce both prosocial and antisocial behaviour. *Perspectives on Psychological Science, 6*(5), 415-427. https://doi.org/10.1177/1745691611416992

Ipsos-Mori. (2016). "*Now I will read you a list of different types of people. For each would you tell me if you generally trust them to tell the truth, or not -- Politicians?* " 14th October – 1st November, 2016 in *Veracity Index*. Retrieved from https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/Polls/ipsos-mori-veracity-index-2016-charts.pdf

Joosten, A., Van Dijke, M., Van Hiel, A., & De Cremer, D. (2014). Feel good, do-good!? On consistency and compensation in moral self-regulation. *Journal of Business Ethics*, *123*(1), 71-84. https://doi.org/10.1007/s10551-013-1794-z

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology : A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of*

Personality and Social Psychology, 103(1), 54–69. https://doi.org/10.1037/a0028347

Keating, C. F., Mazur, A., & Segall, M. H. (1977). Facial gestures which influence the perception of status. *Sociometry*, *40*(4), 374-378. https://doi.org/10.2307/3033487

King, A., & Leigh, A. (2009). Beautiful politicians. *Kyklos, 62*(4), 579–593. http://doi.org/10.1111/j.1467-6435.2009.00452.x

Korndörfer, M., Egloff, B., & Schmukle, S. C. (2015). A large scale test of the effect of social class on prosocial behaviour. *PloS one*, 10(7), e0133193. https://doi.org/10.1371/journal.pone.0133193

Kristofferson, K., White, K., & Peloza, J. (2014). The nature of slacktivism: How the social observability of an initial act of token support affects subsequent prosocial action. Journal of Consumer Research, 40(6), 1149-1166.

Krumhuber, E., Manstead, A. S. R., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behaviour. *Emotion, 7*(4), 730-735. https://doi.org/10.1037/1528-3542.7.4.730

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*(3), 390–423. https://doi.org/10.1037/0033-2909.126.3.390

Lawson, C., Lenz, G. S., Baker, A., & Myers, M. (2010). Looking like a winner: Candidate appearance and electoral success in new democracies. World Politics, 62(4), 561–593. http://doi.org/10.1017/S0043887110000195

Lee, A. B., Nadler, B., & Wasserman, L. (2008). Treelets—An adaptive multi-scale basis for sparse unordered data. *The Annals of Applied Statistics, 2*(2), 435–471. http://doi.org/10.1214/07-AOAS137

Lin, C., Adolphs, R., & Alvarez, R. M. (2018). Inferring Whether Officials Are Corruptible From Looking at Their Faces. Psychological Science, 29(11), 1807–1823. https://doi.org/10.1177/0956797618788882

Mattes, K., & Milazzo, C. (2014). Pretty faces, marginal races: Predicting election outcomes using trait assessments of British parliamentary candidates. *Electoral Studies*, 34, 177–189. https://doi.org/10.1016/j.electstud.2013.11.004

Mullen, E., & Monin, B. (2016). Consistency versus licensing effects of past moral behaviour. *Annual Review of Psychology*, 67, 363-385. https://doi.org/10.1146/annurev-psych-010213-115120

Mobius, Markus, M., and Tanya S. Rosenblat. 2006. Why Beauty Matters. American Economic Review, 96 (1): 222-235. 10.1257/000282806776157515

Norris, P. (2011). *Democratic deficit: Critical citizens revisited.* Cambridge, Massachusetts: Cambridge University Press.

Olivola, C. Y., Eastwick, P., Finkel, E., Ariely, D., & Todorov, A. (2011). "A picture is worth a thousand inferences: Appearance-based first impressions predict leader selection and mate choice", in NA -

Advances in Consumer Research Volume 39, eds. Ahluwalia, R., Chartrand, T. L., & Ratner, R. K., Duluth, MN : Association for Consumer Research, Pages: 276-277.

Olivola, C. Y., Eubanks, D. L., & Lovelace, J. B. (2014). The many (distinctive) faces of leadership: Inferring leadership domain from facial appearance. *The Leadership Quarterly*, *25*(5), 817-834. https://doi.org/10.1016/j.leaqua.2014.06.002

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, *18*(11), 566–570. https://doi.org/10.1016/j.tics.2014.09.007

Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., & Todorov, A. (2012). Republicans prefer Republican-looking leaders. *Social Psychological and Personality Science, 3*(5), 605–613. http://doi.org/10.1177/1948550611432770

Olivola, C. Y., Tingley, D., & Todorov, A. (2018). Republican voters prefer candidates who have conservative-looking faces: New evidence from exit polls. *Political Psychology*, *39*(5), 1157-1171.

Olivola, C. Y., & Todorov, A. (2010a). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, *34*(2), 83-110. http://doi.org/10.1007/s10919-009-0082-1

Olivola, C. Y., & Todorov, A. (2010b). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*(2), 315-324. https://doi.org/10.1016/j.jesp.2009.12.002

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Pattie, C., & Johnston, R. (2012). The electoral impact of the UK 2009 MPs' expenses scandal. *Political Studies, 60*(4), 730–750. http://doi.org/10.1111/j.1467-9248.2011.00943.x

Poutvaara, P., Jordahl, H., & Berggren, N. (2009). Faces of politicians: Babyfacedness predicts inferred competence but not electoral success. *Journal of Experimental Social Psychology*, *45*(5), 1132-1135. https://doi.org/10.1016/j.jesp.2009.06.007

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behaviour. *PLoS ONE*, *7*(3), e34293. https://doi.org/10.1371/journal.pone.0034293

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7, 331–363. doi:10.1037/1089-2680.7.4.331

Rosar, U., Klein, M., & Beckers, T. (2008). The frog pond beauty contest: Physical attractiveness and electoral success of the constituency candidates at the North Rhine-Westphalia state election of 2005. *European Journal of Political Research, 47*(1), 64–79. http://doi.org/10.1111/j.1475-6765.2007.00720.x

Rosenblat, T. S. (2008). The Beauty Premium: Physical Attractiveness and Gender in Dictator Games. *Negotiation Journal*, 24(4), 465–481. https://doi.org/10.1111/j.1571-9979.2008.00198.x

Rule, N. O., & Ambady, N. (2008). The face of success. *Psychological Science, 19*(2), 109–111. http://doi.org/10.1111/j.1467-9280.2008.02054.x

Rule, N. O., Ambady, N., Adams Jr, R. B., Ozono, H., Nakashima, S., Yoshikawa, S., & Watabe, M. (2010). Polling the face: prediction and consensus across cultures. *Journal of Personality and Social Psychology*, *98*(1), 1-15. https://doi.org/10.1037/a0017673

Rush, M., & Giddings, P. (2011). *Parliamentary Socialisation. Parliamentary Socialisation*. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230316850_7

Rychlowska, M., Jack, R. E., Garrod, O. G., Schyns, P. G., Martin, J. D., & Niedenthal, P. M. (2017). Functional smiles: Tools for love, sympathy, and war. *Psychological Science*, *28*(9), 1259-1270. https://doi.org/10.1177/0956797617706082

Shea, C. M., & Howell, J. M. (1999). Charismatic leadership and task feedback: A laboratory study of their effects on self-efficacy and task performance. *The Leadership Quarterly*, *10*(3), 375–396. https://doi.org/10.1016/S1048-9843(99)00020-X

Smart, M., & Sturm, D. M. (2013). Term limits and electoral accountability. *Journal of Public Economics*, *107*, 93-102. https://doi.org/10.1016/j.jpubeco.2013.08.011

Spisak, B. R., Blaker, N. M., Lefevre, C. E., Moore, F. R., & Krebbers, K. F. (2014). A face for all seasons: Searching for context-specific leadership traits and discovering a general preference for perceived health. *Frontiers in Human Neuroscience*, *8*, 792. https://doi.org/10.3389/fnhum.2014.00792

Spisak, B. R., Dekker, P. H., Krüger, M., & van Vugt, M. (2012). Warriors and peacekeepers: Testing a biosocial implicit leadership hypothesis of intergroup relations using masculine and feminine faces. *PLoS ONE, 7,* e30399. https://doi.org/10.1371/journal.pone.0030399

Spisak, B. R., Homan, A. C., Grabo, A., & van Vugt, M. (2012). Facing the situation: Testing a biosocial contingency model of leadership in intergroup relations using masculine and feminine faces. *The Leadership Quarterly*, *23*(2), 273-280. https://doi.org/10.1016/j.leaqua.2011.08.006

Stoker, J. I., Garretsen, H., & Spreeuwers, L. J. (2016). The facial appearance of CEOs: Faces signal selection but not performance. *PLOS ONE*, *11*(7), e0159950. https://doi.org/10.1371/journal.pone.0159950

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*(5728), 1623–1626. http://doi.org/10.1126/science.1110589

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on

social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460. https://doi.org/10.1016/j.tics.2008.10.001

Trichas, S., & Schyns, B. (2012). The face of leadership: Perceiving leaders from facial expression. *The Leadership Quarterly*, *23*(3), 545–566. https://doi.org/10.1016/j.leaqua.2011.12.007

Trichas, S., Schyns, B., Lord, R., & Hall, R. (2017). "Facing" leaders: Facial expression and leadership perception. *The Leadership Quarterly*, *28*(2), 317–333. https://doi.org/10.1016/j.leaqua.2016.10.013

Valla, J. M., Ceci, S. J., & Williams, W. M. (2011). The accuracy of inferences about criminality based on facial appearance. *Journal of Social, Evolutionary, and Cultural Psychology, 5*(1), 66–91. doi: 10.1037/h0099274

Ward, C., Flowe, H. D., & Humphries, J. E. (2012). The effects of masculinity and suspect gender on perceptions of guilt. *Applied Cognitive Psychology, 26*, 482-488. http://doi.org/10.1002/acp.2823.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7), 592–598. http://doi.org/10.1111/j.1467-9280.2006.01750.x

Wang, D., Nair, K., Kouchaki, M., Zajac, E. J., & Zhao, X. (2019). A case of evolutionary mismatch? Why facial width-to-height ratio may not predict behavioral tendencies. *Psychological Science*, 30(7), 1074-1081. https://doi.org/10.1177/0956797619849928

Wong, E. M., Ormiston, M. E., & Haselhuhn, M. P. (2011). A face only an investor could love: CEOs' facial structure predicts their firms' financial performance. *Psychological Science*, *22*(12), 1478-1483. https://doi.org/10.1177/0956797611418838

Zebrowitz, L. A. (2018). *Reading faces: Window to the soul?* New York: Routledge. https://doi.org/10.4324/9780429493188

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, *2*(3), 1497–1517. https://doi.org/10.1111/j.1751-9004.2008.00109.x

Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998). The relationship between appearance and personality across the life span. *Personality and Social Psychology Bulletin*, *24*(7), 736–749. https://doi.org/10.1177/0146167298247006

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, *22*(12), 1258–1269. https://doi.org/10.1177/01461672962212006

Zebrowitz, L.A. & McDonald, S.M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior, 15*(6), 603-623. https://doi.org/10.1007/BF01065855

Zhong, C.-B., Liljenquist, K., & Cain, D. M. (2009). Moral self-regulation: Licensing and compensation. In D. De Cremer (Ed.), *Psychological perspectives on ethical behaviour and decision making* (pp. 75–89). Charlotte, North Carolina: Information Age Publishing.