Short communication

# Is a disease leader attractive? Six tests of whether the COVID-19 pandemic affected follower preferences for attractiveness, health and other traits in political and non-political leaders

Lasse Laustsen [a,*], Asmus Leth Olsen [b]

[a] Centre for the Experimental-Philosophical Study of Discrimination, Department of Political Science, Aarhus University, Denmark
[b] Department of Political Science, University of Copenhagen, Denmark

## ARTICLE INFO

## ABSTRACT

Attractive political candidates receive more votes on Election Day compared to their less attractive competitors. One well-cited theoretical account for this attractiveness effect (White et al., 2013) holds that it reflects an adaptive psychological response to disease threats. Voters are predicted to upregulate preferences for attractiveness because it constitutes a cue to health. The global COVID-19 pandemic constitutes an ecologically relevant and realistic setting for further testing this prediction. Here, we report the results from six tests of the prediction based on two large and nationally representative surveys conducted in Denmark (n = 3297) at the outbreak of the pandemic and one year later. Utilizing experimental techniques, validated individual difference measures of perceived disease threat and geographic data on COVID-19 severity, we do not find that disease threats like the COVID-19 pandemic upregulate preferences for attractive and healthy political or non-political leaders. Instead, respondents display heightened preferences for health in socially proximate relations (i.e. colleagues). Moreover, individuals who react aversively to situations involving risks of pathogen transmission (scoring high in Germ Aversion) report higher importance of a wide range of leadership traits, rather than for health and attractiveness in particular. Results are discussed in relation to evolutionary accounts of leadership and followership.

The COVID-19 crisis has caused widespread implementation of comprehensive and hitherto unprecedented government regulation of citizen mobility and freedom of assembly, affecting the daily routines of hundreds of millions of citizens across the globe. According to recent research on the behavioral immune system, disease threats like COVID-19 also hold great potential for affecting human social behavior and attitudes across a wide range of domains (Hartman et al., 2020; Schaller, Murray, & Bangerter, 2015; Schaller, 2016; Tybur, Lieberman, Kurzban, & DeScioli, 2012). In this article, we focus on one such domain: leader preferences.

A vast number of studies find that leaders and political candidates benefit from evoking a positive first impression based on their physical appearance (Antonakis & Dalgas, 2009; Antonakis & Eubanks, 2017; Giacomin & Rule, 2020; Little, 2014; Todorov, 2017; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). In particular, physical attractiveness is shown to positively affect candidates' electoral success (e.g. Berggren, Jordahl, & Poutvaara, 2010; Rosar, Klein, & Beckers, 2008; Laustsen, 2014). One popular explanation for this effect is the so-called "attractiveness halo effect": the tendency to ascribe positive attributes to attractive individuals (Langlois et al., 2000; Verhulst, Lodge, & Lavine, 2010). However, recent evolutionary accounts of leadership and followership—sometimes referred to as adaptive followership psychology—offer an alternative explanation for followers' preferences for attractive leaders (e.g. Van Vugt, 2006; von Rueden & van Vugt, 2015; Laustsen, 2021). A central premise of these theories is that humans evolved to prefer different individuals as their leader depending on the kinds of problems confronting their group. In particular, anthropological records as well as experimental studies reveal that followers generally prefer more masculine, dominant and strong leaders in contexts characterized by intergroup conflict (Laustsen & Petersen, 2015, 2016, 2020). Likewise, human followership psychology might also react to other evolutionarily recurring problems and contexts, causing specific trait preferences in leaders. One type of evolutionarily relevant context that could have molded human followership psychology is situations of disease threats such as the COVID-19 pandemic. Based on such evolutionary rationales, White et al. (2013) suggest that followers might not prefer attractive individuals because of a "halo effect." Rather, they prefer

attractive leaders as a consequence of evolved psychological mechanisms related to disease avoidance and followership. Following this line of reasoning, White et al. (2013) predict that because attractiveness constitutes a cue to health, preferences for attractive leaders—and political candidates—will also increase as a function of disease threat. Moreover, because "group members are relatively more dependent on leaders than on other group members," White et al. (2013: p. 2430) predict that preferences for health and attractiveness rise more in leaders than in other social categories when disease threats intensify. Importantly, White et al. (2013) find support for their prediction across four studies using both experimental primes and geographical variance in disease threat as well as different outcome measures.

In this article, we test whether the results from White et al. (2013) generalize and replicate in relation to the disease threat instilled by the COVID-19 pandemic. Originally, White et al. (2013) tested their theory and its predictions in relation to macro-level indicators of pathogen and disease threat (Study 1 in White et al. (2013)), verbal depictions in a *disease-threat story* (Studies 2 and 3), and visual cues of disease threat through photos (Study 4). Because we wanted to test the theory concretely in connection with the ongoing global pandemic, we tested whether experimental treatments priming the COVID-19 pandemic (compared to control conditions asking subjects to think of pre- and post-COVID-19 times) would also cause heightened preferences for attractive and healthy leaders. In other words, we aimed to test White et al. (2013) theory at the height of a global health crisis, thus maximizing its ecological relevance. Moreover, our studies also contribute in other important ways to uncovering and testing how disease threats potentially mold leader preferences. First, we report results from experimental studies conducted with well-powered, large and representative samples of Danish citizens. That is, in contrast to the original studies by White et al. (2013), we move beyond the use of relatively small convenience samples, permitting us to test the key theoretical prediction on more diverse, heterogeneous and much larger samples. Second, testing the theory—that disease threats upregulate preferences for attractive leaders—outside the US setting directly addresses its generalizability across countries with different histories of disease threat, institutional settings and political systems. Finally, to supplement our experimental conditions and geographical indicators of disease threat, one of our surveys included a measure of subjects' Perceived Vulnerability to Disease (Duncan, Schaller, & Park, 2009) —a key individual difference capturing perceived vulnerability to infectious diseases. Consequently, we can test whether subjects display stronger preferences for attractive and healthy leaders the more vulnerable they perceive themselves in connection to disease threats.

Below, we first provide an overview of our available data sources, consisting of two surveys composed of various experimental tests conducted at two different time points during the COVID-19 pandemic. Next, we report key results obtained across our six tests. Finally, the general discussion addresses the total amount of evidence for the claim that disease threat upregulates preferences for attractive leaders, both in concrete connection with the COVID-19 pandemic and in general, and discusses theoretical implications of the results.

## Overview of data, studies and tests

To test the prediction that disease threats like the COVID-19 pandemic might upregulate preferences for attractive and healthy leaders, we conducted two surveys, each containing two different survey experiments as well as data on local variation with respect to COVID-19 severity. Both surveys were based on large and representative samples of the adult (ages 18–65) Danish population, collected in collaboration with the Danish polling company Epinion. This section describes the time at which our surveys were in the field, discusses

Denmark as the test country for the theoretical prediction, and presents how the different experiments contained in Surveys 1 and 2 comprise a solid point of departure for testing whether the threat from COVID-19 enhances preferences for attractive and healthy leaders.

### Timing of data collection

The two surveys were fielded at two different points in time during the COVID-19 pandemic. Survey 1 was collected during the early weeks of the first shutdown of Danish society, beginning March 13, 2020. This lockdown involved sending non-essential public employees home and closing educational institutions to curb the initial spread of coronavirus. In the following week a ban on public meetings of more than 10 people was introduced, with bars and restaurants also closing down. Survey 1 was collected from March 20 to 27, 2020, and in this period additional restrictions were introduced and existing restrictions were extended. Testifying to the perceived severity of the situation, our subjects reported an average of 8.5 on a 0–10 scale (10 reflecting "Most severe threat in my lifetime") when asked how severely they perceived the threat from coronavirus. Survey 2 was conducted one year later, from March 20 to 31, 2021. At this time, the second shutdown of Danish society had been in place for several months, beginning December 18, 2020. Comparatively, this second lockdown reflects a qualitatively different setting, with Danes in Survey 2 reporting a lower perceived threat from coronavirus (average of 7.4, same scale as above). Furthermore, the second shutdown was characterized by increased political polarization related to the necessary level of restrictions, and a growing public fatigue caused by limited mobility, restrictions on gatherings, etc. (Lindholt & Petersen, 2021). In total, the two settings surrounding Surveys 1 and 2 comprise a solid combination for testing how the disease threat from COVID-19 possibly affects leader preferences. While Survey 1 permits us to test the role of immediate threat responses, Survey 2 makes it possible to test the long-term consequences of living with constant disease threats and corresponding restrictions.

Despite the described strengths of these settings for testing the theoretical prediction, one potential challenge is also worth considering. As described in detail below (see "Experimental Tests and analyses across surveys 1 and 2"), our experiments involve random assignment of respondents to disease threat conditions (asking them to think of the COVID-19 crisis) or control conditions. However, the omnipresence of the COVID-19 threat surrounding our data collections might cause us to obtain null results (i.e. negligible differences between disease threat and control conditions) even if the theory is in fact correct. To ensure our ability to test this possibility, experiments in Survey 2 included manipulation checks of perceived risks in direct connection to disease threat conditions (because of length restrictions on Survey 1, such manipulation checks were not included). Moreover, if we were to obtain null effects across our experiments and such results were caused by the omnipresent COVID-19 threat, then we should expect exaggerated preferences for attractiveness and health in leaders across our different tests. Hence, despite the possibility that the COVID-19 pandemic pre-treated our subjects to hold exaggerated disease threat perceptions, we see the presence of the COVID-19 crisis as a strength rather than a weakness of our research design. As reported above, subjects in both surveys were clearly aware of risks related to disease spread, and thus our studies permit us to test how well White et al. (2013) prediction that disease threat upregulates preferences for attractive and healthy leaders generalizes to the context of the global COVID-19 pandemic.

### Denmark as test country

White et al. (2013) tested how disease threats affected preferences for attractive and healthy leaders among American subjects who directly elect political leaders based on first-past-the-post electoral sys-

tems; that is, the candidate who receives the largest vote share wins the election. In addition to testing the prediction concretely in connection with the COVID-19 pandemic, we also move the test to another institutional setting. Denmark employs a proportional electoral system in local and national elections, in which each party nominates a series of candidates in each municipality/district. The party then receives the number of seats in the district that correspond to its vote share. Next, a party distributes its seats to candidates in accordance with personal votes.[1] Consequently, Danish voters choose to vote for either a party or a candidate from a party (in this case, the vote first counts at the party level and subsequently at the candidate level within the party). According to Statistics Denmark, approximately 50 and 75 percent of voters cast a personal vote in national and local elections, respectively. Thus, while Danish elections are probably less candidate-centered than American elections, Danes do to a large degree vote for specific candidates. Testifying to a growing focus on political personas, the Danish National Election Study also includes ratings of party leaders in recent election surveys (data and surveys are online available via www.valgprojektet.dk). That is, political candidates play a major role in Danish elections, attracting voters to or repelling them from the parties they represent. Moreover, recent studies also find that candidates' visual appearances—as measured by facial attractiveness, competence and dominance—matter in Danish elections (Laustsen & Petersen, 2016, 2020; Laustsen, 2014) as well as in electoral contests resembling the Danish (Berggren et al., 2010; Rosar et al., 2008). In sum, based on these characteristics of the Danish electoral system and prior findings regarding appearance effects on electoral outcomes, we believe Denmark constitutes a valuable case for testing whether disease threats upregulate preferences for attractive and healthy political leaders. Moreover, to further test the scope of the theory outside the domain of politics, we follow White et al. (2013) and test whether the COVID-19 pandemic also affects preferences for attractiveness and health in non-political leaders. For these tests, political institutional differences between the United States and Denmark seem less relevant.

*Experimental tests and analyses across surveys 1 and 2*

Below, we report the results from six tests across two surveys, each of which contain two experiments. In the experiments, subjects were randomly assigned to either a disease threat condition—asking them to consider the ongoing threat from COVID-19—or a control condition not priming coronavirus. The six tests cover two distinct time periods and real-world contexts related to COVID-19. Tests 1–3 cover the immediate threat reaction following the first shutdown of Danish society. Tests 4–6 cover the situation one year later during the second shutdown, characterized by political disagreement and citizen fatigue. In each of the two surveys, one experiment explores subjects' explicit preferences for attractiveness and health in leaders (Tests 1 and 4) and another investigates subjects' subtle and implicit preferences for attractive leaders (Tests 2 and 5). The six tests also cover trait preferences related to political (Tests 1, 2 and 4) and non-political leaders (Tests 4 and 5). Finally, our surveys permit us to test the role played by individual differences in perceived disease threat in explaining leader trait preferences (Tests 4 and 5) as well as the role played by geographic variation in COVID-19 risks (Tests 3 and 6). Taken together, we believe that this package constitutes a solid empirical point of departure for detecting any meaningful effects of COVID-19 threat on preferences for attractive and healthy leaders.

Following recent trends and recommendations in the social sciences for openness and transparency (Antonakis, 2017; Nosek et al., 2019; Open Science Collaboration, 2015), we pre-registered our hypotheses using the Open Science Framework platform (Survey 1: https://mfr.osf.io/render?url = https://osf.io/7mjzh/?view_only =

e9cce66c1a56466eb9789b9daf4e9332%26direct%26mode = render% 26action = download%26mode = render; Survey 2: https://mfr.osf.io/ render?url = https://osf.io/mkvt7/?view_only = e9cce66c1a56466eb9 789b9daf4e9332%26direct%26mode = render%26action = download %26mode = render). Unless otherwise noted, results follow the pre-registrations. Throughout our analyses we employ an alpha-level of 5 percent following standard practices.[2] Finally, all statistical tests were conducted using Stata 14, while figures were created using R 4.0.2.

### Test 1: Disease threats and desired leadership characteristics at the outbreak of the COVID-19 pandemic

*Method*

*Design:* Test 1 investigated subjects' preferences for attractiveness and health in a Danish prime minister when they were reminded of the COVID-19 crisis and when they were not during the first lockdown of Danish society (March 20 to 27, 2020). Importantly, we contrasted this test with a similar test of preferences for attractiveness and health in a colleague. Consequently, Test 1 sheds light on whether the COVID-19 pandemic enhances preferences for attractiveness and health in leaders in particular or whether, instead, such preferences reflect wider social preferences across various relations. Specifically, we conducted a 2 × 2 between-subjects design, with subjects randomly assigned to state their trait preferences in a Danish prime minister or a colleague (target conditions) in either a COVID-19 or a control condition (context conditions). In addition to rated importance of attractiveness and health, subjects also rated the importance of competence, trustworthiness, dominance and power. Only predictions for attractiveness and health were pre-registered, but we also briefly report results for the remaining traits.

*Participants.* We surveyed 1782 Danes representative with respect to age (>18 years old; M = 49.9, SD = 18.5), gender (51.9 percent females), and geographical region. The Danish polling company Epinion fielded the survey among their financially compensated panelists. The sample size was determined based on budgetary constraints and is approximately eight times larger than the largest sample employed by White et al. (2013). With a 7-point scale employed as dependent variable and assuming a 1.8 standard deviation, we can capture the following effects with 90% power in a 2 × 2 design: effects as small as 0.3 scale points for main effects, a difference of 0.4 comparing two individual groups, and interaction effects of about 0.6 (see SOM S.1 for sample descriptives).[3]

*Procedure.* Participants were instructed to state the perceived importance of six traits in a prime minister or a colleague. Importantly, we randomly introduced the task to subjects by highlighting the disease threat related to the COVID-19 crisis (in brackets below). Specifically, the instruction read, "Below you see a set of questions regarding how important you find different characteristics to be in a **Danish prime minister/colleague** [in relation to the ongoing crisis caused by the spread of coronavirus]. Please read the different statements below and indicate how important you find each of the mentioned characteristics. [During the present infection crisis,] it is important that a Danish prime minister/colleague is…" Responses were collected on 1–7 scales reflecting "Not at all important" and "Very important," respectively. Across conditions, the rated importance of health

---

[1] However, some parties depart slightly from this procedure. For more detailed information on the Danish electoral system see Blom-Hansen, Elklit, Serritzlew, and Riis Villadsen (2016).

[2] Importantly, some of the pre-registered hypotheses involve multiple testing (e.g. predicting importance ratings of both health and attractiveness in Tests 1 and 4). Reassuringly, for these pre-registered hypotheses, substantial conclusions reported throughout the article remain unchanged when applying Bonferroni adjustments for multiple testing.

[3] Replication data and command files for Survey 1 is available via the Open Science Framework: https://osf.io/fzh8v/. Importantly, this article reports the results from Test 1, 2 and 3 in detail, while Laustsen (2021) also refers to the results as part of a larger review of the literature on context-sensitive trait preferences in leaders.
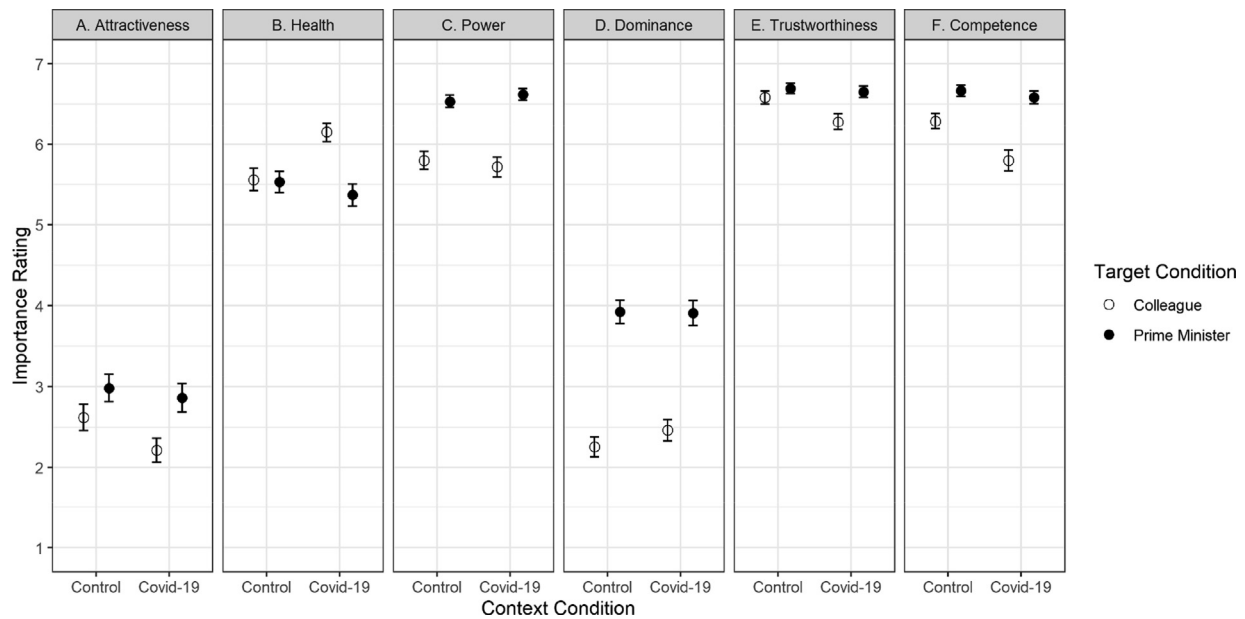
**Fig. 1.** Average rated importance of six traits in prime minister and colleague in Test 1 with 95 percent confidence intervals.

(M = 5.65, SD = 1.44) was approximately twice as large as for attractiveness (M = 2.67, SD = 1.80).[4]

*Results and discussion*

For attractiveness we found an non-significant negative difference between rated importance of attractiveness in a prime minister in the COVID-19 compared to the control condition (b = -0.12, SE = 0.12, p = 0.330). These results largely replicated when we predict preferences for health in a prime minister with a negative and non-significant difference (b = −0.16, SE = 0.10, p = 0.097). That is, these results do not support the prediction that the COVID-19 pandemic upregulates preferences for attractive and healthy leaders. Although not pre-registered, we also tested whether the COVID-19 condition changed the rated importance of competence, trustworthiness, dominance and power, but all tests yielded non-significant differences between the COVID-19 and control conditions (ps > 0.05). Average ratings across all traits are provided across conditions in Fig. 1.

Although importance of leader attractiveness and health did not increase as a function of disease threat, we still formally tested whether the COVID-19 condition enhances preferences for attractiveness and health more strongly in a prime minister compared to a colleague. Predicting attractiveness importance from assigned target condition, context condition and their interaction, we found a positive and non-significant interaction (b = 0.29, SE = 0.17, p = 0.090). Next, we also predicted health importance from assigned target condition, context condition and their interaction. This revealed a significant and strong *negative* interaction (b = -0.75, SE = 0.13, p < 0.001). In other words, the COVID-19 condition caused subjects to weight health much more when evaluating a colleague than a prime minister. This pattern is primarily caused by a very strong effect of the COVID-19 condition on importance of health in a colleague (b = 0.59, SE = 0.09, p < 0.001), corresponding to a medium-sized effect with Cohen's *d* = 0.43 (see SOM S.3 for full models). In addition, exploratory analyses for the remaining traits found positive and significant interactions between target condition and context condition for com-

petence (b = 0.41, SE = 0.10, p < 0.001) and trustworthiness (b = 0.26, SE = 0.08, p = 0.002), and non-significant interactions for power and dominance (ps > 0.05). For both competence and trustworthiness, the positive interaction is driven by the reduced importance of the traits in the COVID-19 compared to the control condition for a colleague (see SOM S.4 for full models for competence, trustworthiness, power and dominance).

Across the reported analyses, three main patterns stand out. First, the non-significant (and negative) effect of the COVID-19 condition on the importance of attractiveness and health in a prime minister does not support the underlying theoretical rationale when tested in connection with the COVID-19 pandemic. Second, comparing the effect of the COVID-19 condition for the prime minister and colleague targets, respectively, provides further support *against* the notion that COVID-19 shapes perceived importance of attractiveness in leaders in particular. While the COVID-19 condition only non-significantly affected perceived importance of attractiveness and health in a prime minister, it significantly moved the importance of attractiveness (negatively) and health (positively) in a colleague. Thus, while Test 1 did not support the theoretical expectation, the analyses revealed that subjects did react to the COVID-19 condition, given that the rated importance of health in a colleague was significantly higher in the COVID-19 condition compared to the control condition. A related concern is that the omnipresence of the COVID-19 pandemic surrounding our experiment had already heightened preferences for health and attractiveness in leaders from the outset. This does not seem to be the case, with the rated importance of prime minister attractiveness and health around 3 and 5.5 (on 1–7 scales), respectively. These results suggest that the experimental null findings for prime minister ratings are likely not caused by failed manipulation of disease threat in our experiment, nor by exceptionally high preferences for an attractive and healthy prime minister caused by the general COVID-19 situation outside our experiment. Third, the non-significant results from the exploratory analyses for rated importance of competence, trustworthiness, dominance and power in a prime minister further suggest that the COVID-19 condition generally did not greatly affect prime minister trait preferences.

Because Test 1 relied on explicitly stated preferences for attractiveness in a political leader, it is possible that social desirability bias may dampen subjects' inclinations to express such preferences explicitly. Test 2 therefore investigates whether the COVID-19 pandemic height-

---

[4] Because the underlying theory for our predictions builds on attractiveness being a cue to health, we tested whether ratings of the traits were correlated. The correlation is significant for both targets and also substantially strong for prime minister: $r_{\text{colleague}} = 0.07$, $p = 0.031$; $r_{\text{Prime Minister}} = 0.33$, $p < 0.001$. See SOM S.2 for correlations across all six trait ratings separately for the two targets.

ens preferences for attractive leaders using a more subtle manipulation of attractiveness through visual appearance.

## Test 2: disease threats, candidate facial attractiveness and competence evaluations at the outbreak of the COVID-19 pandemic

*Method*

*Design:* Test 1 found no support for upregulated importance of attractiveness and health in leaders as a function of disease threat. However, it could be that asking subjects about their explicit trait preferences gets at the theoretical phenomenon of interest too blatantly—especially if these preferences work at a more subtle level. To test this, we report the results from a second experiment. Specifically, we conducted another $2 \times 2$ between-subjects experiment, asking subjects to rate how well-suited for being prime minister a given target individual low or high in attractiveness (attractiveness conditions) seemed. Moreover, subjects stated their impression in relation to the current COVID-19 crisis or in a neutral scenario (context conditions).

*Participants.* Subjects were the same 1782 Danes who participated in the first experiment (see Test 1). Conditions across the two experiments were assigned at random to make sure that any sequence effects canceled out. However, the experiment for Test 2 was always presented after the trait rating experiment (Test 1) with unrelated filler questions in between (see Test 1 for power calculations).

*Procedure.* Based on first impressions, participants were instructed to state how well-suited they found the assigned target individual to be for the job of Danish prime minister. Importantly, we introduced the task to subjects by randomly highlighting the COVID-19 crisis (in brackets below). Specifically, the instruction read "Now we are interested in your first impressions about the appearance of a good leader. We ask you to imagine that the person shown on the screen below is prime minister in Denmark [during the current health crisis caused by the spread of coronavirus]. How good a prime minister does the person in your opinion appear to be [in this situation]?" Responses were collected on a 1–7 scale, reflecting "Very bad" and "Very good" respectively (M = 3.38, SD = 1.24). The attractiveness condition randomly assigned subjects to evaluate a low- or high-attractiveness target as prime minister. The target faces were collected from the Chicago face database (Ma, Correll, & Wittenbrink, 2015), which contains front face target faces that are pre-rated on relevant characteristics such as attractiveness. Due to survey space restrictions, we only included one low- and one high-attractiveness target, respectively. To hold constant age, gender and ethnicity, we chose Caucasian male targets with almost identical perceived age (30.9 and 28.6 years, respectively). Furthermore, we sought to maximize rated attractiveness for the targets through our choice of faces—2.65 and 4.59 on a 1–7 scale for the low and high target, respectively—while keeping other factors as constant as possible (see SOM S.5. for information about used face materials). We designed the experiment based on a difference-in-differences logic, which compares ratings of the high-attractiveness target face in the COVID-19 condition to ratings in the control condition and compares this difference to the same contextually induced rating difference for the low-attractiveness target face (following the design logic of Study 2 in Bøggild & Laustsen, 2016). That is, if the COVID-19 pandemic enhances preferences for attractive leaders, the high-attractiveness target should receive better ratings as prime minister than the low-attractiveness target *and* the COVID-19 condition should enhance this difference compared to the control condition.

*Results and discussion*

We first tested whether the high-attractiveness target received preferential evaluations with regard to prime minister potential.

Results show that the low- and high-attractiveness targets received ratings of 2.99 and 3.78, respectively, which also constitutes a statistically significant difference (b = 0.79, SE = 0.06, p < 0.001). Consequently, in line with a vast number of published studies (e.g. Berggren et al., 2010; Verhulst et al., 2010), attractiveness enhances subjects' leadership evaluations of these unknown target leaders. Second, we tested the key prediction that the preferential leadership evaluation of the attractive target is particularly strong in the COVID-19 condition compared to the control condition. However, the two-way interaction between the attractiveness condition and the context condition remains non-significant (b = 0.07, SE = 0.11, p = 0.531). That is, the preferential rating of the high- over the low-attractiveness target is indistinguishable in the COVID-19 (b = 0.83, SE = 0.08, p < 0.001) and the control conditions (b = 0.76, SE = 0.07, p < 0.001) (See SOM S.6 for full models). This is further illustrated in Fig. 2.

One main pattern stands out across Tests 1 and 2: Experimentally induced COVID-19 threats did not increase preferences for attractiveness in leaders at the outbreak of the COVID-19 pandemic in Denmark. However, it might be that subjects' preferences for leader attractiveness and health respond to geographic variation in COVID-19 threat. Test 3 investigates this possibility across the experiments reported under Tests 1 and 2.

## Test 3: Geographical threat intensity and leader preferences at the outbreak of the COVID-19 pandemic

*Method*

*Design:* The central organizational units of the Danish healthcare system are the five Danish regions. During the first phase of the COVID-19 pandemic, Danish health authorities also communicated geographical variance in disease intensity to the public based on these regions. At later stages of the pandemic, health authorities were able to provide information about disease intensity at a more local and, thus, fine-grained scale (see Test 6). However, limited by data availability from this initial stage of the pandemic, we focus on regional dif-
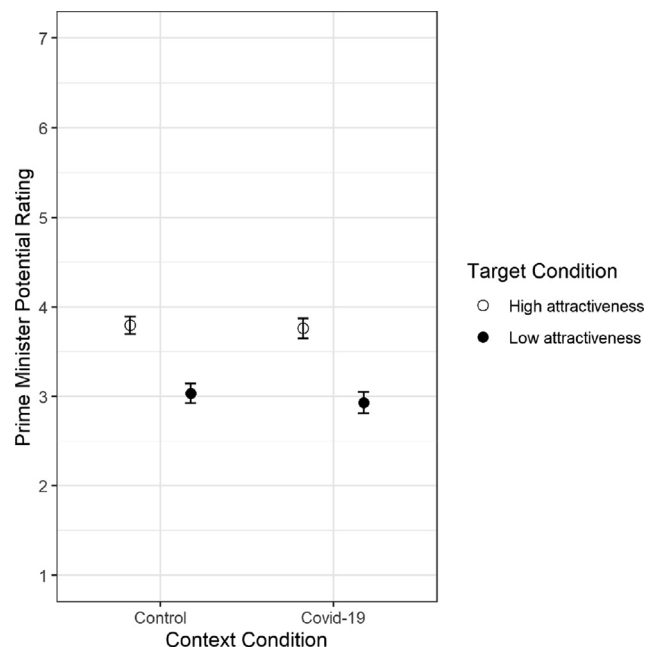


**Fig. 2.** Average rated prime minister potential across experimental conditions from Test 2 with 95 percent confidence intervals.

ferences. The capital region (Greater Copenhagen) quickly emerged as the most severely affected region, with more than 50% of all cases, intensive care patients and deaths while only 30% of the total population live there (Danish Statistics, 2021).

*Participants.* Subjects were again identical to the 1782 representative Danes used for Tests 1 and 2. Of these, 565 resided in Greater Copenhagen, with the remaining 1217 subjects living throughout the other four Danish regions.

*Procedure.* As described above, Greater Copenhagen quickly emerged as the region most intensely affected by COVID-19. This holds across standard metrics of disease severity such as number of reported COVID-19 deaths and numbers of total and intensive care hospitalizations. Consequently, our preregistered hypotheses entail that in the context of the trait-rating experiment (Test 1), we compare subjects' stated importance of attractiveness and health as leader characteristics for residents of Greater Copenhagen versus the remaining regions. Likewise, in the context of the visual attractiveness experiment (Test 2), we compared subjects' ratings of prime minister potential for the attractive versus the less attractive targets across subjects residing in Greater Copenhagen versus any other region. When estimating regional differences we also control for gender, age and education (although results remain unchanged without the control variables).

*Results and discussion*

We first investigated regional differences in relation to rated importance of attractiveness and health as leader characteristics (Test 1). Subjects residing in Greater Copenhagen did not rate either attractiveness ($b = 0.01$, $SE = 0.13$, $p = 0.931$) or health ($b = -0.13$, $SE = 0.10$, $p = 0.203$) as more important in a prime minister than subjects who resided outside the Danish capital. Moreover, preferences for attractiveness and health in a colleague were unaffected by regional residence. Finally, exploratory analyses reveal non-significant regional differences in trait ratings for prime minister and colleague across the four remaining traits, with only one exception. Subjects residing in Greater Copenhagen reported reduced importance for prime minister power compared to subjects in the rest of Denmark (see SOM S.7.1. for full models).

Next, we tested whether ratings of prime minister potential between the low- and high-attractiveness targets (obtained as part of the experiment for Test 2) were larger for Copenhageners compared to subjects residing elsewhere. Again, results are unsupportive of the prediction, with the two-way interaction between attractiveness condition (low vs. high) and geographical residence (Greater Copenhagen vs. rest of Denmark) remaining non-significant ($b = 0.04$, $SE = 0.12$, $p = 0.720$). Furthermore, geographical residence remained non-significantly related to ratings of both the low- ($b = -0.04$, $SE = 0.08$, $p = 0.645$) and high-attractiveness target ($b = 0.00$, $SE = 0.09$, $p = 0.965$) (see SOM S.7.2 for full model).

In sum, across the two sets of analyses of geographic differences, we found hardly any evidence that the regional intensity of the spread of COVID-19 during the intense and early stages of the pandemic affected subjects' preferences for attractiveness and health (nor any other surveyed traits) in Danish prime ministers. However, it may be the case that any effects of disease threats like COVID-19 on leader preferences do not appear immediately. Perhaps instead it takes some time after an outbreak before followers and citizens modify their trait preferences in leaders? Moreover, White et al. (2013) originally tested their theory on leader preferences both inside and outside the domain of politics, while Tests 1–3 focused on political leadership. Thus, it could be that effects of COVID-19 on leader preferences are stronger for non-political leaders. For these reasons, we fielded another survey one year into the COVID-19 pandemic.

## Test 4: Disease threats and desired leadership characteristics one year into the COVID-19 pandemic

### Method

*Design:* Test 4 replicated the trait rating experiment from our first survey (Test 1) with three essential adjustments. First, we substituted colleague with "Boss" at the subject's job as the contrast category to prime minister for rated importance of attractiveness, health, competence, trustworthiness, dominance and power. That is, subjects were randomly assigned to rate trait importance for a boss or prime minister (target conditions) under either COVID-19 or a control condition (context conditions). Second, in order to test whether individual differences in tendencies to fear diseases predict leader preferences, we included the 15-item measure of Perceived Vulnerability to Disease (PVD) in our survey. Specifically, PVD captures "chronic concerns about the transmission of infectious diseases" (Duncan et al., 2009: p. 541) ($M = 0.46$, $SD = 0.16$, $\alpha = 0.82$). The PVD scale consists of Perceived Infectability ($M = 0.34$, $SD = 0.22$, $\alpha = 0.90$) and Germ Aversion ($M = 0.57$, $SD = 0.19$, $\alpha = 0.73$), respectively.[5] While Perceived Infectability taps "beliefs about immunological functioning and personal susceptibility to infectious diseases," Germ Aversion measures "aversive affective responses to situations that connote a relatively high likelihood of pathogen transmission" (Duncan et al., 2009: p. 542). In our analyses, we first employed the general PVD variable as predictor of trait importance and then the two subscales. Relatedly, subjects also answered five questions tapping whether the subject herself or her close friends or relatives faced severe risks if infected by coronavirus. Subjects could answer yes or no to each question. Aggregating across the five answers, we formed a 0–5 scale for self-reported COVID-19 risks ($M = 0.71$, $SD = 0.93$). Third, the second survey included manipulation checks asking subjects to report their perceived risks from COVID-19 in experimental situations (COVID-19 vs. control conditions). Perceived risk was captured on a five-point scale with 1 and 5 reflecting "Very little risk" and "Very severe risk," respectively ($M = 2.46$, $SD = 0.93$).

We fielded the second survey from March 26, 2021 and finished data collection on March 31, 2021 (during the second major lockdown of Denmark). We focus on the pre-registered tests for attractiveness and health but also report results for remaining traits.

*Participants.* 1515 Danes representative with respect to age (>18 years old; $M = 50.9$, $SD = 18.3$), gender (51.8 percent females) and geographical region were collected and financially compensated through Epinion. Similar to the first survey, sample size was determined based on budgetary constraints. Yet, with the dependent variables measured on a 7-point scale and assuming a 1.8 standard deviation, we can capture the following effects with 90% power in a $2 \times 2$ design: effects as small as 0.3 scale points for main effects, a difference of 0.5 comparing two individual groups, and interaction effects of about 0.7 (see SOM S.8 for sample descriptives).[6]

*Procedure.* As already stated, Test 4 followed the same procedure described for Test 1, with subjects randomly assigned to rate trait importance in a boss or a Danish prime minister (target conditions) under either a COVID-19 or a control condition (context conditions). At the time of the data collection, Danes had lived with coronavirus for more than a year. To minimize risks of subjects in the control condition thinking about coronavirus, the instructions in the control condition read: "We ask you to imagine that society has returned to its normal conditions, as before the coronavirus pandemic hit Denmark." Instructions in the COVID-19 condition reused the text from Test 1. Responses for

---

[5] Because the use of Chronbach's alpha as a measure of scale reliability has been questioned (Open Sciene Collaboration, 2000; Osburn, 1997), we also calculate McDonald's Omega, which shows similar levels of satisfactory reliability: $\Omega_{PVD} = 0.81$; $\Omega_{Perc.\ Infect.} = 0.90$; $\Omega_{Germ\ Avers.} = 0.74$.

[6] Replication data and command files for Test 4, 5 and 6 (contained in Survey 2) is publicly available via the Open Science Foundation: https://osf.io/fzh8v/.
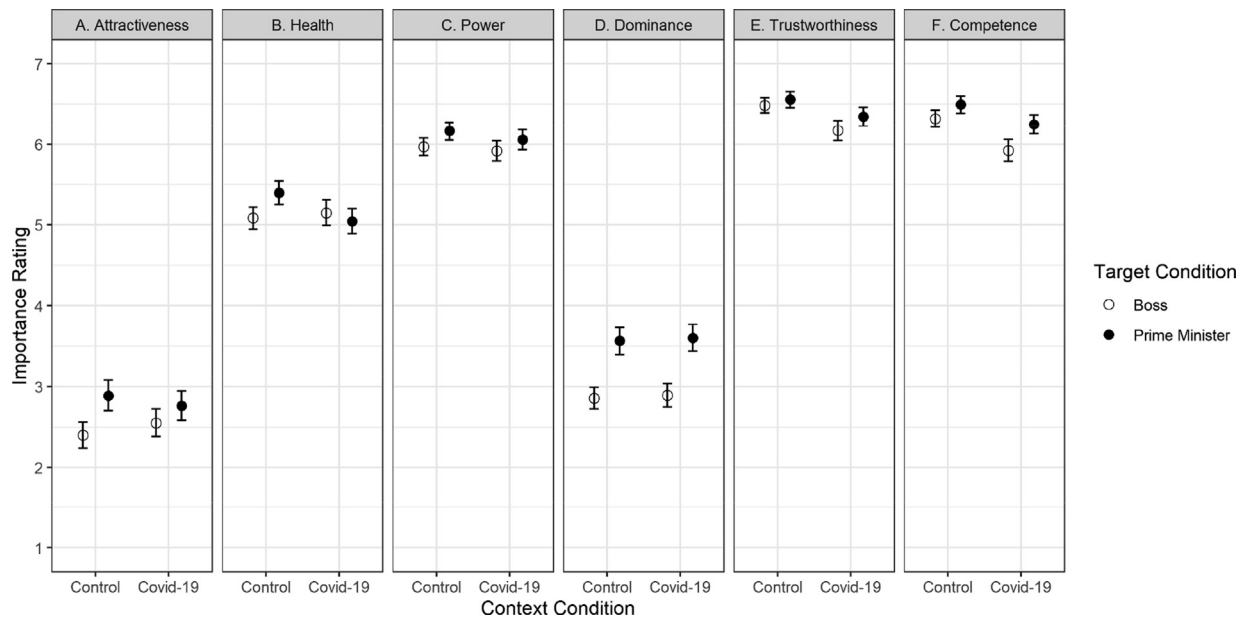
**Fig. 3.** Average rated importance of six traits in prime minister and boss (Test 4) with 95 percent confidence intervals.

rated trait importance were collected on 1 to 7 scales reflecting "Not at all important" and "Very important," respectively. Across conditions, rated importance of health (M = 5.16, SD = 1.48) was approximately twice as large as for attractiveness (M = 2.64, SD = 1.74), with absolute levels close to the results from Test 1.[7]

*Results and discussion*

Manipulation checks revealed that subjects assigned to the COVID-19 condition perceived a larger risk from coronavirus than subjects in the control condition ($M_{COVID-19}$ = 2.63, $M_{control}$ = 2.29, $t_{diff}$ = 7.13, p < 0.001).[8] Thus, the manipulation worked as intended. Next, we tested whether the COVID-19 condition upregulates preferences for health and attractiveness in a Danish prime minister and boss. For health, we found a negative and significant effect for prime minister (b = −0.35, SE = 0.11, p = 0.001) corresponding to a substantially small effect size with Cohen's *d* = −0.24. In contrast the effect remained non-significant for boss (b = 0.07, SE = 0.011, p = 0.532). For attractiveness, both analyses yielded non-significant differences (prime minister: b = −0.13, SE = 0.13, p = 0.347; boss: b = 0.15, SE = 0.12, p = 0.197). In sum, these results do not support the prediction that the COVID-19 pandemic enhanced preferences for healthy and attractive leaders. If anything, subjects exhibited reduced preferences for a healthy prime minister in the COVID-19 condition (Test 1 found an non-significant difference in the same direction). Although not pre-registered, we also tested whether the COVID-19 condition changed rated importance across the four remaining traits. Interestingly, results reveal that the COVID-19 condition lowered rated importance of competence and trustworthiness in both a prime minister ($b_{competence}$ = −0.25, SE = 0.08, p = 0.003; $b_{trustw.}$ = −0.21, SE = 0.08, p = 0.008) and a boss ($b_{competence}$ = −0.40, SE = 0.09, p < 0.001; $b_{trustw.}$ = −0.32, SE = 0.08, p < 0.001). For dominance and power, we found non-

significant differences for both prime minister ($b_{dominance}$ = 0.04, SE = 0.12, p = 0.747; $b_{power}$ = −0.11, SE = 0.09, p = 0.209) and boss ($b_{dominance}$ = 0.04, SE = 0.10, p = 0.725; $b_{power}$ = −0.05, SE = 0.08, p = 0.557). Average trait ratings across all conditions are displayed in Fig. 3.

As illustrated in Fig. 3, the COVID-19 condition affected trait ratings for prime minister and boss in similar ways except for health ratings. Thus, it is hardly surprising that interactions between context conditions and target conditions remained non-significant for attractiveness, competence, trustworthiness, dominance and power ratings (ps > 0.05), while for health a stronger negative effect of COVID-19 was found for prime minister compared to boss (b = −0.42, SE = 0.15, p = 0.006) (see SOM S.10.1 and S.10.2 for full models). Finally, because both Test 1 and Test 4 obtained ratings of trait importance in a prime minister, we can report results from exploratory tests of whether Danes changed trait preferences in their political leader between March 2020 and March 2021. Interestingly, ratings for attractiveness remained unaffected (ps > 0.05), while rated importance of health, competence, trustworthiness, dominance and power were slightly reduced (ps < 0.05) (see SOM S.11 for illustration of average prime minister trait ratings across Tests 1 and 4).

Following these experimental tests, we tested whether subjects' self-reported risks related to coronavirus and Perceived Vulnerability to Disease (PVD) predicted rated importance of health and attractiveness in a prime minister and a boss. Controlling for subjects' gender, age and education, self-reported risks related to coronavirus positively and significantly predicted rated importance of health in a prime minister (b = 0.14, SE = 0.06, p = 0.023). To evaluate the substantial strength of this relationship we used our model to calculate the difference in reported importance of health corresponding to a difference of an interquartile range (IQR, the 25th and the 75th percentiles) in self-reported risks. We found such a difference to equal 0.14 scale points (on a 1–7 scale), which we conceive of as a small effect. Next, we found positive but non-significant relationships between self-reported risks and ratings of health in a boss (b = 0.09, SE = 0.06, p = 0.108) and for ratings of attractiveness for both prime minister (b = 0.13, SE = 0.07, p = 0.081) and boss (b = 0.09, SE = 0.06, p = 0.178). Likewise, self-reported risks remained non-significantly

---

[7] Similar to results reported for Test 1, ratings of health and attractiveness are positively correlated for both targets: $r_{Prime\ Minister}$ = 0.31, p < 0.001; $r_{Boss}$ = 0.26, p < 0.001. See SOM S.9 for correlations across all six trait ratings separately for the two targets.

[8] Further testifying to the success of the manipulation, 65.4 percent of the subjects in the COVID-19 condition compared to 27.0 percent in the control condition ($t_{diff.}$ = 16.2, p < 0.001) reported that they thought of the last year and the coronavirus pandemic when answering the trait rating questions.
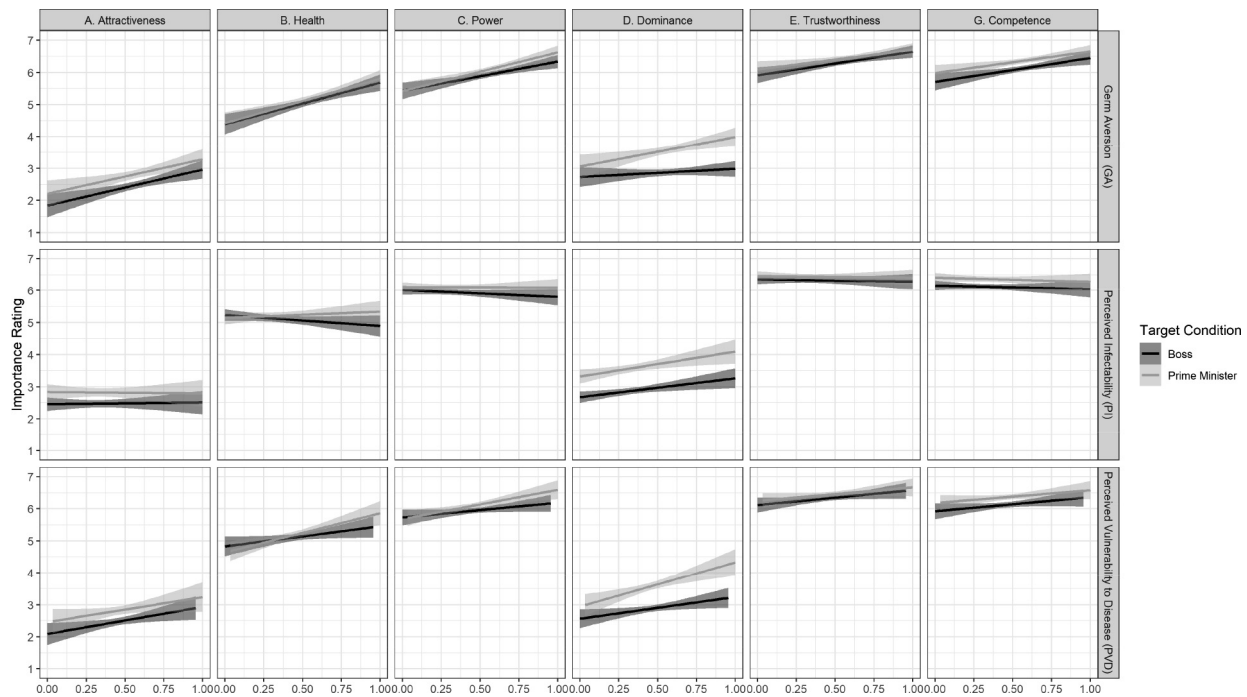
**Fig. 4.** Predicted relationships between Perceived Vulnerability to Disease and its sub-scales, Perceived Infectability and Germ Aversion, and six traits rated for prime minister and boss (Test 4) with 95 percent confidence intervals.

related to ratings of competence, trustworthiness, dominance and power (ps > 0.05) (see SOM S.12 for full models).[9]

Subsequently, we estimated relationships between PVD and trait importance (with the same control variables), finding a positive and significant relationship between PVD and rated importance of health in a prime minister (b = 1.06, SE = 0.35, p = 0.002) and a positive relationship that falls just short of conventional levels of statistical significance for prime minister attractiveness (b = 0.82, SE = 0.42, p = 0.053). To evaluate the substantial strength of the former relationship, we calculated the predicted difference in reported health importance for prime minister corresponding to a difference of one IQR in subject PVD. This yielded a difference of 0.24 scale points (on a 1–7 scale), which we see as a rather small effect. For trait ratings in a boss, PVD relates positively to both health (b = 0.47, SE = 0.32, p = 0.146) and attractiveness (b = 0.83, SE = 0.35, p = 0.019), although only the latter relationship is significant (for this relationship a difference of one IQR in PVD yields a substantially small effect of 0.19 scale points in rated attractiveness). For the remaining traits, PVD emerged as a significant and positive predictor of rated importance of dominance (b = 1.50, SE = 0.38, p < 0.001) and power (b = 0.65, SE = 0.27, p = 0.016; for competence and trustworthiness: ps > 0.05) in a prime minister. For a boss, PVD only significantly predicted ratings of dominance (b = 0.59, SE = 0.30, p = 0.047; for remaining traits: ps > 0.05) (see SOM S.13 for full models). The lower panel of Fig. 4 illustrates these relationships.[10]

These results support the theoretical prediction that PVD relates positively to rated importance of attractiveness and health in leaders —although not always significantly so. However, as illustrated in Fig. 4, PVD also predicts rated importance of dominance and power (only for prime minister). This suggests that subjects who are chronically concerned about the transmission of infectious disease might pre-

fer a wider set of traits in leaders than only attractiveness and health. Departing slightly from the pre-registration, we further provide exploratory tests of how each of the PVD subscales—Perceived Infectability and Germ Aversion—relate to rated trait importance. Although positively correlated (r = 0.26, p < 0.001), very different results emerged for the two subscales. Perceived Infectability only related positively to rated importance of dominance in a prime minister (b = 0.61, SE = 0.28, p = 0.027; for remaining traits: ps > 0.05). In stark contrast, Germ Aversion positively predicted all traits across prime minister ($b_{health}$ = 1.25, SE = 0.29, p < 0.001; $b_{attrac.}$ = 1.06, SE = 0.36, p = 0.003; $b_{competence}$ = 0.52, SE = 0.22, p = 0.018; $b_{trustw.}$ = 0.41, SE = 0.21, p = 0.057; $b_{dominant}$ = 1.28, SE = 0.32, p < 0.001; $b_{power}$ = 1.00, SE = 0.23, p < 0.001) and boss ($b_{health}$ = 1.11, SE = 0.28, p < 0.001; $b_{attrac.}$ = 1.10, SE = 0.31, p < 0.001; $b_{competence}$ = 0.64, SE = 0.22, p = 0.004; $b_{trustw.}$ = 0.62, SE = 0.21, p = 0.003; $b_{dominant}$ = 0.46, SE = 0.26, p = 0.084; $b_{power}$ = 0.78, SE = 0.22, p = 0.001), and almost all relationships are significant (see SOM S.14.1 and S.14.2 for full models). To evaluate the substantial strength of these relationships we again calculated the predicted difference in trait importance (based on the regression models) caused by a difference of one IQR in Germ Aversion. For ratings of health this yields differences of 0.34 and 0.30 scale points (1–7 scale) for prime minister and boss, respectively, which we see as moderately small. These results are illustrated in the middle (Perceived Infectability) and upper panels (Germ Aversion) in Fig. 4. Finally, to explore whether the relationships between PVD (or the two subscales) and rated trait importance are indeed linear, we also estimated similar models using ordered logit regression, as well as a locally estimated scatterplot smoother (LOESS). These alternatives give rise to the same substantial conclusions (see SOM S.15 and S.16).

Across analyses of individual differences related to perceived disease threats, both self-reported risks related to COVID-19 and PVD relate positively to rated importance of health and attractiveness but also dominance and power in leaders. However, exploring the relationships at the level of PVD's subscales, Germ Aversion emerges as the strongest predictor of rated trait importance in both a prime minister and a boss. Importantly, this holds not only for attractiveness and

---

[9] Furthermore, interactions between target condition and self-reported risks remained non-significant across all six traits (ps > 0.05).

[10] While reported tests controlled for gender, age and education, Fig. 4 reports bivariate relationships between PVD (and sub-scales) and trait ratings. Reassuringly, patterns remain similar for both set of tests.

health but also for competence, trustworthiness, dominance and power. Thus, rather than supporting the theoretical expectation that individuals who fear diseases (and, thus, COVID-19) should hold particularly strong preferences for healthy and attractive leaders, our results suggest that these individuals hold stronger preferences for a wide range of leadership traits. In other words, individuals who are prone to fearing diseases just seem to prefer leaders who score higher on any imaginable trait. We return to this point in the general discussion.

One might possibly argue that this trait rating experiment (like Test 1) employs a research design that too blatantly measures subjects' preferences for leader attractiveness. Test 5 therefore tests the prediction that disease threat heightens preferences for attractive leaders using a research design that more subtly investigates subjects' preferences for visual attractiveness in non-political leaders when primed to think of the COVID-19 pandemic.

### Test 5: Disease threats and preferences for attractive leaders one year into the COVID-19 pandemic

#### Method

*Design:* To test whether COVID-19 threat upregulates subtle preferences for attractiveness in leaders, we conducted another experiment, this time closely building on Study 4 reported in White et al. (2013). In our experiment, subjects were asked to imagine themselves working in a private company that is about to hire both a new immediate boss and a new coworker relative to the respondent. Physical contact with the coworker and the boss was described in similar ways, approximating one weekly meeting. Subjects were assigned to either a control or a disease threat condition, with the latter asking them to think of the most severe time during the COVID-19 pandemic (context conditions). Next, all subjects were shown—in random order—six different male individuals and asked whether each individual should be hired as a coworker or boss. Importantly, the six male individuals varied on perceived attractiveness, with three high- and low-attractiveness targets, respectively. All six targets were either current or former Norwegian members of parliament and thus resembled Danish male individuals, maximizing Test 5's realism. Specifically, the high- and low-attractiveness photos were chosen based on a rating study with unmorphed versions of all 20 candidate photos employed in Laustsen and Petersen (2020) (the rating study measured attractiveness on a 1–7 point scale with the following result for the materials used here: $M_{high\ attrac.} = 4.57$, $M_{low\ attrac.} = 2.65$, *t*-test for diff.: t = 23.22, p < 0.001). Because all respondents evaluated all six male individuals, the attractiveness conditions (high vs. low) constitute a within subject experimental factor, whereas the context conditions constitute a between subjects factor.

*Participants.* Subjects were the same 1515 Danes who participated in Test 4. Conditions across experiments for Tests 4 and 5 were assigned at random to make sure that any sequence effects canceled out. However, subjects always participated in the Test 5 experiment after participating in the Test 4 experiment (unrelated filler questions appeared between the two experiments).

*Procedure.* Based on first impressions from photos of the six applicants, subjects indicated whether a given applicant should be hired as coworker or boss, using a 6-point scale with endpoints 1 and 6 reflecting "Definitely coworker" and "Definitely boss," respectively (M = 3.43, SD = 1.25) (Study 4 reported by White et al. (2013) used an identical response scale). Finally, after the hiring decision questions, we included a manipulation check identical to the question described for Test 4, asking subjects to indicate perceived risk from coronavirus in the situation they thought of while answering the hiring questions (1–5 scale, with 5 reflecting "Very severe risk": M = 2.00, SD = 0.91).

The design for this experiment builds on a difference-in-differences logic by testing whether subjects differ in their hiring decisions for high- and low-attractiveness targets and whether this difference varies between control and COVID-19 conditions. If the theoretical expectation holds, we should expect subjects to preferentially hire high-attractiveness targets as their boss in the COVID-19 condition. In the context of this experiment, this entails a significant and positive interaction between context condition and attractiveness condition. Because each subject made six hiring decisions (recall that target attractiveness varied within subject), we cluster standard errors at the subject level in the reported analyses.

#### Results and discussion

Manipulation checks revealed that subjects assigned to the COVID-19 condition perceived a greater risk from coronavirus than subjects in the control condition ($M_{COVID-19} = 2.10$, $M_{control} = 1.90$, $t_{diff} = 4.35$, p < 0.001). Thus, the manipulation worked as intended, although not as strongly as for Test 4. Next, we tested the key prediction that subjects will hire high attractiveness targets as bosses with greater frequency in the COVID-19 compared to the control condition. However, the two-way interaction between attractiveness condition and context condition remained non-significant (b = 0.08, SE = 0.05, p = 0.133). Moreover, target attractiveness remained non-significant related to subjects' hiring decisions in the control (b = −0.03, SE = 0.04, p = 0.473) as well as the COVID-19 condition (b = 0.05, SE = 0.04, p = 0.150) (see SOM S.17 for full models). Fig. 5 below illustrates these non-significant results.

Next, we tested whether self-reported risks and PVD positively predicted preferences for attractive leaders. In the context of the hiring experiment, this entails a two-way interaction between either self-reported risk or PVD and the attractiveness condition. Because the two individual difference variables were observed rather than randomly assigned, we controlled for subjects' gender, age and education and included the context condition in the models. For self-reported risks we found a significant but negative interaction with assigned attractiveness condition (b = −0.08, SE = 0.03, p = 0.007), which means that subjects reporting higher risks from COVID-19 were less,
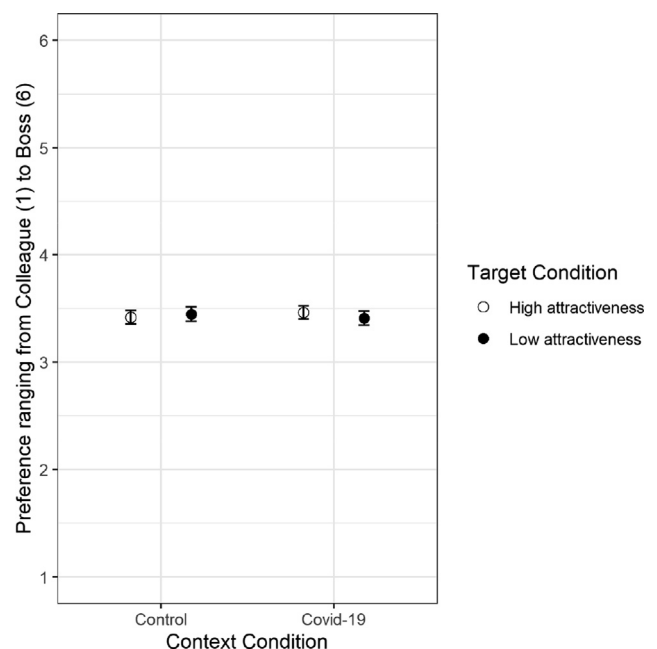
**Fig. 5.** Average hiring decisions across low- and high-attractiveness targets and context conditions in Test 5 with 95 percent confidence intervals.

rather than more, likely to hire the attractive targets as their boss. For PVD we found an non-significant interaction with target attractiveness (b = 0.19, SE = 0.18, p = 0.302). Furthermore, we conducted similar exploratory tests using each of PVD's sub-scales and found non-significant interactions for both Perceived Infectability (see SOM S.18 for full model).

In total, the results from this hiring experiment (Test 5) do not support the theory that COVID-19 enhances preferences for visually attractive leaders. Our results do not support either that subjects assigned to the experimental COVID-19 context or that subjects high in self-reported risks from COVID-19 or PVD preferentially hire attractive individuals as bosses rather than as coworkers. In contrast, the only significant finding from Test 5 shows that individuals who self-reported facing larger risks from COVID-19 were less likely to hire attractive targets as bosses.

### Test 6: Geographical threat intensity and leader preferences one year into the COVID-19 pandemic

Our final test of the overall prediction—that COVID-19 threat upregulates preferences for attractive and healthy leaders—links real-world variation across Danish municipalities in disease threat from COVID-19 to the two experiments reported as Tests 4 and 5. Thus, Test 6 mirrors Test 3 from our first survey. However, when we fielded the second survey, Danish health authorities had begun offering data on local variation in COVID-19 contamination risks at the municipality rather than the regional level. Consequently, Test 6 employs this more fine-grained geographical variable for COVID-19 severity.

*Method*

*Design:* To test how geographical variation in COVID-19 severity potentially relates to preferences for attractive and healthy leaders, we collected data from Statistics Denmark on the total number of confirmed COVID-19 cases per 100,000 people for each of the 98 Danish municipalities as of March 26, 2021 (M = 3959, SD = 1705). Because Epinion provided municipal belonging for each of our subjects, we were able to append the variable for municipal COVID-19 severity to our survey responses.

*Participants.* Subjects were the same 1515 Danes who participated in Tests 4 and 5.

*Procedure.* In the context of the trait rating experiment (Test 4), our preregistered hypotheses entail testing whether subjects rate attractiveness and health as more important in a prime minister and a boss the stronger the local threat from COVID-19. Likewise, in the context of the hiring experiment (Test 5), we predicted subjects' tendencies to preferentially hire attractive targets as their boss would grow with increasing levels of local COVID-19 threat. When estimating municipal differences we also controlled for subjects' gender, age and education as well as the relevant experimental conditions.

*Results and discussion*

First, we linked local COVID-19 severity to the trait rating experiment (Test 4). Local COVID-19 severity remained unrelated to both health (b = −0.00, SE = 0.00, p = 0.237) and attractiveness (b = −0.00, SE = 0.00, p = 0.124) in a prime minister. Similarly, non-significant relationships were found for boss with respect to health (b = −0.00, SE = 0.00, p = 0.585) and attractiveness (b = −0.00, SE = 0.00, p = 0.522). Finally, exploratory analyses for all remaining traits also yielded non-significant relationships between local COVID-19 severity and rated trait importance for both targets (ps > 0.05) (see SOM S.19 for full model). Second, we tested whether local COVID-19 severity related positively to preferentially

hiring the attractive targets as boss rather than colleague in the hiring experiment (Test 5). Specifically, we tested whether local COVID-19 severity interacts with target attractiveness. This is not the case (b = −0.00, SE = 0.00, p = 0.813). Moreover, local COVID-19 severity remained non-significantly (and negatively) related to hiring both high- and low-attractiveness targets as boss (ps > 0.05) (see SOM S.20 for full model). In sum, we found no signs across the data collected in our second survey (Tests 4 and 5) that local disease threats stemming from COVID-19 relate to trait preferences in leaders or to preferences for attractive leaders.

### General discussion

*Results summary and practical implications*

In this article we report the results from six well-powered and preregistered tests of whether the COVID-19 pandemic affected trait preferences in leaders. Building on the theory and results reported by White et al. (2013) that disease threat upregulates preferences for attractive and healthy leaders, we predicted that the threat from COVID-19 would also enhance preferences for attractiveness and health among our subjects. Specifically, we tested whether i) subjects experimentally assigned to COVID-19 conditions (compared to control conditions), ii) subjects with strong predispositions against disease threat and risks (as measured by self-reported risks related to COVID-19 or by Perceived Vulnerability to Disease), or iii) subjects living in geographic regions most severely affected by the COVID-19 pandemic upregulated their preferences for attractive and healthy political and non-political leaders. In addition, we gathered data at two very different time points during the COVID-19 pandemic: at the early stages of the first lockdown of Danish society and one year into the pandemic. Moreover, we tested subjects' explicit trait preferences in leaders as well as implicit preferences for visual attractiveness in leaders. Yet across all tests and analyses we find very little evidence that the COVID-19 pandemic enhanced follower preferences for attractiveness and health in leaders. In contrast, the clearest results actually speak against the theoretical prediction in interesting ways. First, using colleague as contrast category to prime minister, Test 1 revealed a significant rise in importance of health in a colleague for subjects assigned to the COVID-19 rather than the control condition. This suggests that disease threat might primarily regulate trait preferences in more proximate social relations—such as a colleague who potentially constitutes a real risk of contamination—rather than in leaders with whom followers more rarely interact. Second, Test 4 revealed that individual differences in Germ Aversion—tapping aversive affective responses to situations that connote a relatively high likelihood of pathogen transmission—relate positively to all six traits included in the study. In other words, individuals who generally fear disease and contamination report enhanced preferences for a wide range of leadership traits in political and non-political leaders alike rather than for attractiveness and health in particular. The most straightforward interpretation of this finding seems to be that followers who are chronically fearful of disease and illness seek out more archetypal leaders across the board.

*Theoretical implications*

Theoretically, our results hold important implications in several ways. Fist, if disease threats based on a global pandemic like COVID-19 do not activate evolved followership psychology leading to enhanced preferences for attractiveness and health, one might conclude that this falsifies the theory presented by White et al. (2013). However, before drawing this conclusion, one should bear in mind that White et al. (2013)'s original studies manipulated disease threat through verbal disease threat stories and visual cues in photos and also

used macro-level indicators of general pathogen and disease threat as opposed to our stimuli and indicators all focusing on COVID-19. Consequently, future research should seek to further test the theory, seeking to clarify whether it works more strongly for certain modes and primes of disease threat than others. Such studies will also help address the broader question of why it is that followers and voters tend to like attractive politicians (Berggren et al., 2010; Rosar et al., 2008).

Second, recent studies find that disease threats enhance conservatism, nationalism, anti-immigrant sentiments and authoritarianism (e.g. Beall, Hofer, & Schaller, 2016; Hartman et al., 2020). Relatedly, given that preferences for having a strong and dominant leader are often politically connected to these attitudinal dimensions, one might generate the prediction that disease threats should also upregulate preferences for dominant and powerful leaders. Yet results from Tests 1 and 4 show that this is seemingly not the case. This suggest that disease threats like the COVID-19 pandemic do not uniformly cause authoritarian swings in behavioral and attitudinal dispositions. Rather, such effects vary across the more specific attitudinal dimension in question. An important question for future research is therefore which specific attitudes, often seen as parts of a larger authoritarian package, are affected by disease threats and which are not.

Third, the absent effects of disease threat contexts are interesting given the well-established effects of intergroup conflict on trait preferences (dominance) in leaders (e.g. Laustsen & Petersen, 2016; Little, 2014). The idea that intergroup conflict and disease threat might affect leader preferences stems from evolutionary psychological theories of leadership and followership. These theories argue that humans possess a set of adapted psychological mechanisms (sometimes referred to as "a followership psychology") regulating leader preferences as a function of the intensity of evolutionarily relevant problems for coordination among group members (e.g. Van Vugt, 2006; von Rueden & van Vugt, 2015; Laustsen, 2021). Importantly, both intergroup conflict and disease threat constitute two evolutionarily relevant problems that most likely have molded human psychology in multiple ways. However, based on the absent effects of the COVID-19 contexts employed in Tests 1–6 on a range of traits as well as on preferences for facial attractiveness in leaders, one might consider whether intergroup conflict and disease threat constitute equally important contexts for understanding followership psychology. Based on the absent effects reported in this article, it could be possible that intergroup conflict has exerted much greater evolutionary pressures on human followership psychology and that intergroup conflict is also a much stronger contextual factor in regulating leader preferences among individuals in modern large-scale societies. Yet disease threat might still also affect leader preferences on traits not covered in the tests reported here. Hence, further studies that explore alternative theoretical accounts linking disease threat to leader trait preferences are warranted before one can draw any firm conclusions about the relative importance of disease threat vis-à-vis other evolutionarily relevant contexts for followership and leadership psychology.

*Limitations and future research*

One obvious risk involved with testing how experimental primes of COVID-19 threat potentially affect leader preferences in the middle of an ongoing pandemic relates to pre-treatment effects: subjects are already affected by the COVID-19 pandemic before entering our experimental tests. However, three findings suggest that the omnipresent COVID-19 threat did not drive the lack of significance in our experimental tests. First, as already stressed, Test 1 found that subjects upregulate the importance of health in a colleague when primed to think of the COVID-19 pandemic. This is a meaningful response given that contact with ill colleagues could possibly contaminate the subject. Second, manipulation checks from Tests 4 and 5 show that experimental conditions based on COVID-19 did affect perceived disease risk as expected. Third, if pre-treatment were driving our results, we should

also observe that subjects—across tests and conditions—exhibit preferences for attractiveness and health in leaders that are close to scale maximums (i.e. signs of ceiling effects). Figs. 1–5 clearly show that this is not the case. Thus, it seems unlikely that the COVID-19 crisis caused the insignificance of our experimental tests.

One might also wonder whether faces varying in perceived health would have yielded different results compared to the faces varying in attractiveness employed in Tests 2 and 5. Interestingly, Spisak, Blaker, Lefevre, Moore, and Krebbers (2014) used faces directly manipulated in perceived health and found that followers generally prefer healthy-looking leaders across a range of contexts, suggesting that health, like attractiveness, comprises a so-called first-order trait that followers value regardless of contextual conditions. Nevertheless, future research might consider further testing whether disease threat contexts enhance the importance of face-based impressions of health in leaders.

Relatedly, one might also consider how well our experimental manipulations succeeded in evoking the intended perceptions of disease threat and—in the case of the tests involving visual stimuli—trait impressions. This concern relates to the concept of "immersion" and the idea that researchers should strive to "improve realism […] to increase the level of immersion experienced by participants" (Aguinis & Bradley, 2014: p. 361). Because our data was collected during the COVID-19 pandemic, problems of immersion (if they exist) most likely relate to subjects in the control conditions *also* thinking about disease threat (rather than participants in the COVID-19 condition *not* thinking of disease threat). Yet our research designs sought to deal with this issue in two ways. First, we collected data at two different points in time with different wordings of the vignettes manipulating perceived disease threat, which suggests that the null findings are not caused by specific materials or timing of data collection. Second, the manipulation checks embedded in Tests 4 and 5 further underline that the COVID-19 and control conditions did indeed induce the intended difference in perceived disease threat. With respect to the face materials employed in Tests 2 and 5, we relied on visual stimuli that were all previously rated to vary as intended on perceived attractiveness and, thus, we see no strong reasons to expect that participants did not perceive the target faces as intended. Despite these features of the employed research designs, we obviously cannot eliminate all concerns about participant immersion. One interesting possibility worth pursuing for future projects would be to investigate whether politicians elected in real elections during the COVID-19 pandemic were more attractive or healthy-looking compared to politicians elected by the same electorate in the election before the COVID-19 pandemic hit. Unfortunately, such tests lie outside the possibilities of our datasets.

Based on the above considerations and the new questions that arose from our six tests, we urge future research to continue the mapping of follower preferences for traits and characteristics in relation to disease threat, intergroup conflict and other evolutionarily relevant and recurrent contexts.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

## Author contributions

LL conceived the study idea. LL designed the studies and ALO provided critical feedback. LL conducted the analyses with critical inputs from ALO. Both authors contributed to the manuscript preparation.

## Ethical approval

Denmark does not have an institutional review board for research outside biomedical research. In accordance with the national scientific guidelines, formal ethical approval for the present research was not obtained (for Danish rules on ethical approval of various research projects see https://www.rm.dk/sundhed/faginfo/forskning/de-videnskabsetiske-komiteer/anmeldelse/hvilke-projekter-skal-anmeldes/). Furthermore, the data gathering was administered by the Danish polling company, Epinion, who recruited respondents from their standing panel of Danish citizens. Epinion also financially compensated the respondents according to their payment schemes. Finally, the survey included a careful briefing and a short debriefing. Given that no deception was involved in the experimental treatments this strategy seemed sufficient.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.leaqua.2021.101574.

## References

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods, 17*(4), 351–371.

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly, 28*(1), 5–21.

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science, 323*(5918).

Antonakis, J., & Eubanks, D. L. (2017). Looking Leadership in the Face. *Current Directions in Psychological Science, 26*(3), 270–275.

Beall, Alec T., Hofer, Marlise K., & Schaller, Mark (2016). Infections and Elections: Did an Ebola Outbreak Influence the 2014 U.S. Federal Elections (and if so, How)? *Psychological Science, 27*(5), 595–605. https://doi.org/10.1177/0956797616628861.

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.

Blom-Hansen, J., Elklit, J., Serritzlew, S., & Riis Villadsen, L. (2016). Ballot position and election results: Evidence from a natural experiment. *Electoral Studies, 44*, 172–183.

Bøggild, T., & Laustsen, L. (2016). An intra-group perspective on leader preferences: Different risks of exploitation shape preferences for leader facial dominance. *The Leadership Quarterly, 27*(6), 820–837.

Danish Statistics (2021). SMIT1: COVID-19 spread of infection per day (eksperimentel statistik) by key figures. Url (retrieved 28-08-2021): https://www.statistikbanken.dk/statbank5a/SelectVarVal/Define.asp?Maintable=SMIT1&PLanguage=1.

Duncan, L. A., Schaller, M., & Park, J. H. (2009). Perceived vulnerability to disease: Development and validation of a 15-item self-report instrument. *Personality and Individual Differences, 47*(6), 541–546.

Giacomin, M., & Rule, N. O. (2020). How static facial cues relate to real-worl leaders' success: A review and meta-analysis. *European Review of Social Psychology, 31*(1), 120–148.

Hartman, T. K., Stocks, T. V. A., McKay, R., Gibson-Miller, J., Levita, L., Martinez, A. P., ... Bentall, R. P. (2021). The authoritarian dynamic during the COVID-19 pandemic: Effects on nationalism and anti-immigrant sentiment. *Social Psychological and Personality Science, 12*(7), 1274–1285. https://doi.org/10.1177/1948550620978023.

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.

Laustsen, L. (2014). Decomposing the relationship between candidates' facial appearance and electoral success. *Political Behavior, 36*(4), 777–791.

Laustsen, L. (2021). Candidate evaluations through the lens of adaptive followership psychology: How and why voters prefer leaders based on character traits. *Political Psychology*. https://doi.org/10.1111/pops.12738.

Laustsen, L., & Petersen, M. B. (2015). Does a competent leader make a good friend? Conflict, ideology and the psychologies of friendship and followership. *Evolution and Human Behavior, 36*(4), 286–293.

Laustsen, L., & Petersen, M. B. (2016). Wining faces vary by ideology: How nonverbal source cues influence election and communication success in politics. *Political Communication, 33*(2), 188–211.

Laustsen, L., & Petersen, M. B. (2020). Why are right-wing voters attracted to dominant leaders? Assessing competing theories of psychological mechanisms. *The Leadership Quarterly, 31*(2), 101301. https://doi.org/10.1016/j.leaqua.2019.06.002.

Lindholt, M. F., & Petersen, M. B. (2021). Politisk polarisering af holdninger og adfærd blandt danske borgere under COVID-19-pandemien. Unpublished working paper. Online available from: https://github.com/mariefly/HOPE/raw/master/Politisk_polarisering_af_holdninger_og_adf%C3%A6rd_blandt_danske_borgere_under_COVID-19-pandemien_20210209.pdf.

Little, A. C. (2014). Facial appearance and leader choice in different contexts: Evidence for task contingent selection based on implicit and learned face-behaviour/face-ability associations. *The Leadership Quarterly, 25*(5), 865–874.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122–1135.

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., van 't Veer, A. E., & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences, 23*(10), 815–818.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. Science 349, aac4716. doi: 10.1126/science.aac4716.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5*(3), 343–355.

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*, 64–79.

Schaller, M. (2016). The behavioral immune system. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (2nd ed., pp. 206–224). New York, NY: Wiley.

Schaller, M., Murray, D. R., & Bangerter, A. (2015). Implications of the behavioural immune system for social behaviour and human health in the modern world. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*, 20140105.

Spisak, B. R., Blaker, N. M., Lefevre, C. E., Moore, F. R., & Krebbers, K. F. B. (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*, 1–9.

Todorov, A. (2017). *Face Value: The Irresistible Influence of First Impressions*. Princeton, United States: Princeton University Press.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annuela Review of Psychology, 66*(1), 519–545.

Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2012). Disgust: Evolved function and structure. *Psychological Bulletin, 120*(1), 65–84.

Van Vugt, M. (2006). Evolutionary origins of leadership and followership. *Personality and Social Psychology Review, 10*(4), 354–371.

Verhulst, B., Lodge, M., & Lavine, H. (2010). The attractiveness halo: Why some candidates are perceived more favorably than others. *Journal of Nonverbal Behavior, 34*(2), 111–117.

von Rueden, C., & van Vugt, M. (2015). Leadership in small-scale societies: Some implications for theory, research and practice. *Leadership Quarterly, 26*(6), 978–990.

White, A. E., Kenrick, D. T., & Neuberg, S. L. (2013). Beauty at the ballot box: Disease threats predict preferences for physically attractive leaders. *Psychological Science, 24*(12), 2429–2436.