Evolutionary leadership theory and economic voting: Warmth and competence impressions mediate the effect of economic perceptions on vote

Alexander Bor

Aarhus University, Department of Political Science

alexander.bor@ps.au.dk

In press: The Leadership Quarterly. DOI: 10.1016/j.leaqua.2019.05.002

Leaders' persona and the state of the economy are among the two most salient topics during election campaigns. Existing scholarship treats these as two independent or even competing factors. Economic perceptions are overlooked as cues for leader evaluations, while leader evaluations rarely enter considerations of the economic vote. This article builds on evolutionary leadership theory to bridge these distant literatures. It proposes that evaluating leaders' performance based on the resources available to group members may have improved followers' fitness ancestrally. Accordingly, it predicts that the effect of economic perceptions on vote choice is mediated by leaders' warmth and competence impressions in modern democracies. To test these predictions, the article first analyzes representative survey data from seventeen elections in three countries (USA, Australia and Denmark). Second, it relies on two original, well-powered manipulation-of-process experiments to test the validity of the causal claims.

Keywords: leader evaluation, economic voting, evolutionary psychology, warmth, competence

Introduction

Citizens pay particular attention to the personalities of political leaders, especially during election campaigns (Lobo & Curtice, 2015). Research in political science has established that voters form impressions on multiple trait dimensions (Funk, 1999; Kinder, Peters, Abelson, & Fiske, 1980) and that these impressions have a substantial and causal effect on general evaluations and vote choice (Laustsen & Bor, 2017; McGraw, 2011). Evolutionary psychologists have offered an appealing ultimate explanation for this phenomenon. Monitoring potential leaders and thereby promoting and sustaining the leadership of benevolent and able persons yielded fitness benefits for followers ancestrally (Van Vugt, Hogan, & Kaiser, 2008; Van Vugt & Kurzban, 2007).

Meanwhile, a large literature in political science and economics has established that one of the best determinants of vote choice is perceptions of the economy (Fair, 1978, 2009; Key, 1966; Kramer, 1971; Lewis-Beck & Stegmaier, 2000). This literature highlights the rewarding or punishing of incumbents based on the performance of the economy as the main mechanism of democratic accountability (Anderson, 2007). Indeed, "it is virtually a universal belief among politicians, political commentators, and even voters that elections are referenda on the economy" (Duch & Stevenson, 2008, p. 1).

Interestingly, most researchers treat leader traits and economic perceptions as independent or even competing determinants of vote choice (but see Bartels, 2002; Jacquart & Antonakis, 2015). Students of leader evaluations have explored a large array of cues affecting trait impressions but paid little attention to the economy. Thereby, they ignore a potentially important determinant of leader evaluations.[1] Meanwhile, scholars of economic voting scarcely included leader ratings in their

[1] That said, evolutionary psychology has found that the number of resources available in the environment (an important part of contemporary economies) has large and long-lasting consequences for human development and behavior (Del Giudice, Gangestad, & Kaplan, 2016), which may affect leader preferences (Rule & Tskhay, 2014; Safra et al., 2017).

models, even though their theories at least implicitly hinge on citizens' ability and propensity to link leaders' actions to the state of the economy. Studying the psychological mechanisms fueling economic voting may offer new insights about when and why citizens are more likely to keep leaders accountable.

This article seeks to bridge the economic voting and leader evaluations. It aims to open the black box of causality by exploring the mechanisms of economic voting. Taking an explicitly evolutionary approach to leader evaluations, it proposes that the resources (e.g. food, shelter) available to followers may have been a valid cue of leader performance ancestrally. The article argues that followers could improve their fitness (i.e. reproductive success) by relying on cues of their economic well-being to infer their leaders' character traits. Accordingly, it predicts that economic perceptions affect both competence and warmth impressions, which in turn affect support for leaders and vote choice. Whereas previous works have theorized about – but have not tested – the competence mechanism, the warmth mechanism is a novel prediction of this article. The predictions are tested on a large dataset of seventeen representative surveys from three countries (USA, Australia and Denmark) and data from two original well-powered survey experiments.

The article highlights the benefits of applying an evolutionarily informed approach to "suggest important domains of inquiry" and to "predict previously unobserved phenomena" (Buss, 1991, p. 477). It shows to students of leader evaluation that informational cues related to economic perceptions are important determinants of leader evaluations. It demonstrates to scholars interested in economic voting that incumbent trait impressions partially mediate the economic vote and, more broadly, that our evolved followership psychology is sensitive to economic cues. The latter finding may help to explain the cross-cultural consistency and the large effect of economic voting. More generally, the article demonstrates the crucial role that our evolved followership psychology plays in making us good democratic citizens who keep political leaders accountable.

Leader trait impressions from an evolutionary perspective

This article proposes that leader trait impressions are key to understanding why people vote with the economy. It is therefore fitting to begin the discussion by asking why followers would pay attention to their leaders' traits. Taking an evolutionary perspective provides a forceful ultimate explanation to this question (Van Vugt & Ahuja, 2010; Van Vugt & Kurzban, 2007). Evolutionary leadership theory proposes that leadership may have been an important social mechanism enabling our ancestors to reap the benefits of extensive collective action. Cooperation between multiple unrelated individuals is notoriously difficult because free-riding may prevent cooperators from enjoying the benefits of their efforts (Boyd & Richerson, 2005). Monitoring the activities of group members, punishing free-riders if necessary and organizing group activities in general can increase cooperation, but performing these activities can become a (second-order) collective action problem in itself (Oliver, 1980; Ostrom & Walker, 2005). Leaders perform these cooperation-facilitating activities and thereby provide benefits to the entire group (Van Vugt & De Cremer, 1999).

To be able to achieve this, leaders are "accorded differential influence within a group over the establishment of goals, logistics of coordination, monitoring of effort, and reward and punishment" (von Rueden, Gurven, Kaplan, & Stieglitz, 2014, p. 539). Leaders emerge spontaneously in situations requiring collective action, and they improve group performance in public goods games (Glowacki & von Rueden, 2015; Stein, 1975). Anthropological studies further find that leader attributes (physical formidability, kin support and trustworthiness) predict group performance in a hunter-gatherer society (von Rueden et al., 2014). In short, diverse literatures found evidence that leaders produce fitness gains (i.e. increases in reproductive success) to followers through facilitating more and better cooperation.

However, the logic of evolution also posits that leadership and followership could not have evolved unless they produced fitness gains for leaders as well. To give leaders the necessary incentive to lead, followers grant them higher status, which translates into higher reproductive success (von Rueden,

Gurven, & Kaplan, 2011). In essence, leadership and followership are themselves a form of cooperation. To perform the tasks of leadership (cost), leaders receive the reward of higher status and possibly more resources (benefit). To grant the leader higher status and more resources (cost), followers receive the fruits of extensive cooperation, such as better access to food, shelter or protection (benefit) (Price & Van Vugt, 2014). Importantly, both parties are interested in reducing their costs and increasing their benefits. Followers are motivated to grant as few privileges to leaders as possible while enjoying as many of the products of cooperation as possible. Naturally, leaders are interested in undermining both of these activities. On the cost side, leaders might demand (higher) contributions from a follower. On the benefit side, they affect the total number of resources available to the group and may influence how these resources are distributed. Consequently, it is reasonable to assume that a followership psychology includes modules designed to monitor leaders' propensity to inflict costs and to acquire and grant benefits (Bor, 2018).

One set of these modules should be monitoring leaders' tendencies to favor or impede the followers' well-being. A follower should keep track of whether a leader helps or harms them. If a leader is mostly harmful, the follower is strongly motivated to change their relationship with the leader. This could occur either because the leader is a dominant, exploitative individual or because he or she is primarily concerned about the welfare of others in the group. If, on the contrary, the leader is mostly helping them, a follower should do their best to preserve the leader and their relationship with them. This article refers to character traits tapping into followers' impressions of leaders' helpful or harmful intention as *warmth*.[2]

Another set of leader-monitoring modules should be focusing on leaders' abilities to perform the tasks of leadership. These abilities are relevant for followers because they may influence the number of resources and public goods available to group members. All else equal, a better leader entails

[2] Other works refer to essentially the same concept as morality (van Leeuwen, Park, & Penton-Voak, 2012) or character (Bittner, 2011).

more food, safety, peace in the group and success in intergroup conflicts. This article refers to character traits tapping into followers' impressions of leaders' abilities as *competence*.

A long line of empirical evidence shows that these two dimensions are present in political leader evaluations (Abelson, Kinder, Peters, & Fiske, 1982; Bittner, 2011; Funk, 1996; Kinder & Sears, 1985; Laustsen & Bor, 2017; McCurley & Mondak, 1995; Wojciszke & Klusek, 1996). However, monitoring warmth and competence is not unique to followership psychology. Barclay (2013) argues that warmth and competence (labeled as tendencies and qualities) are important to assess the value of potential cooperation partners in biological markets in general. Fiske et al. (2007; 2002) propose that warmth and competence are universal dimensions of social cognition. Importantly, the distinction may hold water even outside human psychology: Chimpanzees show understanding of the difference between willingness (i.e. warmth) and ability (i.e. competence) to perform an action (Call, Hare, Carpenter, & Tomasello, 2004; Melis, Hare, & Tomasello, 2006).

Given the fitness benefits of accurately assessing the intensions and competence of a leader, it is reasonable to assume that followers rely on a wide variety of informational cues to form these impressions. Empirical evidence from psychology and political science confirms this prediction. Trait impressions are influenced by physical attributes of leaders, such as their formidability (Murray, 2014; Stulp, Buunk, Verhulst, & Pollet, 2013), race (Livingston & Pearce, 2009; Moskowitz & Stroh, 1994), gender (Cassese & Holman, 2017; Huddy & Terkildsen, 1993; Winter, 2010) and facial appearance (Laustsen & Petersen, 2016; Todorov, Mandisodza, Goren, & Hall, 2005; Van Vugt & Grabo, 2015). They are also updated based on the views that politicians express, such as leftist views yielding better warmth ratings, whereas rightist views increase perceived competence (Bittner, 2011; Laustsen, 2017; Rapoport, Metcalf, & Hartman, 1989). Last, but not least, a leader's actions and policy performance also influence their perceived traits. Here, again, liberal leaders' focus on social welfare improves their perceived warmth, whereas conservative leaders' attention to taxation, business and a firm foreign policy increases their perceived competence (Bittner, 2011; Hayes, 2005).

Scandals primarily erode warmth ratings, particularly if they concern the political as opposed to the private life of a candidate (Funk, 1996; Maier, 2011). Relatedly, leaders' violations of procedural fairness norms also hurt their perceived warmth (Bøggild & Petersen, 2016).

This article focuses on a cue for trait impressions that has so far received very little attention from researchers concerned with leader evaluations: perceptions of the economy. This omission is surprising given that a voluminous literature in political science demonstrates that perceptions of the economy is one of the largest, most consistent determinants of vote. This implies that economic perceptions may influence leader evaluations, too. Indeed, some theories of economic voting explicitly posit that economic perceptions are taken as indicators of incumbents' economic policy performance and are therefore used as a cue for incumbent competence (Duch & Stevenson, 2008). However, students of economic voting rarely look into leader evaluations. Below, I first review the economic voting literature's main conclusions that are relevant for leader evaluations. Next – building on evolutionary leadership theory – I offer a novel, ultimate explanation for followers' attention to the state of the economy.

Linking economic voting and leader evaluations

Voters around the world turn to the economy[3] in forming their vote choice (Anderson, 2007; Lewis-Beck & Stegmaier, 2000; Wilkin, Haller, & Norpoth, 1997). Extensive research on the topic demonstrates that subjective evaluations of the state of the national economy are more predictive of vote than impressions of personal economic well-being; that retrospective (past-oriented) evaluations are more important than prospective (future-oriented) evaluations. These individual-level attitudes translate into a positive relationship between objective economic indicators and

[3] For the present purposes, it is best to think of economy as "the structure or conditions of the production, distribution and consumption of goods and services in an area" (Merriam-Webster, 2018). An important benefit of such a non-technical definition is that it travels well from modern to ancestral times.

election results (Lewis-Beck & Paldam, 2000). Moreover, experimental evidence demonstrates that there is a causal relationship between economic perceptions and vote intentions (Simonovits, 2015).

On what basis do people evaluate the economy? This is a pressing issue because given the complexity of modern economies, it is a gargantuan task to (objectively) assess the state of the economy. Luckily, people appear comfortable forming *subjective* evaluations of the economy even without a degree in economics (Boyer & Petersen, 2018). First, they may follow the news and are thus frequently exposed to elite discussions about the state of the economy (Hetherington, 1996). More importantly, however, there is increasing evidence that people also rely on their personal experiences and on the economic conditions of people similar to them (Ansolabehere, Meredith, & Snowberg, 2014; Mutz & Mondak, 1997; Newman, Velez, Hartman, & Bankert, 2015). In a recent study, Bisgaard and colleagues (2016) found that citizens' economic perceptions are predicted well by the number of unemployed people residing in an 80-meter radius around a respondent's home.

Why do people vote with the economy? Most works in political science build on theories of democratic accountability, often assuming rational voters. Lewis-Beck and Stegmaier (2000, p. 212) suggest that "[v]oters, regardless of the democracy in which they live, assess national economic conditions and reward or punish the politicians responsible for those conditions". Similarly, Anderson (2007) argues that citizens who are motivated to keep leaders accountable but have limited willingness or capacity to process complex political information vote with the economy because it is a salient and easy-to-process issue. This argument thus presumes that voters' economic evaluations feed into leader evaluations but remain agnostic about the precise mechanisms.

Duch and Stevenson's (2008, p. 2) theory of economic voting offers the most detailed explanation of the role of leader evaluations: "[R]ational voters condition their vote on the incumbent's record of economic performance because this is the optimal way to identify and elect competent economic managers". Although the authors test and confirm several implications of this "competency model of economic voting", interestingly, their analysis does not explicitly test the role of competence

impressions, nor have other students of economic voting investigated the relationship between economic perceptions and leader impressions.

In a similar fashion, leadership scholars have long distinguished between perceptions based on organizational outcomes and perceptions based on observed traits and behavior – paralleling economic performance and trait impressions, respectively (Lord & Maher, 1991; see also Erickson & Krull, 1999). Jacquart and Antonakis (2015) build on this literature and propose an interaction between performance signals and trait impressions. Relying on data from US presidential elections and a realistic scenario experiment in a business context, they demonstrate that charisma impressions matter more when economic performance signals are ambiguous. Yet, these results cannot be directly employed to understand whether and how perceptions from objective outcomes inform leader trait impressions. This article takes the first steps to fill this gap in the literature.

An ultimate explanation of economic voting

This section seeks to offer a novel answer to the question of why people vote with the economy. I apply the logic of evolutionary thinking to ask what may have been the evolutionary function of conditioning support for a leader on the economic cues. A plausible theory about the positive fitness consequences of this behavior would offer an ultimate explanation of economic voting. It would explain how it helped our ancestors to survive and reproduce, why the behavior was thus selected for and why it may have become part of human psychology.

The previous sections established that followers must monitor leaders to make sure that they perform their duties competently and with good intentions. This means that followers should pay attention to correlates of these traits. Here, I argue that economic cues may have been among the best diagnostic tools to establish a leader's warmth and competence. This is easy to see if we consider that a large part of the social function of leaders was economic. By facilitating cooperation in the group, leaders contributed to the production of goods and may have personally influenced the

distribution and consumption of these goods. Consequently, the abundance or scarcity of goods was a function of leaders' qualities.

Consider the example of hunting. Anthropological records show that "small foraging groups are common ethnographically, [and] many foragers seasonally or periodically aggregate into larger groups" (Kelly, 2013, p. 174). In other words, foraging may have been a social, collective effort ancestrally to a significant extent. Good leaders may help to organize a hunting party from the best hunters in the group. They could ensure that everyone shows up and is willing to join an expedition. They may help to resolve disputes emerging between members of the hunting party. They may rely on their personal expertize to make crucial, quick decisions during the hunt. All else equal, having an experienced and reliable leader would result in more large game killed and thus more valuable food for the group. Although followers undoubtedly monitor several aspects of a leader's behavior, ultimately, they are most interested in the output – in this case, whether there is more or less meat available. If there is a lot of food from hunts, followers are better off inferring that they are led by a good leader and following him or her in the next hunt, too.

Hunting and the abundance or scarcity of meat is just one example where people could rely on the amount of resources available to infer the qualities of a leader. Leadership is also prominent in inter-group conflicts, where leaders could contribute to acquiring resources from enemies and protecting the group's own resources from them. Conflict resolution is another area where leaders have a tangible effect on the group as a unit of production. An intelligent and trustworthy leader may help to prevent or quickly resolve costly disputes that could undermine the production or protection of resources. Rewarding leaders if the group was doing well and punishing them if the group was doing poorly may thus have enhanced the fitness of followers ancestrally. A reasonable objection may be that the close proximity of ancestral leaders to followers offer sufficient information to form accurate trait impressions. Why would a follower pay attention to a noisy cue such as the availability of resources if they could simply watch the leader in action? However, forming accurate trait

impressions is a difficult task. Hill and Kintigh (2009) demonstrate that distinguishing between good and poor hunters – a comparatively simple, task-specific competence – may require years of quantitative data because of large temporal variation in individual hunting successes. Assessing the traits of leaders may have been even more difficult given that facilitating group coordination and cohesion requires continuous interactions with multiple individuals. Many, if not most of these interactions may go unseen by a given follower, yet may have serious consequences for her.

The leap from evaluating leaders based on the resources available to the group in a small-scale society to evaluating leaders based on perceptions of the national economy in large-scale societies is not as large as it may seem. As I argue above, people evaluate the national economy partly based on the well-being of neighbors and in-group members – informational cues that have been available since the Pleistocene. Of course, modern citizens are also exposed to novel cues such as macro-economic analyses. However, it is plausible that such novel cues may be reinterpreted to also feed into our evolved followership psychology (Petersen, 2016), especially with the aid of news media (Iyengar, 1994) and social networks (Aarøe & Petersen, 2018). To sum up, our intuition to assess political leaders based on the economy may be part of an adaptive psychological toolkit designed to monitor and select leaders whose actions – among other things – influenced followers' economic circumstances.

The psychological process of economic voting

This ultimate explanation helps to generate novel insights about the *proximate* psychological mechanisms involved in economic voting. In other words, whereas the previous section focused on why paying attention to economic cues (broadly defined) may have yielded fitness gains ancestrally, this section seeks to offer a plausible model for *how* the mind processes these cues to inform leader evaluations. I argue that trait impressions play an important mediating role. Humans in general "model others' behavior as a product of unobservable internal states" (Boyer & Barrett, 2016, p. 163; see also Dennett, 1987). Followers may be forming trait impressions in an attempt to model leaders'

behavior. Consequently, followers will update these trait impressions based on past behavior to maximize the accuracy of these models. Therefore, the question is which leader traits influence their performance to deliver benefits.

First, leaders' competence by definition affects their performance. Leaders who are more competent provide more public goods and resources to group members and, thus, increase followers' fitness. If followers could improve their fitness ancestrally by updating leader competence impressions based on their welfare (Bor, 2017), citizens today might still rely on economic perceptions to form competence impressions. Leader competence impressions are in turn an important predictor of support (McGraw, 2011). Importantly, this evolutionarily informed theory and classic economic voting literature (Duch & Stevenson, 2008) converge on this prediction, but to the best of my knowledge, no empirical attempts have been made to test its validity. This takes us to the first hypothesis:

H1. The relationship between perceptions of the economy and support for a leader is mediated by competence impressions.

Competence, however, is not the only trait influencing leader performance to deliver resources. *Warmth* also affects performance in at least two ways. First, leaders may be motivated to escape their duties as leaders. In the terminology used above, this means that leaders are better off reaping the benefits of leadership status without paying the costs of facilitating cooperation. Second, and perhaps more important, leaders' warmth affect their propensity to distribute costs and benefits fairly among followers. This is a crucial aspect of leaders' performance, especially in large and heterogeneous groups, where leadership is particularly important (von Rueden et al., 2014).

Although a competent leader may reduce the costs of individuals and increase the amount of resources to distribute, costs are never zero, and resources are always scarce. Consequently, leaders have ample opportunities to discriminate against specific followers. An individual's net benefit from a hunt may depend on whether a leader invites them to join the hunting party, on what role the leader

assigns them and on which parts of the spoils they receive. Similarly, a leader who intervenes less in conflicts involving an individual or who often favors the individual's opponents leaves the individual worse off. Even subtle, yet systematic acts of discrimination can accumulate to substantial fitness costs, which is reflected in followers' concern about procedural fairness of political decisions (Bøggild & Petersen, 2016).

In short, from a follower's perspective, a bad economy may indicate either a leader who lacks the intention to perform his or her duties or, worse, one who performs these duties but systematically undervalues the follower's welfare in allocating costs and benefits. Both of these cases are strong cues of a leader's low warmth. Consequently, it is reasonable to assume that when perceiving a bad economy, followers reliably downregulate impressions of leaders' warmth, which in turn has a large influence on support for leaders (Bittner, 2011; Laustsen & Bor, 2017).

The insight that political actors may differentially distribute access to resources is not alien to classic political science research on economic voting (Anderson, 2007). Citizens lacking resources may imply that leaders are benefitting others and thereby link their economic perceptions to warmth impressions (Hibbs, 1977). Nonetheless, to the best of my knowledge, the prediction that economic perceptions may feed into warmth impressions has not been formulated before. This is therefore a novel hypothesis:

H2. The relationship between perceptions of the economy and support for a leader is mediated by warmth impressions.

Overview of studies

The two hypotheses in this article concern causal mediation effects in political leader evaluation. Testing these hypotheses is a challenging task for both conceptual and statistical reasons. First, although the hypotheses focus on a universal followership psychology, leaders' evaluations are also influenced by the peculiarities of particular elections and national contexts. Testing the hypotheses in

several different cultural, institutional, economic and electoral contexts therefore increases the validity of this study. Second, many citizens have strong impressions about their leaders. Manipulating these impressions may be difficult and unethical. Consequently, conducting experimental studies with real politicians is problematic. The researcher is often left facing a difficult tradeoff between analyzing observational data with high external validity but potentially biased causal estimates and conducting experimental research with high internal validity but potentially low generalizability to real-world phenomena.

This article circumvents this dilemma by relying on both observational and experimental data. Study 1 looks at seventeen elections in three countries (USA, Australia and Denmark). These countries provide considerable variation on a number of important political and economic factors known to influence economic voting and thereby ensure that the findings have high generalizability, at least in the Western, developed world. However, Study 1 reports purely correlational evidence, and thus, its estimates may be biased (Antonakis, Bendahan, Jacquart, & Lalive, 2010). Studies 2 and 3 therefore offer original experiments designed specifically to test the causal claims. The experiments were fielded on a well-powered and relatively diverse sample of Americans through Amazon's Mechanical Turk. The two studies combined provide a rigorous test of the article's predictions.

Study 1: Observational analysis

Method

Study 1 relies on representative survey data from the USA, Australia and Denmark. These three countries are selected because they have crucial institutional and economic differences, which may affect the extent to which warmth and competence impressions are a function of economic perceptions. In particular, they vary considerably regarding the chief executive's responsibility for economic policy and the size and openness of the national economy. The US is a prototypical high clarity of responsibility country with a strong president. Moreover, it has the largest and, thus, one of the least vulnerable economies in the world. US citizens should have an easy time making the link

between the economic situation and their president's warmth and competence. Denmark, conversely, has low clarity of responsibility and a relatively open economy. Denmark has a multiparty system where coalitional governments are the norm. Moreover, as a typical Scandinavian country, Denmark has powerful unions and business associations that are involved in policy making, especially on issues related to the labor market. Although Denmark is a rich country, it is also a small one. The fate of the Danish economy is intertwined with trends in other large economies in the European Union and the world. Danish citizens might therefore have a hard time linking their economic perceptions to the incumbent prime minister's warmth and competence.[4] Finally, Australia is a mixed case. It is a parliamentarian regime with single-party majorities where the chief executive has *de facto* legislative control, too. This results in high clarity of responsibility. However, the Australian economy is in between the American and Danish economies in terms of size and openness. Australia is also an interesting case because comparative studies of electoral behavior identify it as a country where leaders traditionally have a large impact on voting (Bean, 1993; Bean & Mughan, 1989). Online appendix A offers objective indicators for each of these factors.

Study 1 relies on election studies data. These are large representative surveys designed for studying electoral behavior and fielded shortly after national elections. Election studies were screened for relevant variables in each country. Whereas these studies routinely measure vote choice and economic perceptions, only a subset includes items on candidate trait impressions. All data with items related to warmth and competence are included in the analysis. The analysis is thus based on public opinion data from seventeen elections (USA: 7 elections 1984-2008, Australia: 8 elections 1993-2013, Denmark 2 elections 2005-2007) with over 26,000 observations. To avoid overly complicated models and because the study does not have strong predictions for inter-country variation, the analysis is performed for each country separately.

[4] Furthermore, American and Danish citizens are also markedly different on cultural values such as individualism and collectivism (Nelson & Shavitt, 2002).

Measures[5]

Perceptions of the economy

The study's primary independent variable is a question asking respondents about the state of the economy in their country over the past few years (i.e. retrospective sociotropic evaluations). This variable has a particularly strong association with vote choice (Lewis-Beck & Stegmaier, 2000). That said, admittedly, this strong association is partially due to motivated reasoning – people who have favorable views of the incumbent may adjust their economic evaluations favorably (Bisgaard, 2015). As attitudes towards the incumbent are much less likely to bias respondents' evaluation of their own or their family's economic situation (i.e. pocketbook evaluations), all models are repeated with this alternative variable as a robustness check.

Warmth and competence impressions

Warmth is measured with items on the incumbent's compassion and caring for people in the US, reliability and trustworthiness in Australia and reliability in Denmark. In the US and Australia, competence is measured with items on intelligence and knowledge; in Denmark, only knowledgeable ratings are available (see Online Appendix B for an overview). Although previous research has convincingly demonstrated that all of these items tap into their respective trait dimensions (Fiske et al., 2002), variation on the items makes it difficult to compare the three countries. However, cross-country comparison is not an important ambition for this study. Indeed, relying on multiple indicators of the same concept increases external validity (Mutz, 2011). Items within countries are sufficiently consistent between elections to justify pooling data for countries. For the sake of simplicity and consistency across models, in surveys where multiple items are available for a trait dimension, the simple mean of the items is used as the indicator for that trait. This approach enables us to include surveys with only a single indicator for a trait, but cannot guard against measurement

[5] Correlation matrices and descriptive statistics of the variables for all studies are available in Appendix A at the end of the manuscript.

error (Ree & Carretta, 2006). Models using purging measurement error by estimating latent variables are run as a robustness check on the subset of the data that contain multiple indicators.

Vote for the incumbent

The main dependent variable in the models below is voting for the incumbent. In the US, this means a simple question prompting the respondents to reveal which presidential candidate they voted for. In Australia, respondents are asked about their first party preference, honoring the fact that Australian voters are asked to rank all parties. In Denmark, respondents are simply asked which party they voted for. These categorical variables are then recoded into a binary variable taking the value 1 if the respondent voted for the incumbent and 0 otherwise. Respondents who did not vote or did not reveal their votes are excluded from the analysis.

In most cases, identifying the incumbent is a straightforward task because the incumbent president or prime minister ran for office. In three of the seven American elections (1988, 2000, 2008), however, the incumbent was barred from running by term limits. In such cases, electoral and campaign experts usually identify the candidate of the incumbent party as the incumbent (Popkin, 2013).[6] Indeed, in the former two cases, the incumbent party candidates (George H. W. Bush and Al Gore) served as vice presidents in the previous administration, so identifying them as incumbents is an uncontroversial decision. Senator John McCain in 2008, however, had no function in the second Bush administration. Indeed, his relationship with Bush was outright tense, especially for a prominent fellow Republican politician. It is, therefore, unclear to what extent voters associated him with the actions of the incumbent leaders. Robustness checks are performed to ensure that identifying incumbents does not influence results (Online Appendix G).

[6] Voters seem to realize that any single individual's power is limited even if they are president or PM of a country. Accordingly, they keep track of which party holds the chief executive office and readily attribute responsibility to other leaders of that party.

Control variables

A wide range of control variables was included in the analysis to improve the consistency of estimates (Antonakis et al., 2010). Most importantly, the models control for partisanship, which is among the strongest predictors of vote and may also affect candidate evaluations or even perceptions of the economy. In the US, the standard 7-point scale of partisanship is rescaled to 0-1 and, if needed, flipped to always indicate higher identification with the incumbent party. In Australia, a new categorical variable is created that indicates whether a respondent identifies with the incumbent party (in-partisans), with the main opposition party (out-partisans) or with one of the smaller parties. In Denmark, the party identification variable measures if the respondent identifies with the incumbent party (the Liberal Party in both elections), one of the other larger parties (Social Democrats, Conservatives or Danish People's Party). Supporters of the minor parties (the Social Liberals, the Socialist People's Party, Christian Democrats, New Alliance, Centrum Party, Red-Green Alliance) are lumped together to improve model convergence.

In addition to party identification, all models control for political interest, income, age, education and sex. A variable measuring urbanization (type of hometown) is included in Australia and Denmark; in the US, the models control for race. These demographic variables may potentially confound the relationship between our variables of interest (economic evaluations, trait impressions and vote choice) by simultaneously affecting multiple variables. For example, more politically sophisticated and highly educated voters may be more critical of both the incumbent and the economy. Sex or race may affect both incumbent impressions and vote choice (especially if they share the sex or the race of the respondent). People with higher income are more likely to have positive economic evaluations and vote for a right-wing candidate, which may both strengthen and weaken the correlation in a given election depending on the incumbent's ideology *etc*. Finally, all models include fixed effects for election years to account for idiosyncratic variation in incumbent impressions and vote share between individual elections.

Analysis strategy

The two predictions of this article concern two parallel mediation paths (warmth and competence), which are estimated best with simultaneous equation modeling (SEM) (James, Mulaik, & Brett, 2006). The SEM framework also allows to freely estimate the covariance between the two mediator disturbances as constraining this to zero leads to the unjustifiable assumption that all correlation in warmth and competence impressions is explained by perceptions of the economy. With SEM, it is also easy to relax the assumption that mediation effects are normally distributed by calculating confidence intervals with bootstrapping (Preacher & Hayes, 2008).[7] Nevertheless, these models cannot offer causal evidence for the hypotheses. Instead, they provide novel evidence from seventeen elections in three countries that economic perceptions are correlated with trait impressions, which in turn are correlated with vote choice.

The *lavaan* package in R statistical software was used to estimate the models (Rosseel, 2012). To account for the binary endogenous variable, a link function is employed similarly to a standard logistic regression. It is worth noting that the comparison of logit coefficients across different models is problematic (Allison, 1999). This analysis therefore focuses primarily on whether each model provides evidence for the two predictions and whether the pattern of the results across the three countries is similar. All computer code necessary for the reproduction of these results is deposited in the article's OSF depository (https://osf.io/8srja/).

Results

Both hypotheses are firmly supported by the data (Online Appendix D). Both competence (H1: $ß_{USA}$ = 0.14, $p < 0.001$, 95%CI [0.11; 0.18], $ß_{Australia}$ = 0.10, $p < 0.001$, 95%CI [0.07; 0.12], $ß_{Denmark}$ = 0.18, $p <$

[7] Another frequently mentioned benefit of SEM is its ability to use overidentification tests to assess model fit. Unfortunately, most models in this analysis cannot utilize this benefit. Lacking latent measures or instruments, multiple path models are just identified, that is, with zero degrees of freedom, and overidentification tests cannot be performed (for the few exceptions, see Online Appendix I).

0.001, 95%CI [0.12; 0.24]) and warmth (H2: $ß_{USA}$ = 0.39, p < 0.001, 95%CI 0.34; 0.45], $ß_{Australia}$ = 0.25, p < 0.001, 95%CI [0.22; 0.28], $ß_{Denmark}$ = 0.55, p < 0.001, 95%CI [0.48; 0.64]) significantly mediate the relationship between evaluations of the economy and vote for the incumbent. Interestingly, the proportion mediated by warmth impressions is substantially and significantly larger than that of competence (Contrasts: $ß_{USA}$ = 0.25, p < 0.001, 95%CI [0.19; 0.32], $ß_{Australia}$ = 0.15, p < 0.001, 95%CI [0.11; 0.20], $ß_{Denmark}$ = 0.38, p < 0.001, 95%CI [0.28; 0.49]). The two mediators combined explain between 33 and 56 percent of economic voting in these countries, but the direct effect remains significant in each case ($ß_{USA}$ = 0.63, p < 0.001, 95%CI [0.47; 0.82], $ß_{Australia}$ = 0.68, p < 0.001, 95%CI [0.55; 0.79], $ß_{Denmark}$ = 0.55, p < 0.001, 95%CI [0.29; 0.84]). The models thereby indicate partial mediation.

Figure 1 displays the main findings of the models, focusing on the proportion mediated by the two traits in each country. It is remarkable how consistent the results are across the three countries, with 10-14 percent of economic voting mediated by competence and 28-43 percent mediated by intention. In fact, the analysis finds considerable overlap in the total economic vote effect between the three countries (Total effects: $ß_{USA}$ = 1.17, p < 0.001, 95%CI: [1.00; 1.34], $ß_{Australia}$ = 1.02, p < 0.001, 95%CI: [0.91; 1.13], $ß_{Denmark}$ = 1.28, p < 0.001, [1.01; 1.55]).

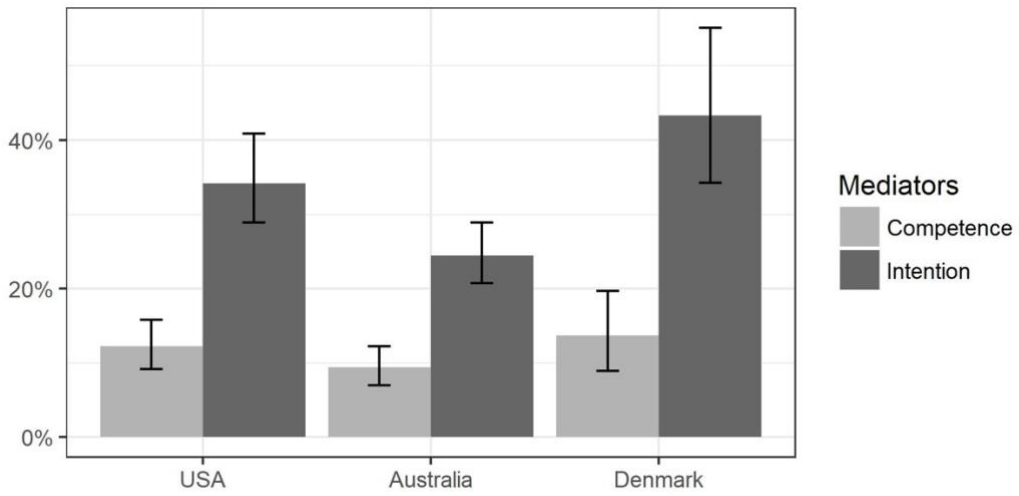

Figure 1 Warmth and to a lesser extent competence impressions mediate the economic vote effect in the USA, Australia and Denmark. The size of the dark and light grey bars show the proportion of the effect mediated by warmth and competence, respectively.

Importantly, these results are robust to a number of alternative model specifications and estimation methods. Online Appendix E demonstrates that results are essentially identical if pocketbook rather than sociotropic evaluations are used to tap into perceptions of the economy. This is important as they lend support to the claim that motivated reasoning, (i.e. supporters of the incumbent adjust their economic evaluations and trait impressions) cannot explain these results. Online Appendix F shows that results hold – in fact strengthen – if a feeling thermometer is used instead of vote choice to tap into general attitudes towards the incumbent. This is reassuring because it indicates that results generalize to less politically engaged citizens who conceal their vote choice or do not vote. Online Appendix G displays results after dropping the three elections from the analyses in which the incumbent parties nominated candidates who were not prominent politicians in the previous administration or government. This slightly strengthens the results.

Concerning issues of estimation, Online Appendix H demonstrates that statistical significance was not achieved by adding covariates. In fact, controlling for factors that influence both the mediators and the outcome attenuates the indirect effects, which is a known property of observational mediation analysis. Online Appendix I shows models run on surveys that contain multiple trait measures for at least one of the two dimensions. This allows us to rely on latent measures to tap into trait impressions without measurement error and show substantively similar results. Online Appendix I also provides a discussion of model fit, which again remind us that inferences from these models should be cautionary.

Finally, Online Appendix J establishes that the results hold if instead of election fixed effects, random intercepts are used to account for clustering in the data. These models exploit the flexibility of the *mediation* R-package (Imai, Keele, Tingley, & Yamamoto, 2011) to include multilevel models, but they suffer from the limitation of testing one mediator at a time. Nonetheless, results remain substantially similar.

Discussion

Study 1 provides firm support for both predictions and demonstrates that in line with previous predictions in the literature, competence impressions mediate the effect of economic perceptions on vote choice. It also highlights that warmth impressions also mediate the economic vote. Study 1 is high on external validity as it integrates evaluations of incumbents from seventeen elections in three countries and measures citizens' actual voting behavior. Despite the considerable institutional and economic variation between the analyzed countries, the effects are remarkably consistent. This gives credence to the generalizability of these results and provides some support for the evolutionary reasoning, too.

Admittedly, however, this analysis suffers from internal validity problems and should therefore not be interpreted as a valid test of causality. Estimating consistent causal mediation estimates from observational data is notoriously difficult (Bullock, Green, & Ha, 2010). Although the models reported control for a wide range of potential confounds, it is impossible to prove that there are no omitted variables affecting both the mediators and the outcome. In other words, the mediators are endogenous variables, and the models may suffer from endogeneity bias. Future research should investigate the possibility of relying on instrumental variables to purge endogeneity bias from the models (Antonakis et al., 2010). Another limitation of these models is that all variables are taken from the same surveys. Estimates of covariation may therefore be elevated by common method bias (Podsakoff, MacKenzie, & Podsakoff, 2012). Simple self-reports also increase the risk of overestimating the strength of associations because of reverse causation. Future tests of these relationships should build on the rich tradition in the economic voting literature to go beyond simple self-reported variables of economic perceptions (e.g. Bisgaard et al., 2016). Meanwhile, Study 2 aims at strengthening the causal claim of this article by offering experimental evidence.

Study 2: A manipulation-of-process experiment

Participants

Study 2 was conducted on a large sample of Americans recruited through Amazon's Mechanical Turk (N = 1003, age = 39 years, 63% female) to participate in a study on impression formation. Although this is a convenience sample, multiple studies demonstrate that causal estimates from MTurkers generalize well to representative samples of US citizens (Clifford, Jewell, & Waggoner, 2015; Mullinix, Leeper, Druckman, & Freese, 2016) and are not particularly prone to bias due to demand effects (Mummolo & Peterson, 2018). To further improve the quality of the sample, Turkprime's advanced feature to improve the naiveté of respondents was utilized (Litman, Robinson, & Abberbock, 2017). This feature excludes the top 5 percent most avid MTurkers who take part in over 50 percent of all studies fielded on the platform.

Design and procedure

This experiment follows a manipulation-of-process design.[8] This is a procedure for experimentally testing causal mediation hypotheses without measuring mediators (Jacoby & Sassenberg, 2011; Spencer, Zanna, & Fong, 2005). The intuition behind this approach is that a standard experiment manipulating X and measuring Y estimates the total effect of X on Y. If it is possible to block a mediation path by experimentally interrupting the respective psychological process, it is also possible to estimate the direct effect of X on Y. If there is a difference between the total and the direct effect, it can be taken as an indirect evidence that a causal mediation is present (i.e. that the indirect effect through mediator M is not zero). An important advantage of this procedure is that it does not involve measurement of the mediator(s). Accordingly, it can be applied in cases where the measurement is problematic (e.g. leading to common-source bias) or where there is a danger that the measurement

[8] Jabocy and Sassenberg (2011) refer to this procedure as *Testing-a-Process-hypothesis-by-an-Interaction Strategy* or *TPIS*, whereas Spencer and colleagues (2005) call it *moderation-of-process* experiments. I use *manipulation-of-process* for simplicity and to prevent confusion about moderation and mediation.

interferes with the proposed psychological procedure (making it more or less salient or even introducing competing mechanisms (Jacoby & Sassenberg, 2011, p. 183)).

In this experiment, perceptions of the economy and availability of trait impression cues about the incumbent are manipulated in a 2 (economy booming or struggling) × 3 (no trait cue or warmth cue or competence cue) between-subjects design. Participants were instructed to read a fictitious scenario about elections in a foreign country, imagine that they are citizens in that country and then answer some questions related to the elections. The fictitious scenario briefly described an election and gave some background about the country with strong cues about the state of the economy (e.g., *booming*/*struggling* economy, *growing*/*struggling* job market). The control group's text ended here, whereas the two trait treatment groups went on to read a description of the candidate, which either provided direct trait cues about warmth (the incumbent is a *pleasant*, *helpful* and *caring* person) or competence (the incumbent is *insightful*, *savvy* and *smart* person). To make sure that there is no spillover from one trait to the other, both manipulations included cues designed to offset such a halo affect by anchoring the competing mediator to its original value. Importantly, the trait manipulations did not directly address the incumbent's contribution or responsibility for the state of the economy.

In this procedure, participants in the control conditions may infer trait impressions from economic cues and may consequently rely on these trait impression to form a vote choice. The difference in vote between the booming and struggling economy conditions *sans* trait information establishes the total (or baseline) effect of the economy. Meanwhile, the two trait treatment conditions interrupt the standard psychological process by providing direct trait information. This rests on the straightforward assumption that the strong direct cues from the trait manipulations overshadow participants' more fuzzy inferences based on economic cues. Put differently, if presented with contradictory information wherein the economy is doing poorly, but the candidate appears to be competent or having good intention, people would privilege the latter cues when forming impressions about the candidate's competence and warmth, respectively.

To the extent that economic perceptions affect vote by carrying valid cues of the leader's traits, overriding these trait impressions should diminish the effect of the economy compared to the baseline. In other words, if keeping the trait information constant across the economy manipulations interrupts the psychological process, the remaining direct effect of the economy on vote will be smaller than the total effect.[9] In contrast, if the mediation hypotheses are false and participants do not infer trait impressions from the economic cues, providing direct trait information does not override any inferences made in the control groups. In this case, the marginal effect of the economy manipulations should stay constant across the trait conditions. Formally, a statistical interaction between the economy treatment and the manipulation of the process is consistent with the proposed mediation hypotheses.

This design offers consistent causal estimates of a mediation effect if the mediation manipulations block one mediation path at a time (Bullock et al., 2010).[10] Extensive pilot

testing ensured that compared to the control condition, the warmth manipulation does not change average competence impressions and that the high competence manipulation does not change average warmth ratings (see results below).

Manipulation checks were included only in the pilot tests. First, respondents were asked how they would describe the state of the economy in the country to tap into the economy manipulation.

[9] This mirrors the logic of traditional mediation analysis (Baron & Kenny, 1986). Controlling for the mediator diminishes the marginal effect of the independent variable. However, in sharp contrast to traditional mediation analysis, here, the mediators are held constant by experimental and not statistical control.

[10] In fact, the manipulation-of-process design is superior to an alternative design in which the independent variable is experimentally manipulated and both the mediator and the dependent variable (DV) are measured, because in the latter case, omitted variables confounding the relationship between the mediator and the DV bias the mediation estimate (Bullock et al., 2010; Spencer et al., 2005).

Second, they rated the incumbent on eight traits, four tapping into warmth (trustworthy, warm, dishonest, unfair), four into competence (intelligent, qualified, mindless, naïve). The latter two items in both sets are reverse coded. The four items showed high reliability (Cronbach's alphas: $\alpha_{competence}$ = 0.83, warmth = 0.73). Warmth and competence impressions were constructed by taking the mean of the four items.

In the experiment itself, two dependent variables were included, one measuring vote intentions and another measuring general evaluations. Because the two dependent variables yield identical results, only the former is reported here. All these items were measured on 7-point scales and recoded to a continuous 0-1 scale for the analysis. Finally, all respondents indicated their age, gender, education and ideological views. Full experimental materials are shared in the article's OSF repository (https://osf.io/8srja/).

Pilot testing

A pilot study also conducted on MTurk (N = 179, M_{age} = 37 years, 60% female) demonstrated that the treatment materials satisfy all the requirements for testing a mediation effect. First, the economy manipulation was significant and large (state of economy: $M_{struggling}$ = 0.28, $M_{booming}$ = 0.85, $p < 0.001$). Second, compared to the control group, the warmth and the competence groups had significantly higher warmth and competence impressions, respectively (warmth impressions: $M_{control}$ = 0.58, M_{warm} = 0.74, p = 0.001; competence impressions: $M_{control}$ = 0.67, $M_{competent}$ = 0.76, p = 0.02). Third, the trait manipulations did not spill over on the other trait dimension. In other words, the high competence group did not have higher warmth ratings than the control group (competence impressions: $M_{control}$ = 0.67, M_{warm} = 0.66, p = 0.85). Similarly, the warmth group did not have higher competence impressions than the control group (warmth impressions: $M_{control}$ = 0.58, $M_{competent}$ = 0.60, p = 0.62). This ensures that the causal mediation estimates are unbiased (Bullock et al., 2010; Spencer et al., 2005).

Results

Figure 2 displays the results of Study 2. Reassuringly, the results replicate the well-known economic voting effect: Participants in the booming economy condition are more likely to vote for the incumbent than those in the struggling economy condition (focusing on those who received no candidate-specific information: $M_{struggling} = 0.31$, $M_{booming} = 0.71$, $ß_{baseline} = 0.40$, $t = 17.08$, $p < 0.001$). This is crucial because the hypotheses are supported to the extent that this main effect is diminished by interrupting the psychological process with information of candidate competence or warmth.

Figure 2 The main effect of the economy manipulation (distance between grey and black points) is substantially diminished by directly manipulating candidate traits.

Do direct cues of competence diminish the effect of the economy? Yes. Describing the incumbent leader as a competent politician markedly decreases the main effect of the economy manipulation (Interaction effect: $ß_{baseline \times competent} = -0.17$, $t = -4.9$, $p < 0.001$). As expected, despite a struggling economy, participants in the high competence condition had more favorable impressions of the incumbent than participants in the control condition did (vote for incumbent in struggling economy:

$M_{control}$ = 0.31, $M_{competent}$ = 0.55, t = 9.1, p < 0.001). This lends credence to Hypothesis 1 that competence is a causal mediator between perceptions of economy and vote.

Do direct cues of warmth diminish the effect of the economy? Yes. Describing the incumbent leader as a well-intentioned, caring leader substantially decreases the main effect of the economy manipulation (Interaction effect: $ß_{baseline \times warm}$ = -0.14, *t* = -4.2, *p* < 0.001). Again, this is due to higher vote intention in the struggling economy conditions $M_{control}$ = 0.31, M_{warm}= 0.50, t = 7.4, p < 0.001). This provides firm support for Hypothesis 2 that warmth is a causal mediator between perceptions of economy and vote. Interestingly, contrary to the findings of Study 1, warmth is not a more important mediator than competence. The economy manipulation's effect is diminished to a similar extent in the high competence and the high warmth groups ($ß_{contrast}$ = -0.03, *t* = -0.84, *p* = 0.40).

Discussion

Study 2 sought to provide causal evidence for the mediation hypotheses. It succeeded for both competence and warmth. Experimentally interrupting the link between the economy and trait impressions diminished the (total) effect of the economy manipulation by 42.5 and 35 percent, respectively. These results complement the findings of Study 1. Interestingly, the relative effect sizes in the two studies are markedly different. In Study 1, the warmth path explained much more of the variance than the competence path; in Study 2, the difference in the two effects was not statistically significant. Online Appendix K considers potential explanations for this result.

Scenario experiments akin to Study 2 are sometimes criticized for being biased by experimenter demand effects. Scenarios may make it easy for participants to guess the purpose of the experiment and thus to give answers that help confirm the hypotheses (Zizzo, 2010). Could demand effects drive the results reported above? It is unlikely that participants exposed to any one condition could guess the mediation hypotheses tested here or the expected interaction pattern between four experimental conditions (difference between booming and struggling economy across trait treatments). It is more plausible that participants had less nuanced guesses, which nonetheless

correlate positively with the true experimental objectives and, thus, pushed them to confirm with expected behavior. Study 3 seeks to rule out this alternative explanation of the results and allows us to test if the results of Study 2 replicate on an independent sample.

Study 3: Ruling out demand effects

The true experimental objective in Study 2 is to show that the economic manipulation has a smaller effect in the trait treatment conditions than in the control condition. Accordingly, demand effects inflating the baseline effect of the economy manipulation or deflating the same effect under the trait treatment may both lead to false positives. Because the control condition in Study 2 focused very much on the economy, it is plausible that some participants conditioned their support for the incumbent on the economy just to help find an effect and not because they considered the economy to be a relevant cue. In this case, the trait treatments could diminish the economy effect by reducing demand effects and not by undermining the psychological process of economic voting. Meanwhile, the treatment conditions had materials both on the economy and on incumbent's traits. It is plausible – albeit less likely – that some participants ignored the economic cues or conditioned their responses on the trait cues just to help find an effect for the trait manipulation and not because they thought the trait information attenuates the importance of economic cues. The trait treatment could diminish the economy effects by introducing a novel demand effect and not by interrupting the psychological process of economic voting.

It is important to acknowledge that demand effects may also increase false negative error rates in case participants' guesses and the true objectives are negatively correlated. In particular, if participants in the trait treatment conditions guessed that the experiment is primarily about the effects of the economy, that the economy should matter more than or even as much as trait information, the hypothesized interaction effect is less likely to be found. In other words, the trait treatment may fail to diminish the economy effects even though the psychological process of

economic voting was disrupted. Although Study 3 is primarily concerned with demands increasing the likelihood of false positives, Online Appendix L investigates the possibility of the latter.

Participants

Study 3 was conducted on an independent sample of participants recruited on Amazon's Mechanical Turk (N = 499, age = 39 years, 45% female). People who took Study 2 or any of its pre-tests were excluded from the subject pool. Unlike Study 2, Study 3 did not employ Turkprime's premium feature to exclude the most avid MTurkers because of budgetary limitations. If anything, this makes the current investigation more conservative, because Study 3's more experienced subject pool is more likely to be able to guess the experimental objectives than the subject pool of Study 2.

Design and Procedure

The procedure of Study 3 is identical to Study 2, with a number of important modifications. First, two additional control groups were designed that included the same trait information as the trait treatment conditions. Importantly, however, in these placebo conditions, the trait cues described the challengers and not the incumbent prime minister. Thereby, the trait cues do not disrupt psychological process of economic voting, but participants trying to guess the purpose of the experiment rely on similar information across all experimental groups. Pre-tests were run to ensure that the challenger trait cues do not influence the incumbents' trait impressions and, hence, that the two trait treatments remain symmetrical compared to the new baseline, too. Thereby, Study 3 is a 2 (economy: booming or struggling) × 4 (no trait cues, irrelevant trait cues, warmth cues or competence cues) between-subject design.[11]

Second, towards the end of the survey, participants were invited to guess "what the experimenters conducting this study were expecting to find by asking these questions" (Mummolo & Peterson, 2018). Participants were offered a $0.25 bonus if they could correctly guess the answer in an open-

[11] Because of a coding error, there is an imbalance in the size of the experimental groups with the placebo and the control being half the size (86 and 79, respectively) of the trait treatment groups (166 and 168).

ended question. This financial incentive signals to participants a genuine interest in their honest answers and, thus, reduces potential concerns that by revealing their understanding of the experimenter demands, they would "hurt" the experiment (Zizzo, 2010). Open-ended answers were hand-coded by two naïve research assistants. First, they coded each response by the topics mentioned. In particular, they recorded whether participants mention the economy, traits or any other topic. Coders also recorded whether participants mentioning multiple topics make any relational or conditional statements, that is, whether a topic matters more or less than another or whether something matters in certain conditions but not in others and so on. Second, research assistants revisited the subset of responses, which mention the appropriate topics and coded whether the participant makes a clear statement about the experimental objective or simply reiterates materials from the vignettes. This step seeks to differentiate between attention and experimenter demand effects. Some participants may correctly guess the research question, yet fail to identify any expectations (e.g. "I believe the researchers are expecting to find out how people make their voting choices based on economic trends and money"). Conversely, other participants indicate clear ideas about what the study is expecting to find (e.g. "They expect to find that bad economic times lead to people disliking the current leadership"). Neither of these categories can refute (nor prove) that participants adjust their answers to help the experimenter, but it is less likely that they do so if they cannot articulate a directional expectation to which they are supposedly expected to conform. For additional details and analyses about the post-experimental inquiries, see Online Appendix L.

Results

Do results from Study 2 replicate on an independent sample? Yes. The baseline effect of the economic manipulation (ß = 0.50, t = 10.2, $p < 0.001$) is substantially diminished by both the competence manipulation (ß = -0.30, t = 4.92, $p < 0.001$) and the warmth manipulation (ß = -0.22, t = 3.59, $p < 0.001$). These statistically significant negative interaction effects are consistent with the mediation hypotheses (see Figure 3).

Do the trait treatments diminish the economy effect compared to the placebo? Yes. Using the new placebo condition as the baseline instead of the original control has no substantial effect on the results. Although the economy effect is slightly smaller in the placebo condition (ß = 0.41, t = 7.9, p < 0.001) than in the original control condition (ß = 0.50, t = 10.2, p < 0.001), this change is not statistically significant (ß = -0.09, t = 1.4, p = 0.18). Accordingly, both the competence manipulation (ß = -0.21, t = 3.3, p = 0.001) and the warmth manipulation (ß = -0.13, t = 2.0, p < 0.05) substantially diminish the effect of the economy manipulation compared to the placebo condition.

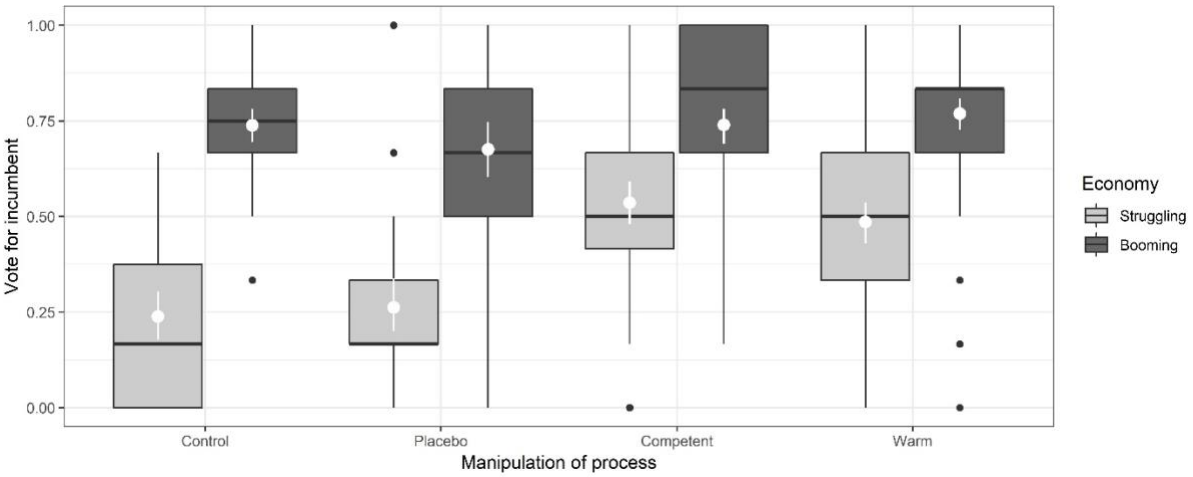

Figure 3 Study 3 replicates and extends the previous findings by showing that trait cues interrupt the psychological process only if they describe the incumbent, but not in the placebo condition, where they describe an opponent.

How many participants experienced experimenter demands? Very few. Only 10 percent of participants in the original control condition indicated that they thought that the experimenters expected to find that they would vote with the economy. The share of participants who experience experimenter demands is reduced by the placebo condition to 4 percent, which indicates that the obfuscation of the objectives was successful. Even less evidence was found for experimenter demands in the trait treatment conditions. Less than 1 percent of participants (3 out of 334) indicated that they believe the experimenters expected to find that they vote based on trait cues regardless of the economy. With so few participants articulating potentially problematic demands, it seems unlikely that the overall results can be explained by demand effects.

That said, more participants correctly identified the topic of the experiment. It cannot be ruled out that some of them (perhaps unconsciously) gave responses to demonstrate the significance of the topics they identified. There is also a third group of respondents, however, making up around a third of the sample. These include participants who explicitly said that they do not know what the experimenters expected to find and who did not admit this, but provided answers which fail to identify any (independent) variable of interest (e.g. "If I would vote for PM Bennett", "Analogies to US political feelings"). Interestingly, a chi-square test of independence shows that the proportion of these participants is stable across conditions (X^2 (3) = 1.81, p = 0.61). It is reasonable to assume that these most naïve participants are the least likely to engage in demand-like behavior, which raises the question whether the results replicate even among them.

Do results replicate among participants least likely to have experienced experimenter demands? Yes. Subsetting the data to participants who failed or refused to make a guess about the experimental objectives reduces the sample to 161 participants. Although this results in a considerable loss of power, the results remain reasonably consistent. The warmth and competence trait manipulations diminish the size of the economy effect compared to the original control (baseline: ß = 0.36, t = 3.6, p < 0.001; competence interaction: ß = -0.19, t = 1.7, p < 0.1; warmth interaction: ß = -0.18, t = 1.5, p = 0.15) and the placebo condition (baseline: ß = 0.38, t = 3.6, p < 0.001; competence interaction: ß = -0.22, t = 1.8, p < 0.1; warmth interaction: ß = -0.20, t = 1.6, p = 0.12).[12]

Discussion

The main objective of Study 3 was to investigate whether the results of Study 2 may suffer from bias because of experimenter demand effects. It is impossible to definitively prove or refute whether participants – consciously or unconsciously – adjust their answers to aid the experimenter. The results of Study 3, however, are much more consistent with a genuine mediation effect from

[12] Combining the original control and the placebo condition into a single baseline category improves power and further increases confidence that the results cannot be explained by chance: (baseline: ß = 0.36, t = 5.2, p < 0.001; competence interaction: ß = -0.20, t = 2.1, p < 0.5; warmth interaction: ß = -0.18, t = 1.9, p < 0.1).

economic perceptions to vote through trait impressions than with a methodological artefact. Study 3 demonstrated that results replicate 1) on an independent sample 2) with a new baseline condition that obfuscates the experimental objectives better and 3) even among participants least likely to experience experimenter demands. Moreover, only 3 percent of the participants in the sample perceived demands that could potentially (but not necessarily) propel them to bias their results in favor of a significant finding. These findings are consistent with recent methodological advances showing that survey experiments may rarely suffer from experimenter demand effects in large part because "with rare exceptions, respondents appear largely unable to engage in demand- like behavior" (Mummolo & Peterson, 2018, p. 27).

That said, there are limits to what a single experimental design can reveal. First, these two studies have tested but one implication of a mediation: If there are straightforward positive cues of candidate trait impressions, positive economic cues contain little information. An untested implication is that the effects are symmetrical for straightforward negative impression cues, which should diminish the effects of negative economic cues (compared to the control). Second, there are alternative experimental designs for testing mediations, which take different approaches to testing causal mediation effects. In a causal chain design, a series of experiments are run that test each link in psychological procedure 1) manipulating the economy, measuring traits, 2) manipulating traits, measuring vote intention, 3) manipulating the economy, measuring vote intention (Spencer et al., 2005). In a crossover design, a first experiment would randomize economic cues and measure both trait impressions and vote intentions. A subsequent experiment (following a washout period) would assign participants to the opposite economic treatment group, but would keep trait impressions constant for each participant at the level observed in the first experiment (see Imai, Tingley, & Yamamoto, 2013 for more technical information). Both of these designs take less assumptions than the manipulation-of-process design employed here, although this comes at a cost in both cases (not estimating the proportion mediated in the first case, employing a costly and complex design in the latter). Yet another alternative approach is relying on instruments to purge endogeneity bias in an

experiment with manipulated economy and measured trait impressions (Emsley, Dunn, & White, 2010), although finding good instruments for competence and warmth impressions could be a significant challenge. In sum, future research should investigate in more detail whether the proposed mediation hypotheses hold under alternative experimental tests.

General discussion

This article argues that citizens rely on their perceptions of the economy to inform impressions of political leaders' warmth and competence. These impressions in turn feed into general evaluations and vote decisions. Data from seventeen elections from three countries provided firm evidence that 33-56 percent of the effect of economic perception on vote choice is mediated through warmth and competence. Two original experiments designed to test the causal mediation effect corroborated these results. These findings have important practical and theoretical implications, which are discussed below.

Decades of political science research demonstrates that economic voting is one of the most important mechanisms for holding political leaders accountable. Democracies hinge on the premise that good leaders are rewarded for their performance, whereas bad leaders are voted out of office. Previous scholarship has convincingly demonstrated that citizens across cultures punish or reward incumbents for their economic performance. However, the psychological processes contributing to this social mechanism have been largely ignored. This article opens the black box of causality and offers an ultimate explanation for economic voting based on evolutionary leadership theory. It predicts and demonstrates that economic perceptions feed into both warmth and competence impressions.

This finding may help to explain the cross-cultural consistency and the large effect of economic voting. This article proposes that updating leader trait impressions based on cues of the economic well-being of the self and fellow in-group members may be part of a universal followership

psychology. It is therefore plausible that domain-specific cognitive mechanisms performing this action routinely develop in all healthy humans across the globe. Indeed, some political scientists have made uncharacteristically bold claims about the universality of economic voting (Norpoth, 1996). The consensus in the literature is perhaps more careful, but the abundance of cross-cultural evidence is undisputed (Duch & Stevenson, 2008). It is remarkable that comparative research shows that institutional (as opposed to cultural or psychological) factors explain international variation in the strength of economic voting. Coalitional governments and convoluted mechanisms of decision-making and of political control decrease citizens' ability to understand who is responsible for what. The lower the clarity of responsibility is, the weaker economic voting is in a country (Anderson, 2000; Powell & Whitten, 1993). This conclusion aligns remarkably well with the central insight of evolutionary psychology that domain-specific cognitive mechanisms are triggered as long as their inputs are reliably detected in the environment but may turn dormant if changes in the environment conceal these inputs (Tooby & Cosmides, 1992).

This article also offers some insights regarding the large effect of economic voting. Warmth and competence are universal dimensions of social perception that, *together*, explain over 80 percent of the variance in perceptions (Fiske et al., 2007). The fact that both warmth and competence impressions are affected by perceptions of the economy may go a long way towards explaining its robust impact on general evaluations and vote. In other words, this article finds that it is true that voters update their competence impressions based on the economy (Duch & Stevenson, 2008), but without the newly established warmth mechanism, economic voting may not have been able to establish its dominant status as the key tool of democratic accountability.

Followership psychology might also help to understand the limits of economic voting-centered democratic accountability. An evolved followership psychology yields normatively accurate judgments to the extent that ancestral and modern environments overlap. Whereas ancestrally, the relationship between a group's welfare and its leader's performance was relatively clear and simple

(if probabilistic), today, it is increasingly blurred. Surely, governments affect the well-being of their citizens through monetary, financial, economic and welfare policies. However, they have little control over broader trends in our globalized economies. Therefore, estimating the impact of leaders on the economy is a difficult and often controversial task even for an academic investigator. Lacking both the motivation and the resources to contemplate this issue, citizens are unlikely to weigh accurately all factors complicating this relationship.

In fact, if voters overestimate their leader's contribution to an economic trend, attention to the state of the economy in evaluating incumbents may even hinder accountability. This may occur if the relevance of other, more important cues is underestimated. For example, people might underestimate a leader's contribution to corruption and overestimate the leader's role in increasing economic well-being while judging that leader's warmth. Thus, they may arrive at the decision that on balance, the leader is worth supporting. In fact, these patterns describe recent history in Russia well. In the early 2000s, the surge in oil prices contributed to a large economic boom, and President Vladimir Putin's approval rating has risen to dizzying heights despite widely known violations of human rights and excessive corruption (Treisman, 2014). Future research should investigate what factors influence how much weight the economy is assigned in forming trait impressions and how competing cues are tallied. Resolving these issues would contribute to a better understanding of democratic accountability.

This article also implies that assuming that leader impressions are but one of the several factors determining vote choice (and not a particularly "rational" one at that) may be misguided (Bartels, 2002). From a psychological perspective, leader impressions do not simply reflect the idiosyncrasies of candidates' public image. Instead, they play a crucial role as a mediator, tallying several different types of informational cues. Hence, their relevance does not hinge exclusively on how much of the variance in vote they explain. Rather, studying leader evaluations may yield important insights about

the causal mechanisms that voters employ in forming their vote choice or in deciding to support or oppose the government.

This study has a number of limitations. Most importantly, it remains an open question whether sensitivity to cues about the economy is part of a universal followership psychology. All data in this study come from WEIRD societies (Henrich, Heine, & Norenzayan, 2010), and much more cross-cultural evidence needs to be collected to strengthen the generalizability of the findings. Given the findings in political science discussed above, this appears to be a fruitful avenue for future research.

This article is also limited in its scope. Although it finds evidence for two psychological mechanisms fueling economic voting, in most analyses, these two indirect effects explain less than 50 percent of the variance. This implies that additional causal mechanisms may be at work, which future research should identify and test. More broadly, although this article argued that trait impressions mediate the effect of economic perceptions on vote, it is plausible that there is also an interaction between the two variables. Convincing evidence shows that charisma impressions have the largest effect on vote choice when economic cues are ambiguous (Jacquart & Antonakis, 2015). It is also plausible that people may be less sensitive to leaders' qualities when the economy is doing well or when (and until) they are personally well-off. Finally, an evolutionarily informed perspective may offer novel insights to cardinal debates in the economic voting literature, for example, about the role of retrospective versus prospective perceptions or about the relevance of sociotropic versus pocketbook evaluations. Future research should investigate these questions in more detail.

Finally, from a statistical perspective, estimating causal mediation effects is notoriously difficult (Bullock et al., 2010). The three studies in this article sought to satisfy both external and internal validity requirements by combining observational and experimental evidence. Nonetheless, they remain subject to the well-known limitations of observational mediation analysis and fictitious vignette experiments, respectively. Future research should seek to come closer to the golden standard of experimental evidence with real leaders from representative samples.

Subject to these limitations, this article is an important reminder of the benefits of generating hypotheses informed by evolutionary theories. Acknowledging the reciprocal relationship between leaders and followers in ancestral societies led to the novel prediction that the economy is a cue for leaders' warmth, not just their competence. Because of the importance of economic voting for democratic accountability, these findings advance our understanding of the psychological machinery that enables people to act as democratic citizens. It also opens up new avenues for future interdisciplinary research in evolutionary political psychology.

Appendix A. Correlation matrices

Table 1. Descriptive statistics and correlation matrix of variables in USA data

	mean	sd	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1 Vote	.52	.50	-																						
2 Feeling thermometer	.59	.31	.73	-																					
3 Competence	.68	.24	.45	.56	-																				
4 Intention	.52	.29	.65	.75	.57	-																			
5 Economy (sociotropic)	.45	.29	.39	.41	.27	.39	-																		
6 Economy (pocketbook)	.54	.27	.23	.25	.17	.24	.34	-																	
7 Party ID	.49	.37	.75	.67	.41	.61	.37	.23	-																
8 Interest in politics	.65	.33	-.05	-.05	.01	-.02	-.05	-.01	-.03	-															
9 Income	2.80	1.35	.05	.03	.01	.02	.06	.15	.04	.05	-														
10 White	.79	.41	.15	.10	.08	.14	.15	.05	.14	.01	.13	-													
11 Age	47.66	16.85	.01	.03	.04	.07	-.01	-.19	.01	.10	-.18	.11	-												
12 Female	.55	.50	-.01	.01	.01	.04	-.10	-.08	-.02	-.07	-.14	-.04	.02	-											
13 Education - higher	.31	.46	.00	-.03	-.06	-.02	.07	.09	.03	.13	.24	.10	-.09	-.10	-										
14 Education - primary	.04	.20	-.03	.00	-.01	.00	-.05	-.09	-.05	-.05	-.17	-.08	.24	-.02	-.14	-									
15 Education - secondary	.36	.48	-.01	.02	.05	.02	-.06	-.07	-.03	-.11	-.18	-.03	.11	.07	-.51	-.16	-								
16 Education - some higher	.28	.45	.02	.01	.01	.00	.01	.03	.03	.01	.02	-.03	-.14	.03	-.42	-.13	-.48	-							
17 Year - 1984	.19	.39	.07	.09	.01	.02	.19	.04	-.01	-.06	.04	.07	-.09	.01	-.09	.04	.07	.00	-						
18 Year - 1988	.16	.37	.02	.06	-.01	.03	.02	.03	.01	-.03	.03	.02	-.02	-.01	-.03	.04	.03	-.02	-.21	-					
19 Year - 1992	.19	.39	-.10	-.09	.01	-.07	-.37	-.10	-.07	.08	.03	-.01	.00	.00	.00	.04	.03	-.04	-.23	-.21	-				
20 Year - 1996	.15	.35	.05	.01	.12	.09	.20	.03	.05	-.05	.00	.02	.06	-.01	.04	-.03	-.02	.00	-.20	-.18	-.20	-			
21 Year - 2000	.15	.36	.01	-.01	.05	.04	.17	.06	.05	-.04	-.08	.01	.04	.00	.07	-.05	-.08	.03	-.20	-.19	-.21	-.18	-		
22 Year - 2004	.11	.32	.00	-.03	-.20	-.06	-.04	.01	.02	.06	-.01	-.04	.02	-.01	.04	-.05	-.05	.03	-.17	-.16	-.17	-.15	-.15	-	
23 Year -2008	.05	.22	-.07	-.06	-.02	-.07	-.27	-.10	-.06	.06	-.03	-.15	.00	.02	-.03	-.02	.01	.03	-.11	-.10	-.11	-.10	-.10	-.08	-

Table 2. Descriptive statistics and correlation matrix of variables in Australian dat

	mean	sd	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1 Vote	.39	.49	-																												
2 Feeling thermometer	4.75	3.40	.62	-																											
3 Competence	.71	.24	.36	.50	-																										
4 Intention	.41	.32	.50	.71	.50	-																									
5 Economy (sociotropic)	.43	.28	.32	.40	.27	.34	-																								
6 Economy (pocketbook)	.45	.25	.22	.26	.20	.23	.52	-																							
7 Political Interest	.72	.26	.01	-.02	.14	.02	.03	.04	-																						
8 Income	.54	.29	.01	-.01	.06	-.01	.07	.20	.09	-																					
9 Age	49.55	16.33	.00	-.02	.08	.02	-.07	-.10	.24	-.29	-																				
10 Female	.51	.50	-.01	.00	-.02	.00	-.03	-.05	-.14	-.04	-.06	-																			
11 Ownparty - Else	.24	.43	-.20	-.08	-.09	-.09	-.03	-.03	-.11	-.02	-.12	-.01	-																		
12 Ownparty - In	.38	.49	.73	.58	.36	.47	.30	.20	.03	.00	.04	.00	-.45	-																	
13 Ownparty - Out	.37	.48	-.55	-.51	-.28	-.39	-.27	-.17	.07	.02	.07	.01	-.44	-.61	-																
14 Urban Rural	.14	.35	-.01	-.01	-.01	.01	-.08	-.05	.00	-.07	.02	-.01	.05	-.01	-.04	-															
15 Urban Smalltown	.08	.28	-.03	-.01	-.01	-.01	.01	.00	-.03	-.10	.05	.01	.02	-.02	.00	-.12	-														
16 Urban Midtown	.08	.27	-.02	.00	-.01	.00	-.02	-.02	-.03	-.07	.03	.00	.03	-.01	-.01	-.12	-.09	-													
17 Urban Large town	.14	.35	.00	.02	-.02	.01	.03	-.01	-.05	-.07	.00	.01	-.01	.01	.00	-.16	-.12	-.12	-												
18 Urban City	.56	.50	.04	.00	.03	-.01	.04	.05	.07	.19	-.06	.00	-.06	.02	.03	-.45	-.34	-.32	-.46	-											
19 Education Higher	.29	.45	-.03	-.03	.02	-.05	.08	.12	.13	.30	-.13	.04	.08	-.04	-.03	-.10	-.06	-.05	-.06	.17	-										
20 Education Primary	.34	.47	.01	.00	-.03	.02	-.05	-.07	-.12	-.20	.07	.14	-.04	.02	.02	.05	.02	.02	.02	-.07	-.45	-									
21 Education Secondary	.38	.48	.02	.03	.01	.03	-.02	-.04	-.01	-.09	.05	-.16	-.03	.02	.00	.04	.03	.03	.03	-.09	-.49	-.56	-								
22 Year 1993	.16	.36	.07	-.01	.04	.02	-.21	-.09	.02	.03	-.04	-.01	-.05	.04	.01	.21	-.13	.02	-.18	.05	-.10	.08	.01	-							
23 Year 1996	.09	.29	-.03	-.05	.06	.03	-.05	-.01	-.03	.08	-.09	.00	.00	-.01	.01	-.02	.03	.00	.02	-.02	-.02	.03	-.01	-.14	-						
24 Year 1998	.10	.29	.00	.05	-.07	.10	.06	.04	.01	-.05	-.08	-.02	.00	-.03	.03	-.02	.02	.01	.02	-.01	-.03	-.02	.04	-.14	-.11	-					
25 Year 2001	.10	.30	.02	.07	-.05	.06	.03	.03	-.04	.01	-.07	.01	.01	.00	.00	-.01	.00	.00	.03	-.01	-.01	.01	.00	-.15	-.11	-.11	-				
26 Year 2004	.09	.29	.04	.08	.03	.02	.19	.09	-.01	.04	-.02	.00	.01	.02	-.03	-.02	.03	-.01	.02	-.01	.01	.00	-.01	-.14	-.10	-.10	-.11	-			
27 Year 2007	.10	.29	-.01	.03	.03	-.01	.13	.05	.02	.05	.02	.01	.02	-.01	.00	-.03	.04	-.02	.03	.00	.03	-.03	.01	-.14	-.11	-.11	-.11	-.10	-		
28 Year 2010	.13	.33	.01	.03	.00	.01	.05	-.03	.02	-.09	.12	.00	-.01	.01	-.01	-.06	.00	.00	.04	.02	.02	.00	-.02	-.16	-.12	-.12	-.13	-.12	-.12	-	
29 Year 2013	.24	.43	-.07	-.14	-.03	-.16	-.10	-.04	.00	-.04	.11	.00	.02	-.02	.00	-.05	.04	.00	.04	-.02	.09	-.07	-.02	-.24	-.18	-.18	-.19	-.18	-.18	-.21	-

Table 3. Descriptive statistics and correlation matrix of variables in Danish data

	mean	sd	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1 Vote	.27	.45	-																								
2 Feeling thermometer	5.57	3.29	.55	-																							
3 Competence	.79	.22	.32	.53	-																						
4 Intention	.60	.32	.49	.77	.53	-																					
5 Economy (sociotropic)	.65	.20	.23	.33	.24	.34	-																				
6 Economy (pocketbook)	.60	.22	.12	.16	.13	.18	.33	-																			
7 Interest	.68	.24	-.04	-.06	.01	-.05	.03	-.02	-																		
8 Income	.39	.23	.10	.09	.07	.13	.14	.27	.09	-																	
9 Age	51.08	16.14	.07	.09	.08	.06	-.01	-.22	.13	-.08	-																
10 Female	.46	.50	-.08	-.14	-.05	-.14	-.13	-.08	-.11	-.19	-.02	-															
11 Party ID - Danish People'	.05	.21	-.13	.13	.03	.10	.03	-.02	.00	-.07	.05	-.05	-														
12 Party ID - Conservative P	.06	.23	-.13	.15	.10	.14	.08	.04	.04	.07	.05	-.03	-.05	-													
13 Party ID - Other party	.57	.50	-.19	-.23	-.14	-.21	-.08	.00	-.07	.01	-.17	.08	-.26	-.28	-												
14 Party ID - Sociodemocrat	.17	.37	-.26	-.30	-.16	-.26	-.15	-.09	.03	-.08	.09	.00	-.10	-.11	-.51	-											
15 Party ID - Venstre	.16	.37	.67	.45	.27	.40	.19	.08	.04	.05	.08	-.06	-.10	-.11	-.51	-.20	-										
16 Urban - Rural	.20	.40	.08	.08	.06	.07	.02	.01	-.03	-.02	.05	-.04	.02	-.01	-.05	-.02	.09	-									
17 Urban - Smalltown	.21	.41	.03	.02	.01	.02	.00	-.02	-.04	-.03	.04	.00	.02	.01	-.04	.01	.03	-.26	-								
18 Urban - Midtown	.24	.42	.01	.03	-.01	.01	.01	-.03	.03	.03	.07	.00	.01	.01	-.02	.02	-.01	-.28	-.29	-							
19 Urban - Largetown	.18	.38	-.06	-.04	-.02	-.02	-.02	.01	-.01	-.03	-.06	-.01	-.03	.01	.02	.04	-.06	-.23	-.24	-.26	-						
20 Urban - Metropolitan	.17	.38	-.07	-.10	-.06	-.08	-.02	.04	.05	.06	-.10	.04	-.02	-.02	.10	-.06	-.06	-.23	-.24	-.25	-.21	-					
21 Education Higher	.45	.50	-.04	-.05	.00	-.03	.01	.06	.20	.29	-.01	.10	-.09	.04	.13	-.11	-.03	-.08	-.08	.02	.03	.11	-				
22 Education Primary	.28	.45	-.01	-.02	-.02	-.02	-.03	-.03	-.13	-.29	-.04	-.01	.08	-.05	-.07	.08	.00	.07	.03	-.03	-.03	-.05	-.55	-			
23 Education Secondary	.28	.45	.06	.07	.03	.05	.03	-.03	-.09	-.02	.06	-.10	.02	.01	-.07	.05	.03	.02	.05	.01	-.01	-.07	-.56	-.38	-		
24 Year 2005	.35	.48	.04	.06	.05	.00	-.05	.08	-.05	-.14	-.05	-.05	.00	-.03	.00	.03	-.01	.09	.01	-.02	-.07	.00	-.12	.02	.11	-	
25 Year 2007	.65	.48	-.04	-.06	-.05	.00	.05	-.08	.05	.14	.05	.05	.00	.03	.00	-.03	.01	-.09	-.01	.02	.07	.00	.12	-.02	-.11	-1.00	-

Table 4. Descriptive statistics and correlation matrix of variables in Study 2

	mean	sd	1	2	3	4	5
1 Vote	.61	.28	-				
2 Age	39.18	12.18	.04	-			
3 Female	.63	.48	.06	.08	-		
4 Ideology	.43	.27	.02	.14	.00	-	
5 Highered	.66	.47	.00	.02	-.05	-.02	-
6 Group: Booming Control	.17	.37	.18	-.01	.00	.07	.01
7 Group: Booming Competent	.17	.37	.28	.04	.02	.04	-.01
8 Group: Booming Good.Intentions	.17	.38	.26	-.02	-.02	-.05	-.02
9 Group: Struggling Control	.17	.37	-.47	.01	.01	-.01	-.03
10 Group: Struggling Competent	.17	.37	-.09	.02	.01	.00	.01
11 Group: Struggling Good.Intention:	.16	.37	-.16	-.03	-.01	-.05	.03

Table 5. Descriptive statistics and correlation matrix of variables in Study 3

	mean	sd	1	2	3	4	5
1 Vote	.58	.29	-				
2 Age	39.06	12.93	.00	-			
3 Female	.45	.50	.00	.13	-		
4 Ideology	.41	.29	.02	.10	-.06	-	
5 Highered	.67	.47	-.01	-.01	-.02	.00	-
6 Group: Booming Control	.08	.28	.16	-.11	-.11	.03	-.07
7 Group: Booming Placebo	.08	.27	.10	.02	.08	-.02	-.05
8 Group: Booming Competent	.17	.38	.25	.01	.05	-.05	.09
9 Group: Booming Good.Intentions	.17	.37	.29	.00	-.05	.02	-.02
10 Group: Struggling Control	.09	.28	-.37	-.03	.06	-.03	.01
11 Group: Struggling Placebo	.08	.27	-.33	.07	.05	-.02	-.01
12 Group: Struggling Competent	.17	.37	-.07	.05	.00	-.02	.01
13 Group: Struggling Good.Intention:	.17	.37	-.15	-.02	-.06	.07	.01

Appendix B. Supplementary data

Additional supporting information for this article can be found online at https://doi.org/10.1016/j.leaqua.2019.05.002. All original data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/8s-rja/.

References

Aarøe, L., & Petersen, M. B. (2018). Cognitive Biases and Communication Strength in Social Networks: The Case of Episodic Frames. *British Journal of Political Science*, 1–21. https://doi.org/10.1017/S0007123418000273

Abelson, R. P., Kinder, D. R., Peters, M. D., & Fiske, S. T. (1982). Affective and semantic components in political person perception. *Journal of Personality and Social Psychology*, *42*(4), 619–630. https://doi.org/10.1037/0022-3514.42.4.619

Allison, P. D. (1999). Comparing logit and probit coefficients across groups. *Sociological Methods and Research*, *28*(2), 186–208. https://doi.org/10.1177/0049124199028002003

Anderson, C. J. (2000). Economic voting and political context: A comparative perspective. *Electoral Studies*, *19*(2–3), 151–170. https://doi.org/10.1016/S0261-3794(99)00045-1

Anderson, C. J. (2007). The End of Economic Voting? Contingency Dilemmas and the Limits of Democratic Accountability. *Annual Review of Political Science*, *10*, 271–296. https://doi.org/10.1146/annurev.polisci.10.050806.155344

Ansolabehere, S., Meredith, M., & Snowberg, E. (2014). Mecro-economic voting: Local information and micro-perceptions of the macro-economy. *Economics and Politics*, *26*(3), 380–410. https://doi.org/10.1111/ecpo.12040

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *Leadership Quarterly*, *21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution*

and Human Behavior, *34*(3), 164–175. https://doi.org/10.1016/j.evolhumbehav.2013.02.002

Baron, R. M., & Kenny, D. a. (1986). The Moderator-Mediator Variable Distinction in Social The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bartels, L. M. (2002). The Impact of Candidate Traits in American Presidential Elections. In A. King (Ed.), *Leaders' Personalities and the Outcomes of Democratic Elections* (pp. 44–69). Oxford University Press. https://doi.org/10.1093/0199253137.003.0002

Bean, C. (1993). The Electoral Influence of Party Leader Images in Australia and New Zealand. *Comparative Political Studies*, *26*(1), 111–132. https://doi.org/10.1177/0010414093026001005

Bean, C., & Mughan, A. (1989). Leadership Effects in Parliamentary Elections in Australia and Britain. *The American Political Science Review*, *83*(4), 1165–1179. https://doi.org/10.2307/1961663

Bisgaard, M. (2015). Bias Will Find a Way: Economic Perceptions, Attributions of Blame, and Partisan-Motivated Reasoning during Crisis. *Journal of Politics*, *77*(3), 849–860. https://doi.org/10.1086/681591

Bisgaard, M., Dinesen, P. T., & Sønderskov, K. M. (2016). Reconsidering the Neighborhood Effect: Does Exposure to Residential Unemployment Influence Voters' Perceptions of the National Economy? *The Journal of Politics*, *78*(3), 719–732. https://doi.org/10.1086/685088

Bittner, A. (2011). *Platform Or Personality?: The Role of Party Leaders in Elections*. Oxford: Oxford University Press.

Bøggild, T., & Petersen, M. B. (2016). The Evolved Functions of Procedural Fairness : An Adaptation for Politics. In T. K. Shackelford & R. D. Hansen (Eds.), *The Evolution of Morality* (pp. 247–276). New York: Springer.

Bor, A. (2017). Spontaneous categorization along competence in partner and leader evaluations. *Evolution and Human Behavior*, *38*(4), 468–473. https://doi.org/10.1016/j.evolhumbehav.2017.03.006

Bor, A. (2018). *Intention and Competence Impressions in Political Leader Evaluations*. Aarhus: Forlaget Politica.

Boyd, R., & Richerson, P. J. (2005). Culture and the evolution of human cooperation. In S. Levinson (Ed.), *Evolution and Culture*. Cambridge, MA: MIT Press.

Boyer, P., & Barrett, H. C. (2016). Intuitive Ontologies and Domain Specificity. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (pp. 161–179). Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119125563.evpsych105

Boyer, P., & Petersen, M. B. (2018). Folk-economic beliefs: An evolutionary cognitive model. *Behavioral and Brain Sciences*, *41*, e158. https://doi.org/10.1017/S0140525X17001960

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*(4), 550–558. https://doi.org/10.1037/a0018933

Buss, D. M. (1991). Evolutionary Personality Psychology. *Annual Review of Psychology*, *42*, 459–491.

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). "Unwilling" versus "unable": chimpanzees' understanding of human intentional action. *Developmental Science*, *7*(4), 488–498. https://doi.org/10.1111/j.1467-7687.2004.00368.x

Cassese, E. C., & Holman, M. R. (2017). Party and Gender Stereotypes in Campaign Attacks. *Political Behavior*. https://doi.org/10.1007/s11109-017-9423-7

Clifford, S., Jewell, R. M., & Waggoner, P. D. (2015). Are samples drawn from Mechanical Turk valid for research on political ideology? *Research & Politics*, *2*(4), 1–9. https://doi.org/10.1177/2053168015622072

Del Giudice, M., Gangestad, S. W., & Kaplan, H. S. (2016). Life History Theory and Evolutionary Psychology. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology.* (pp. 88–114). Hoboken, NJ: Wiley.

Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Duch, R. M., & Stevenson, R. T. (2008). *The Economic Vote - How Political and Economic Institutions*

Condition Election Results. Cambridge: Cambridge University Press.

Emsley, R., Dunn, G., & White, I. R. (2010). Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Statistical Methods in Medical Research*, *19*(3), 237–270. https://doi.org/10.1177/0962280209105014

Erickson, D. J., & Krull, D. S. (1999). Distinguishing Judgments About What From Judgments About Why: Effects of Behavior Extremity on Correspondent Inferences and Causal Attributions. *Basic and Applied Social Psychology*, *21*(1), 1–11. https://doi.org/10.1207/s15324834basp2101_1

Fair, R. C. (1978). The Effect of Economic Events on Votes for President. *The Review of Economics and Statistics*, *60*(2), 159. https://doi.org/10.2307/1924969

Fair, R. C. (2009). Presidential and Congressional Vote-Share Equations. *American Journal of Political Science*, *53*(1), 55–72. https://doi.org/10.1111/j.1540-5907.2008.00357.x

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Funk, C. L. (1996). The impact of scandal on candidate evaluations: An experimental test of the role of candidate traits. *Political Behavior*, *18*(1), 1–24. https://doi.org/10.1007/BF01498658

Funk, C. L. (1999). Bringing the Candidate into Models of Candidate Evaluation. *Journal of Politics*, *61*(3), 700–720. https://doi.org/10.2307/2647824

Glowacki, L., & von Rueden, C. (2015). Leadership solves collective action problems in small-scale societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20150010. https://doi.org/10.1098/rstb.2015.0010

Hayes, D. (2005). Candidate Qualities through a Partisan Lens: A Theory of Trait Ownership. *American Journal of Political Science*, *49*(4), 908–923.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hetherington, M. J. (1996). The Media's Role in Forming Voters' National Economic Evaluations in 1992. *American Journal of Political Science*, *40*(2), 372–395. https://doi.org/10.2307/2111629

Hibbs, D. A. (1977). Political Parties and Macroeconomic Policy. *American Political Science Review*, *71*(04), 1467–1487. https://doi.org/10.2307/1961490

Hill, K., & Kintigh, K. (2009). Can Anthropologists Distinguish Good and Poor Hunters? Implications for Hunting Hypotheses, Sharing Conventions, and Cultural Transmission. *Current Anthropology*, *50*(3), 369–378. https://doi.org/10.1086/597981

Huddy, L., & Terkildsen, N. (1993). Gender Stereotypes and the Perception of Male and Female Candidates. *American Journal of Political Science*, *37*(1), 119–147. https://doi.org/10.2307/2111526

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the Black Box of Causality: Learning About Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review*, *105*(04), 765–789. https://doi.org/10.1017/S0003055411000414

Imai, K., Tingley, D., & Yamamoto, T. (2013). Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society: Series A*, *176*(1), 5–51. https://doi.org/10.1111/j.1467-985X.2012.01032.x

Iyengar, S. (1994). *Is Anyone Responsible? How Television Frames Political Issues*. Chicago: University of Chicago Press.

Jacoby, J., & Sassenberg, K. (2011). Interactions do not only tell us when, but can also tell us how: Testing process hypotheses by interaction. *European Journal of Social Psychology*, *41*(2), 180–190. https://doi.org/10.1002/ejsp.762

Jacquart, P., & Antonakis, J. (2015). When Does Charisma Matter for Top-Level Leaders? Effect of Attributional Ambiguity. *Academy of Management Journal*, *58*(4), 1051–1074. https://doi.org/10.5465/amj.2012.0831

James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A Tale of Two Methods. *Organizational Research Methods*, *9*(2), 233–244. https://doi.org/10.1177/1094428105285144

Kelly, R. L. (2013). *The lifeways of hunter-gatherers - The foraging spectrum*. Cambridge: Cambridge University Press.

Key, V. O. (1966). *The responsible electorate: Rationality in presidential voting, 1936–1960*. Cambridge, MA: Harvard University Press.

Kinder, D. R., Peters, M. D., Abelson, R. P., & Fiske, S. T. (1980). Presidential prototypes. *Political Behavior*, *2*(4), 315–337. https://doi.org/10.1007/BF00990172

Kinder, D. R., & Sears, D. O. (1985). Public opinion and political action. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 2, pp. 659–741). New York: Random House.

Kramer, G. H. (1971). Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964. *American Political Science Review*, *65*(01), 131–143. https://doi.org/10.2307/1955049

Laustsen, L. (2017). Choosing the Right Candidate: Observational and Experimental Evidence that Conservatives and Liberals Prefer Powerful and Warm Candidate Personalities, Respectively. *Political Behavior*, *39*(4), 883–908. https://doi.org/10.1007/s11109-016-9384-2

Laustsen, L., & Bor, A. (2017). The relative weight of character traits in political candidate evaluations: Warmth is more important than competence, leadership and integrity. *Electoral Studies*, *49*, 96–107. https://doi.org/10.1016/j.electstud.2017.08.001

Laustsen, L., & Petersen, M. B. (2016). Winning Faces Vary by Ideology: How Nonverbal Source Cues Influence Election and Communication Success in Politics. *Political Communication*, *32*(3), 188–211. https://doi.org/10.1080/10584609.2015.1050565

Lewis-Beck, M. S., & Paldam, M. (2000). Economic voting: an introduction. *Electoral Studies*, *19*(2–3), 113–121. https://doi.org/10.1016/S0261-3794(99)00042-6

Lewis-Beck, M. S., & Stegmaier, M. (2000). Economic Determinants of Electoral Outcomes. *Annual Review of Political Science*, *3*(1), 183–219. https://doi.org/10.1146/annurev.polisci.3.1.183

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data

acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Livingston, R. W., & Pearce, N. A. (2009). The Teddy-Bear Effect. *Psychological Science*, *20*(10), 1229–1236. https://doi.org/10.1111/j.1467-9280.2009.02431.x

Lobo, M. C., & Curtice, J. (Eds.). (2015). *Personality Politics? The Role of Leader Evaluations in Democratic Elections*. Oxford: Oxford University Press.

Lord, R. G., & Maher, K. J. (1991). *Leadership and Information Processing: Linking Perceptions and Performance*. Cambridge, MA: Unwin Hyman.

Maier, J. (2011). The impact of political scandals on political support: An experimental test of two theories. *International Political Science Review*, *32*(3), 283–302. https://doi.org/10.1177/0192512110378056

McCurley, C., & Mondak, J. J. (1995). Inspected by #1184063113 : The Influence of Incumbents' Competence and Integrity in U. S. House Elections. *American Journal of Political Science*, *39*(4), 864–885.

McGraw, K. M. (2011). Candidate impression and evaluations. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge Handbook of Experimental Political Science* (pp. 187–200). Cambridge: Cambridge University Press.

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees Recruit the Best Collaborators. *Science*, *311*(5765), 1297–1300. https://doi.org/10.1126/science.1123007

Merriam-Webster. (2018). Economic. Retrieved from https://www.merriam-webster.com/dictionary/economic

Moskowitz, D., & Stroh, P. (1994). Psychological sources of electoral racism. *Political Psychology*, *15*(2), 307–329. https://doi.org/10.2307/3791741

Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2016). The Generalizability of Survey Experiments. *Journal of Experimental Political Science*, *2*(2), 109–138. https://doi.org/10.1017/XPS.2015.19

Mummolo, J., & Peterson, E. (2018). Demand Effects in Survey Experiments: An Empirical Assessment. *American Political Science Review*, 1–13. https://doi.org/10.1017/S0003055418000837

Murray, G. R. (2014). Evolutionary preferences for physical formidability in leaders. *Politics and the Life Sciences*, *33*(01), 33–53. https://doi.org/10.2990/33_1_33

Mutz, D. C. (2011). *Population-Based Survey Experiments*. Princeton, NJ: Princeton University Press.

Mutz, D. C., & Mondak, J. J. (1997). Dimensions of Sociotropic Behavior: Group-Based Judgements of Fairness and Well-Being. *American Journal of Political Science*, *41*(1), 284–308. https://doi.org/10.2307/2111717

Nelson, M. R., & Shavitt, S. (2002). Horizontal and vertical individualism and achievement values: A multimethod examination of Denmark and the United States. *Journal of Cross-Cultural Psychology*, *33*(5), 439–458. https://doi.org/10.1177/0022022102033005001

Newman, B. J., Velez, Y., Hartman, T. K., & Bankert, A. (2015). Are citizens "receiving the treatment"? Assessing a key link in contextual theories of public opinion and political behavior. *Political Psychology*, *36*(1), 123–131. https://doi.org/10.1111/pops.12069

Norpoth, H. (1996). The Economy. In L. LeDuc, R. G. Niemi, & P. Norris (Eds.), *Comparing Democracies - Election and Voting in Global Perspective* (pp. 299–318). Thousand Oaks, CA: Sage Publications.

Oliver, P. (1980). Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations. *American Journal of Sociology*, *85*(6), 1356–1375. https://doi.org/10.1086/227168

Ostrom, E., & Walker, J. (Eds.). (2005). *Trust and Reciprocity: Interdisciplinary Lessons for Experimental Research*. New York, NY: Russell Sage Foundation.

Petersen, M. B. (2016). Evolutionary Political Psychology. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology*. Hoboken, NJ: Wiley.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of Method Bias in Social

Science Research and Recommendations on How to Control It. *Annual Review of Psychology*, *63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Popkin, S. L. (2013). *The Candidate. What it Takes to Win - and Hold - the White House*. Oxford: Oxford University Press.

Powell, B. G., & Whitten, G. D. (1993). A Cross-National Analysis of Economic Voting: Taking Account of the Political Context. *American Journal of Political Science*, *37*(2), 391–414.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Price, M. E., & Van Vugt, M. (2014). The evolution of leader-follower reciprocity: the theory of service-for-prestige. *Frontiers in Human Neuroscience*, *8*(363), 1–17. https://doi.org/10.3389/fnhum.2014.00363

Rapoport, R. B., Metcalf, K. L., & Hartman, J. A. (1989). Candidate Traits and Voter Inferences : An Experimental Study. *Journal of Politics*, *51*(4), 917–932. https://doi.org/10.2307/2131541

Ree, M. J., & Carretta, T. R. (2006). The Role of Measurement Error in Familiar Statistics. *Organizational Research Methods*, *9*(1), 99–112. https://doi.org/10.1177/1094428105283192

Rosseel, Y. (2012). lavaan : An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Rule, N. O., & Tskhay, K. O. (2014). The influence of economic context on the relationship between chief executive officer facial appearance and company profits. *Leadership Quarterly*, *25*(5), 846–854. https://doi.org/10.1016/j.leaqua.2014.01.001

Safra, L., Algan, Y., Tecu, T., Grèzes, J., Baumard, N., & Chevallier, C. (2017). Childhood harshness predicts long-lasting leader preferences. *Evolution and Human Behavior*, *38*(5), 645–651. https://doi.org/10.1016/j.evolhumbehav.2017.05.001

Simonovits, G. (2015). An Experimental Approach to Economic Voting. *Political Behavior*, *37*(4), 977–994. https://doi.org/10.1007/s11109-015-9303-y

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*(6), 845–851. https://doi.org/10.1037/0022-3514.89.6.845

Stein, R. T. (1975). Identifying emergent leaders from verbal and nonverbal communications. *Journal of Personality and Social Psychology*, *32*(1), 125–135. https://doi.org/10.1037/h0076842

Stulp, G., Buunk, A. P., Verhulst, S., & Pollet, T. V. (2013). Tall claims? Sense and nonsense about the importance of height of US presidents. *Leadership Quarterly*, *24*(1), 159–171. https://doi.org/10.1016/j.leaqua.2012.09.002

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

Tooby, J., & Cosmides, L. (1992). The Psychological Foundations of Culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the generation of culture*. New York: Oxford University Press.

Treisman, D. (2014). Putin's popularity since 2010: Why did support for the Kremlin plunge, then stabilize? *Post-Soviet Affairs*, (June 2015), 1–19. https://doi.org/10.1080/1060586X.2014.904541

van Leeuwen, F., Park, J. H., & Penton-Voak, I. S. (2012). Another fundamental social category? Spontaneous categorization of people who uphold or violate moral norms. *Journal of Experimental Social Psychology*, *48*(6), 1385–1388. https://doi.org/10.1016/j.jesp.2012.06.004

Van Vugt, M., & Ahuja, A. (2010). *Selected: Why some people lead, why others follow, and why it matters*. London: Profile Books.

Van Vugt, M., & De Cremer, D. (1999). Leadership in social dilemmas: social identification effects on collective actions in public goods. *Journal of Personality and Social Psychology*, *76*(4), 587–599. https://doi.org/10.1037/0022-3514.76.4.587

Van Vugt, M., & Grabo, A. E. (2015). The Many Faces of Leadership: An Evolutionary-Psychology

Approach. *Current Directions in Psychological Science*, *24*(6), 484–489. https://doi.org/10.1177/0963721415601971

Van Vugt, M., Hogan, R., & Kaiser, R. B. (2008). Leadership, followership, and evolution: some lessons from the past. *The American Psychologist*, *63*(3), 182–196. https://doi.org/10.1037/0003-066X.63.3.182

Van Vugt, M., & Kurzban, R. (2007). Cognitive and Social Adaptations for Leadership and Followership: Evolutionary Game Theory and Group Dynamics. In J. Forgas, M. Haselton, & W. von Hippel (Eds.), *The evolution of the social mind: Evolutionary psychology and social cognition* (pp. 229–244). New York: Psychology Press.

von Rueden, C., Gurven, M., & Kaplan, H. (2011). Why do men seek status? Fitness payoffs to dominance and prestige. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1715), 2223–2232. https://doi.org/10.1098/rspb.2010.2145

von Rueden, C., Gurven, M., Kaplan, H., & Stieglitz, J. (2014). Leadership in an Egalitarian Society. *Human Nature*, *25*(4), 538–566. https://doi.org/10.1007/s12110-014-9213-4

Wilkin, S., Haller, B., & Norpoth, H. (1997). From Argentina to Zambia: A world-wide test of economic voting. *Electoral Studies*, *16*(3), 301–316. https://doi.org/10.1016/S0261-3794(97)00028-0

Winter, N. J. G. (2010). Masculine Republicans and Feminine Democrats: Gender and Americans' Explicit and Implicit Images of the Political Parties. *Political Behavior*, *32*(4), 587–618. https://doi.org/10.1007/s11109-010-9131-z

Wojciszke, B., & Klusek, B. (1996). Moral and competence-related traits in political perception. *Polish Psychological Bulletin*, *27*(1), 319–325.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75–98. https://doi.org/10.1007/s10683-009-9230-z