Trump supported it?! A Bayesian source credibility model applied to appeals to specific American presidential candidates’ opinions

Jens Koed Madsen (j.madsen@bbk.ac.uk)
Department of Psychological Sciences, Birkbeck, University of London, Malet Street, Bloomsbury, London, WC1E 7HX
School of Geography and the Environment, University of Oxford, Manor Road, OX1 3UJ

Abstract

The credibility of politicians is crucial to their persuasiveness as election candidates. The paper applies a parameter-free Bayesian source credibility model (integrating trustworthiness and epistemic authority) in a real-life test predicting participants’ posterior belief in the goodness of an unnamed policy after a named candidate has publically supported or attacked it.

Two studies test model predictions against policy support and attack of five presidential candidates from the USA. Model predictions were measured against observed posterior belief in the goodness of the policy.

The results strongly suggest the model captures essential traits of how participants update beliefs in policies given appeals to a candidate’s support of attack. Further, individual differences suggest that people consider other factors than the ones elicited for the study. More studies into appeals to specific candidates are warranted to construct more accurate models of the influence of source credibility on political reasoning.

Keywords: Bayesian source credibility, trustworthiness, epistemic authority, political reasoning

Introduction

Source credibility influences a range of human cognitive phenomena such as the reception of persuasive messages (Petty & Cacioppo, 1984; Chaiken & Maheswaran, 1994; Tormala & Clarkson, 2007), the development of children’s perception of the world (Harris & Corriveau, 2011), legal reasoning (Lagnado et al., 2013), decision-making (Birnbaum et al., 1976), adherence with persuasion strategies (Cialdini, 2007), and how people are seen in social situations (Fiske et al., 2007; Cuddy et al., 2011).

The normative function and role of source credibility in argumentation is still subject to debate. The dual-process-based Elaboration-Likelihood Model (Petty & Cacioppo, 1981) takes sources as heuristic rather than analytic cues (Petty & Cacioppo, 1984, Briñol & Petty, 2009). Comparatively, recent Bayesian models place source credibility within a rational paradigm (e.g. Hahn et al., 2012; Harris et al., in press). The latter predicts the convincingness of an appeal to expert opinion from a Bayesian perspective as a product of the perceived trustworthiness and epistemic authority of the source (Bovens & Hartmann, 2003; Hahn et al., 2009).

For the purpose of the present paper, source credibility is defined as an amalgamation of epistemic authority and trustworthiness\(^1\). Epistemic authority is “the authority of those with superior knowledge in a specific field – experts” (Harris et al., in press). This differs from administrative authority, which refers to people who have authority bestowed upon them. For example, a judge is an epistemic authority on legal matters whilst a police officer is an administrative authority. Trustworthiness refers to the likelihood that a person will not deliberately present wrong or misleading information. Thus, where epistemic authority refers to the capability of providing good information, trustworthiness refers to the intention to actually do so.

Other sources influence how we perceive the credibility of sources, e.g. facial configurations (Rezlescu et al., 2012), the gender of the candidate (see later), as well as a host of other factors. For example, a misogynist may believe that a woman is both trustworthy and capable, but may still refuse advice from the woman solely based on her gender. It is therefore important to stress that the scope of the model tested only considers trustworthiness and epistemic authority rather than a richer source credibility account.

Whilst the persuasive and behavioural influence of epistemic authority has not been studied extensively in politics (although it would be a highly interesting study given recent anti-governmental sentiments in many democracies), it is clear that trustworthiness is an important factor in politics. It increases compliance with public policy (Ayers & Braithwaite, 1992), influences the choice of political candidate (Hetherington, 1999), increases intention of voting (Householder & LaMarre, 2014), increases societal cooperation (Fukuyama, 1995), and lack of trust may instigate civic participation (see Levi & Stoker, 2000).

How to determine if a person is a trustworthy politician remains an open issue. In political science literature, the main facets cited to describe trustworthiness are integrity, competence, fairness, flip-flopping, honesty, equitable, and being responsiveness to public needs (Miller & Listhaug, 1990; Levi & Stoker, 2000), which mirrors definitions in rhetoric (e.g. McCreoskey, 1997), cognitive and social psychology (e.g. Bovens & Hartmann, 2003; Hahn et al., 2009; Fiske et al., 2007; Cuddy et al., 2011), and reasoning theory (Walton, 1997).

Given the influence of source credibility on politics, it is reasonable to assume that the perception of an election candidate impacts multiple factors for voters. First, it may modulate their beliefs and support for particular policies. Second, trust in a candidate or government correlates with a person’s willingness to adhere to official policies. Third, low credibility spurs anti-incumbent voting behavior.

---

\(^1\) Walton (1997) argues for six traits of source credibility. However, Harris et al (in press) suggest that, in absence of other interlocutors, two may suffice. The model employed here is similar to the one tested in Harris et al.
The studies in the paper test a Bayesian source credibility model that uses trustworthiness and epistemic authority to predict the persuasive potential of references to a candidate. All model parameters are elicited (prior beliefs and conditional probabilities), generating parameter-free predictions subsequently compared against posterior beliefs in the goodness of a policy. This ensures that there is no a posteriori model fitting to observed data. To the author’s knowledge, this is the first study to apply the general Bayesian model in a parameter-free way to a political context to test the persuasive influence of appeals to specific political candidates’ positions on policies.

Modelling source credibility
Bayesian approaches to reasoning take point of departure in subjective, probabilistic degrees of belief in propositions where the posterior degree of belief in a proposition is captured by Bayes’ theorem (Oaksford & Chater, 2007; Howson & Urbach, 1996). The approach is suggested as an alternative to logicist approaches to reasoning (Oaksford & Chater, 1991) and has been applied to argumentation theory (Hahn & Oaksford, 2006; 2007, see also Oaksford & Hahn, 2004; Corner et al., 2011; Harris et al., 2012). The findings suggest Bayesian reasoning can account human information integration in practical reasoning.

Bovens and Hartmann provide a formal foundation for a Bayesian source credibility model (2003, see also Schum, 1981; Hahn et al., 2012). It integrates epistemic authority and trustworthiness and provides predictions for the posterior degrees of belief in the hypothesis (H) given the representation (Rep) by a source:

$$P(H|Rep) = \frac{P(Rep|H) \times P(H)}{P(Rep|H) \times P(H) + P(\neg H) \times P(Rep|\neg H)}$$

Harris et al. (in press) test the model using dialogues and show a good fit between predictions and observed convinciness. The current studies extend the model tested in Harris et al. (in press) by applying the model to predict the convinciness of appeals to specific and known experts in a political domain (election candidates).

Rather than a dichotomous description of an unknown source as completely trustworthy/untrustworthy, the current studies elicit prior trustworthiness and epistemic authority beliefs about five presidential candidates from the USA. The study makes use of real-life candidates and thus relies on participants’ subjective beliefs about these rather than making use of fictitious and abstractly described sources. As such, the study is an extension of the empirical test of the Bayesian model, as it lodges the dialogue in a more natural setting.

Given the previous success of the Bayesian model in Harris et al. (in press), we hypothesise that the model can predict how convinced the electorate will be given their prior beliefs of each of the five candidates together with their conditional probabilities. The model captures different degrees of posterior convinciness depending on the prior beliefs regarding the candidate in question. Thus, if people have low prior beliefs in the trustworthiness and epistemic authority of a particular candidate (e.g. Donald Trump), the model predicts that the persuasiveness of that particular candidate will be low.

Given previous empirical support for the model, we predict that the model will be able to account for a significant amount of the observed posterior degree of belief in the goodness of the policy on a population level. On an individual level, however, we predict greater noise due to the fact that, as mentioned in the above, estimations of other people involve other factors than the two central characteristics measured here and due to the fact that individual predictions are noiser than population estimations in general. Nonetheless, we predict the model to be significantly positively correlated with individual observations as well.

Study 1: Method, design, and respondents
Study 1 tests the predictive potential of the Bayesian source credibility model against appeals to specific candidates when they publicly endorse or attack a policy. This mimicks political discourse in which opinions are formed, not necessarily on the basis of evidence for or against a given policy, but on the basis that a politician that a person trusts and finds expert has supported the policy (mutatis mutandis if a person finds a candidate untrustworthy and inexerpt recommends a policy, it may be grounds for dismissal of the policy). Such reasoning is frequently observed in political debates and it is important to understand the persuasive potential of simply referring to the opinions of known political figures (e.g. a person who finds Bernie Sanders highly credible might believe a policy to be good because Sanders publically supported it).

Design and method
The candidates tested in the study were from the American race for presidential nomination for the 2016 election. Five candidates were included in the study. Candidate choice was influenced by prominence and contemporary attention. This yielded two Democratic (Hilary Clinton and Bernie Sanders) and three Republican candidates (Jeb Bush, Marco Rubio, and Donald Trump)\(^2\).

In order to generate parameter-free predictions, prior beliefs for candidates as well as all conditional probabilities were collected from each participant. The source credibility model merges epistemic authority, P(E), trustworthiness, P(T), and hypothesis, P(H) (see fig. 1).

---

\(^2\) P(Rep|H) = P(Rep|H, exp, T) \times P(Exp) \times P(T) + P(Rep|H, \neg exp, T) \times P(\neg Exp) \times P(T) + P(Rep|H, \neg exp, \neg T) \times P(\neg Exp) \times P(\neg T) + P(Rep|H, exp, \neg T) \times P(Exp) \times P(\neg T); mutatis mutandis for P(Rep|\neg H)

\(^3\) Clinton, Sanders, Bush, and Rubio were expected to be front-runners. Trump was included, as he enjoyed a lot of attention at the time of the study (July 2015).
To elicit relevant prior beliefs relating to source credibility of each politician, participants were asked how trustworthy and politically expert they believe each candidate to be. Participants responded on a sliding scale, ranging from 0 to 100 (0 was complete disagreement with the statement, 100 was complete agreement with the statement). Results varied greatly between candidates with the highest average trust score given to Sanders (61.73) and the lowest to Trump (18.36) and the highest average expertise score given the Clinton (75.63) and the lowest to Trump (19.58).4

Conditional probabilities were elicited, which represent the likelihood that a person would represent something to be true if the person is trustworthy or not, expert or not, and if the hypothesis is actually true or false (e.g. for P(Rep|T, E, H) participants valued the likelihood that a completely trustworthy and politically expert person would support a policy in a world where that policy happens to be good). Table 1 presents the average estimations of the conditional probabilities later used to calculate posterior predictions.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>80.38</td>
<td>58.21</td>
<td>34.63</td>
<td>18.04</td>
<td>22.59</td>
<td>42.3</td>
<td>59.90</td>
<td>71.26</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Conditional probabilities (study 1)

Model predictions were measured against posterior degrees of belief in the goodness of an unknown policy given public support or attack from a candidate. In line with previous studies on Bayesian argumentation (e.g. Harris et al., 2012; Harris et al., in press), P(H) was assumed to be 0.5. In order to convey this assumption, the interlocutor in the study explicitly states that she has “...no idea whether this policy is good or bad”. Having heard this statement, the other interlocutor argues that the policy is good given the fact that “CANDIDATE has publically supported/attacked the policy”. The participant was then asked how likely it is that the policy is good given the candidate’s statement. In total, the participants read ten dialogues (attack and support for each candidate).

The posterior degree of belief in the goodness of the policy was compared against predictions from the model as described in the above. Given the fixation of P(H), there are no free parameters in the model and consequently no possibility for posterior model fitting.

The results are divided into two main sections. The first section looks at averages across the participant population. This tests if the model is capable of predicting posterior degrees of belief in the goodness of policies given public support of attack from named election candidates. The second section looks at predictions from each individual.

Predicted posterior degrees of belief in the goodness of a policy given support were calculated via the above equation for P(H|Rep). Predictions for P(H|¬Rep) (i.e. if a candidate publically attacked the policy) P(H|¬Rep) was calculated from reversed conditional probabilities, e.g. P(¬Rep|H, E, T) = 1 - P(Rep|H, E, T). The conditionals allow for this, as the full probabilistic range was elicited for all conditions.

No gender or age effects were observed for prior beliefs of trustworthiness or expertise for any of the candidates or for conditional probability estimations. However, as will be discussed later, a one-way ANOVA shows that a gender effect was observed on posterior degrees of belief in the goodness of the policy for Hilary Clinton (but not for any other candidate), F (1,251) = 4.984, p = 0.026.

Population predictions

Posterior degrees of belief in the goodness of policies were non-trivial, as participants did indeed take the stance of the candidate into consideration when evaluating the likelihood that the policy was good given the appeal to the opinions of a specific candidate. The average posterior degrees of belief in the policy in the support condition ranged from 31.39 (Trump) to 65.32 (Sanders). Similarly, the average posterior degree of belief in the attack condition ranged from 36.84 (Sanders) to 63.51 (Trump). This suggests that, for preferred candidates, an endorsement is persuasive such that people update their beliefs in the likelihood that the policy will be good. Conversely, if disliked candidates

---

4 As discussed later, most participants identified as Democrats. As could be expected, prior beliefs divided across party lines with self-identified Republicans rating Republican candidates higher and mutatis mutandis for self-identified Democrats.

5 For the purpose of the current study, the non-equal political distribution does not pose a problem, as the difference will be expressed through their prior beliefs in the trustworthiness and epistemic authority of each candidate. Indeed, the results show that participants favour candidates from their preferred party.

6 The time of data collection is relevant, as developments in the candidates’ campaigns could entail significant changes in estimations of candidates’ trustworthiness and epistemic authority.
endorse a policy, people update their belief in the opposite direction and believe the policy to be bad.

A linear regression suggests the parameter-free Bayesian source credibility model captures essential characteristics of how the population update their beliefs given the support or attack of a particular candidate. The $R^2$ of the model predictions against observed average posterior belief is .824 ($p < 0.001$).

**Individual predictions**

Individual predictions are expected to be noisier. As mentioned previously, other factors than trustworthiness and epistemic authority influence how we see and react to other people. The gender difference for posterior beliefs in policies given Clinton’s statements suggest that being a woman influences how men and women react to Clinton’s statements. It is worth noting that this effect is not found for male candidates, suggesting that only being a female candidate influences the posterior degree of beliefs. It would be worth conducting a study on sexism in political reasoning. Other factors may be relevant for each candidate (e.g. Bush being the brother and son of former presidents, Trump being a Washington outsider, Sanders being a socialist, Rubio being Latino, Clinton being married to a former president, etc.). The model does not include these personal differences, but they will, to some degree, influence some voters regarding their persuasive potential.

Also, individual errors cannot be ameliorated given the parameter-free model. Thus, if a participant misunderstands the conditional probability questions, model predictions will be wrong. In some cases, this was observed, as some participants either flat-lined (50 for all conditionals) or reversed the scores presented in Table 1. These have not been excluded for the current analyses, which reduces the predictive potential of the model on an individual level.

Despite the fact that individual predictions are difficult to model, the parameter-free Bayesian source credibility is able to capture a significant correlation in the support and attack conditions. For support, a linear regression between observed and predicted posterior beliefs yield an $R^2$ of .462 for support ($p < 0.001$) and .317 for attack ($p < 0.001$). As expected, the model accounts for less of the variance, but remains highly significantly correlated with observed posterior beliefs.

**Study 2: Method, design, and respondents**

Study 1 suggests the model can be applied successfully as a predictor of the persuasive impact of appeals to specific candidates for an unknown policy. The parameter-free predictions were strongest on a population level and, as may be expected, less strong on an individual level. Study 2 tests if the model can be applied to sequential political reasoning. First, the participant is told that one candidate has publically supported or attacked the policy (replication of study 1). Having provided their posterior belief, the participant is then told that another candidate has publically supported or attacked the policy.

**Design and method**

Study 2 follows the same method as study 1. Priors and conditional probabilities were elicited from participants (conditionals showed similar response patterns as study 1, see table 1 on p. 3). To limit sequential combinatorics, two candidates were chosen (Jeb Bush and Bernie Sanders). They were chosen, as they represented the highest scoring Democratic and Republican candidate from study 1.

It was a 2 (candidate #1 attack/support) x 2 (candidate #2 attack/support) x 2 (order of candidates) design. Half of the participants saw Bush as candidate #1 and Sanders as #2. Participants saw four dialogues with the order of presentation as a between subjects condition.

Model predictions were identical to study 1. However, rather than a one-off elicitation of belief in the policy after each dialogue, participants were asked for two posterior degrees of belief. The first posterior degree of belief was elicited after candidate #1 supported or attacked the policy. The second was elicited after candidate #2 was referenced. As in study 1, initial $P(H)$ was assumed to be 0.5. In order to capture the dynamic element of the paradigm, $P(H|Rep_1)$ or $P(H|\neg Rep_1)$ was taken as $P(H)$ for the subsequent representation ($P(H|Rep_2)$). This allowed for a dynamic belief updating mechanism to be tested against observed degrees of belief in the goodness of the policy.

**Respondents**

511 participants were recruited from Mechanical Turk from the 26th to the 31st of August. Respondents were paid an equivalent of $9/hour to participate. 48.0% were women; mean age was 37.14. The political affiliation was similar to study 1: 56.2% Democrats, 20.7% Republicans, 1.8% tea party, 4.9% libertarian, 16.4% undecided. No effect of gender or age was observed on prior beliefs, conditional probabilities, and posterior degrees of belief were observed.

**Study 2: Results**

Similar to study 1, population and individual observations were compared against model predictions. Responses and model predictions after candidate #1 were identical to the ones from study 1. Consequently, the subsequent analysis focuses on the fit between responses and predictions after candidate #2 has provided his opinion.

**Population predictions**

Notably, Harris et al (in press) report increased noise when introducing interlocutors who disagree with the advice given by the expert. This is further made complicated by the real-life nature of the dialogues. As opposing politicians tend to disagree on crucial political points (especially as political polarisation has increased in American politics), participants may have found the support-support and the attack-attack conditions confusing.

Despite a noisier dialogical structure in which two rather than one candidate offer their opinion on a particular policy, the model has a good fit with observed posterior
A linear regression shows that the model predicts responses on a population level ($R^2 = .690$, $p < 0.001$)

A paired-sample t-test shows an order effect in three of four conditions: BS to SS v SS to BS: mean diff (12.44), $t = 5.405$, $p < 0.001$; BA to SS v SS to BA: mean diff (7.02), $t = 3.501$, $p = 0.003$; BA to SA v SA to BA: mean diff (-7.66), $t = 3.550$, $p < 0.001$). This may be due to a recency effect coupled with the estimation of the candidate, as the order effect was observed when Sanders ended the dialogue. As in study 1, most of the participants were Democrats and Sanders was rated higher in trustworthiness and expertise. Consequently, his opinion may have been given greater weight overall. Further studies in order effects for sequential political opinions should explore this further.

**Individual predictions**

As with study 1, individual predictions were expected to have a less good fit with observed data. As previous, some participants provided different answers to the conditional probabilities than expected. Further, order effects were observed, which further challenges individual predictions.

Despite the above considerations, the model had a good fit with individual responses. Linear regression analyses show the model accounting for 16.3% (Bush to Sanders, $p < 0.001$) and 19.3% (Sanders to Bush, $p < 0.001$).

**General discussion**

The studies applied a general Bayesian source credibility model to appeals to known and named political election candidates involved in the 2016 presidential election in the USA. This provides a real-life test if the model can predict the persuasive potential of stating that the candidate has publicly supported or attacked a particular policy (e.g. 'you shouldn’t like this policy because Donald Trump has publicly supported it'). The model had no free parameters, as prior beliefs regarding trustworthiness and epistemic expertise, conditional probabilities, and posterior degrees of belief in the goodness of the proposed policy given the public support or attack were measured. Study 1 tests appeals to five known candidates (Clinton, Sanders, Bush, Rubio, and Trump) whilst study 2 tests sequential appeals to two known candidates (Bush and Sanders).

Overall, the results suggest that the Bayesian model enjoys a good fit with observed data on a population level and, to a lesser extent, on an individual level. On the population level, the model accounts for 82.4% and 69.0% of the variance in study 1 and 2 respectively. Individual differences were expected to be noisier, as variables that may influence the judgment of one person may be irrelevant to another. Despite greater noise, the model accounts for 31.7-46.2% of the variance in study 1 and 16.3-19.3% in the more complex study 2. This indicates that other variables influence the persuasive potential of election candidates (e.g. as a gender effect was observed for the persuasiveness of Clinton despite the fact that no differences were observed in the estimations of her trustworthiness and expertise, it is reasonable to assume that being a woman influences how persuaded people were of appeals to Clinton). Future studies should determine these additional variables more concretely (variables may vary depending on the culture in question).

Alongside a good model fit, study 2 tested sequential reasoning from appeals to specific candidates. Although only tentative, the results suggest that participants were influenced by a recency effect such that the statement of the final election candidate was more influential than the statement of the first candidate. Although displaying general tendencies that are predictable from the Bayesian model, this suggests that participants were experiencing heuristic biases and therefore are not perfectly Bayesian.

Descriptively, the data suggests that participants may have been expecting political disagreement between candidates from different parties (given recent polarization in American politics, such an assumption may not be unreasonable). In the four cases in which politicians disagreed on the policy (e.g. participants were told that Sanders supported the policy while Bush attacked it), the summed differences between mean predictions and mean observations were 8.31. Comparitively, the summed differences when politicians agreed with each other were 32.51 across four conditions. Although descriptive, this suggests participants were expecting disagreement between the candidates and agreement introduced additional noise into their responses.

**Concluding remarks**

The paper tested a Bayesian source credibility model integrating expertise and trustworthiness to determine the persuasive potential of an appeal to that particular source.

Overall, the model enjoys a good fit with with observed estimations of the likelihood of the goodness of an unknown policy given public support of attack of a candidate. As expected, the model has a better population than individual fit. However, results from the studies suggest that the Bayesian source credibility approach applicable to appeals to specific and known political candidates as a model to predict the persuasive potential of the person. The paper extends previous literature by applying a parameter-free Baysian source credibility model in political context when a named candidate has publically supported or attacked a policy.

More research is required to determine individual differences for integrating evidence from uncertain sources, but that is beyond the scope of the present paper.

**Acknowledgments**

The Danish Council of Independent Research, DFF – 1329-00021B, supported the research. I am grateful to Dr. Adam Harris for very useful comments on an earlier draft.
References


Harris, A., Hsu, A. & Madsen, J. K. (2012) Because Hitler did it! Quantitative tests of Bayesian argumentation using Ad Hominem, Thinking & Reasoning 18 (3), 311-343


