Group differences in broadness of values may drive dynamics of public opinion on moral issues

Kimmo Eriksson and Pontus Strimling

Original Publication:

Kimmo Eriksson and Pontus Strimling, Group differences in broadness of values may drive dynamics of public opinion on moral issues, 2015, Mathematical Social Sciences, (77).
http://dx.doi.org/10.1016/j.mathsocsci.2015.06.004

Group differences in broadness of values may drive dynamics of public opinion on moral issues

Kimmo Eriksson

(Corresponding author) Centre for the Study of Cultural Evolution, Stockholm University, SE-106 91 Stockholm, Sweden, and School of Education, Culture and Communication, Mälardalen University, SE-721 23 Västerås, Sweden. email: kimmoe@gmail.com

Pontus Strimling

Institute for Analytical Sociology, Linköping University, SE-581 83 Linköping, and Centre for the Study of Cultural Evolution, Stockholm University, SE-106 91 Stockholm, Sweden

Abstract

Here we propose the idea that the success of an argument in favor of an issue position should depend on whether the argument resonates with the audience's values. Now consider two groups, one of which has a broader set of values than the other. We develop a mathematical model to investigate how this difference in broadness of values may drive a change on the population level towards positions in line with the more narrow set of values. The model is motivated by the empirical finding that conservative morality rests equally on moral foundations that are individualizing (harm and fairness) and binding (purity, authority, and ingroup), whereas liberal morality relies mainly on the individualizing moral foundations. The model then predicts that, under certain conditions, the whole population will tend to move towards positions on moral issues (e.g., same-sex marriage) that are supported by individualizing moral foundations.

Keywords:

moral foundations, attitude change, moral opinions, micro-macro, mathematical modelling, ingroup bias

1. Introduction

In a recent paper, Grossmann and Hopkins (2015) noted a paradoxical feature of American public opinion: Whereas conservative responses predominate on items measuring ideological self-identification, liberal positions are more popular than conservative positions on most issues. In the present paper we shall argue that such paradoxes may arise as moral psychology drives cultural dynamics.

A key observation is that positions on issues are indeed dynamic. According to the General Social Survey, a biyearly survey of opinions in the US since 1972, the public opinion in the US has become more liberal on a number of morally charged issues [1]. For instance, from 1972 to 2012 we see substantial, sometimes even dramatic, movement towards more liberal opinions on issues such as whether or not it is morally acceptable to have homosexual relations (support increased from 11% in 1972 to 42% in 2012); sex before marriage (support increased from 27% to 56%); allowing anti-religionists to make speeches (support increased from 66% to 76%); allowing communist books in the library (support increased from 53% to 71%); accepting homosexual college teacher (support increased from 48% to 83%); and approval of sex-education (support increased from 79% in 1974 to 89% in 2012). In contrast to these changes in specific moral opinions, the proportions of Americans who self-identify as liberal have changed only slightly and in the other direction during the same time period, going from 31% in 1974 to 28% in 2012. Thus, it seems that moral opinions have been changing in the liberal

[1] http://www3.norc.org/Gss+website/

direction in a way that cannot be accounted for by spread of liberal ideology.

We shall develop a hypothesis about moral opinon change based on a new perspective on the psychological theory of moral foundations. When applied to the dynamics of American public opinion, our hypothesis offers a possible explanation of the above-mentioned phenomenon.

A basic idea of moral foundations theory (Haidt and Joseph, 2004; Haidt and Graham, 2007; Graham et al., 2009) is that moral opinions draw upon a handful of universal human moral foundations: Harm, Fairness, Ingroup, Authority, and Purity. (The definitions of these moral foundations and much other relevant material, including a discussion of additional candidates for moral foundations, are conveniently collected at a single website, www.moralfoundations.org.) Individuals differ in their reliance on each moral foundation, as measured by the Moral Foundations Questionnaire (Graham et al., 2011). The MFQ asks questions about how relevant the respondent finds various concerns (e.g., whether or not someone violated standards of purity and decency) when making moral judgments. Responses are given on a scale from *This consideration has nothing to do with my judgments of right and wrong* to *This is one of the most important factors when I judge right and wrong.* According to Graham et al. (2009), such reports of moral relevance "are likely to be concordant with explicit reasoning during moral arguments" (p. 1031). Thus, individuals' reliance on moral foundations (as measured by the MFQ) is expected to play a role when they evaluate moral arguments.

The Gallup polls indicate that many individuals change their positions on moral issues over time. Although it is difficult to know the cause of an individual changing position, exposure to arguments is an obvious candidate (Chong, 1996; Chong and Druckman, 2007; Keasey, 1973; Lindström, 1995, 1997). The new perspective we offer is that reliance on moral foundations might influence the individual's receptiveness to various arguments that bears on an issue — thereby explaining why an individual who is exposed to an argument on a moral issue may sometimes change position, sometimes not. To illustrate, consider how positions on the issue of same-sex marriage may be taken by two hypothetical individuals, H and PH, where H finds only the moral foundation of Harm to be relevant whereas PH relies both on Purity and Harm. When exposed to a Harm-based argument why same-sex marriage should be legal, both H and PH should find the argument relevant and may update their positions on the issue accordingly. Change of position in the opposite direction may occur when PH is exposed to a Purity-based argument against same-sex marriage, whereas H is assumed not to be receptive to arguments based on Purity and is therefore less likely to change position based on such arguments. The consequence is that moral arguments can sway PH in both directions but H only in one direction. If conservatives are more likely to be PH types and liberals are more likely to be H types, this could form the basis of an explanation of the phenomenon we described in the opening paragraph.

In the coming sections we explore the idea sketched above. First we develop a hypothesis about an individual mechanism, "position-change bias", grounded in several previous lines of research. We then develop a mathematical model to allow investigation of what macro-level dynamics of moral opinons in the population should emerge from the proposed micro-level mechanism. The last two sections present some testable predictions, both at micro-level and macro-level, and discuss our contribution from a broader perspective. Most of the mathematical analysis appears in the Appendix.

2. A hypothesis about an individual bias in moral position change

The basic message of this paper is that population level change towards more liberal positions on morally loaded issues may be a consequence of conservatives tending to endorse a greater diversity of moral foundations than liberals. The difference between liberals and conservatives was first proposed in a seminal paper by Haidt and Graham (2007):

> Our thesis in this article is that there are five psychological foundations of morality, which we label as harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity. Cultures vary on the degree to which they build virtues on these five foundations. As a first approximation, political liberals value virtues based on the first two foundations, while political conservatives value virtues based on all five. (p. 99)

Haidt and Graham (2007) based this statement on the results of a survey to 1,613 Americans, using a precursor to the Moral Foundations Questionnaire. Researchers from the same group have since replicated the finding using a variety of methods, including the final version of the MFQ (Graham et al., 2009; Koleva et al., 2012). Moral foundations theory is not without its critics; see the 2013 special issue of *Journal of Moral Education* (Maxwell and Narvaez, 2013). Of particular relevance to our thesis, some critics have questioned the fundamental nature of the difference in reliance on moral foundations between liberals and conservatives. For instance, studies have found that the difference in moral foundations endorsement is attenuated under cognitive load or situational threat (Wright and Baril, 2011, 2013). These interesting results give new insight into the nature of moral foundations and suggest that other differences between liberals and conservatives might be more fundamental. However, our premise is only that the difference in reliance on moral foundations exists in practice. This premise is not undermined by the new studies; on the contrary, their control conditions replicate the typical difference between liberals and conservatives.

2.1. The relation between moral foundations and individual change of positions on morally loaded issues

In the introduction we noted that individuals cannot be static in their positions on moral issues. The change in population level support of same-sex marriage seems to be too fast to be accounted for only by population turnover, hence individual change in positions must be common. The fact that people may change positions on moral issues has not received enough attention in moral foundations research (Bloom, 2010). Nonetheless, MFQ items on "the most important factors when I judge right and wrong" suggest a view of moral judgment as an ongoing process. Moral judgment could of course be an ongoing process that consistently results in the same judgment for a given issue. However, earlier psychological research provides good reason to expect individuals to fluctuate in their positions when they encounter moral arguments.

For instance, a classic study of Eiser and White (1974) demonstrated that people's position on an issue may change depending on how the issue is framed. In a more recent study, Brewer (2002) let participants read a newspaper article that framed gay rights either as an issue of morality or an issue of equality. When participants later were asked for the basis of their own position on the issue, they tended to refer to arguments within the frame to which they had been exposed. Taken together, these findings suggest that when an individual encounters a moral argument in which the issue is framed as being about a certain moral foundation, this aspect of the issue may become more important and thereby result in a revised judgment. This view is closely related to political scientist Dennis Chong's model of political attitudes changing with the framing of issues (Chong, 1996; Chong and Druckman, 2007).

Our argument then rests on one fundamental assumption: Framing of an issue as being about, say, purity should be effective only to the extent the individual endorses purity as a valid basis for moral judgment. In other words, we assume that people's self-theories about the bases of their moral judgments have at least some grounding in how they actually form moral judgments. This assumption is consistent with a study of a large sample of US residents who had previously disclosed their political orientation on the liberal-conservative spectrum (Koleva et al., 2012). Respondents filled in both the MFQ and a questionnaire asking for their position on a number of political issues (e.g., same-sex marriage) to which moral arguments are commonly applied. Regression analyses of positions on political issues showed that moral foundations were predictive above and beyond political orientation on the liberal-conservative spectrum. For instance, to favor a ban for same-sex marriage was independently predicted by conservatism, endorsement of the Purity foundation, and lack of endorsement of the Harm foundation. Koleva et al. (2012) noted that their correlational results cannot establish a causal order between foundation endorsement, ideology, and issue positions. We think the most natural interpretation of their results is that people's positions on issues are to some extent influenced directly by their reliance on various moral foundations. Our hypothesis is based on this interpretation.

2.2. The position-change bias hypothesis

Chong (1996) presented a dynamical model of framing in which an individual's receptiveness to various framings of an issue was a key parameter. Building on this idea we shall define an individual's *position-change bias* on a given issue as the ratio between, on the one hand, the likelihood of moral arguments "against" making the individual change position when he (or she) is currently "for" and, on the other hand, the likelihood of arguments "for" making the individual change position if he is currently "against". An individual who is always receptive to arguments for the other position than the one he currently holds has no or weak position-change bias. The position-change bias is strong if one of these change likelihoods is much smaller than the other. Once such a strongly biased individual has arrived at the favored position, he is unlikely to be swayed by arguments for the other position.

Above we have put forward a view of moral foundations as a way to understand change of moral opinion. In this view, MFQ scores on the different moral foundations predict how receptive the individual will be to different kinds of moral arguments. For any given issue, the individual's varying receptivity to different arguments for and against then sums up to his (or her) position-change bias on the issue.

From this hypothesis a prediction follows: Individuals who endorse a wider range of moral foundations should tend to have weaker position-change bias than individuals who endorse a smaller range of moral foundations. From the premise that liberals tend to endorse a smaller range of moral foundations than conservatives we then obtain a second prediction: Conservatives should tend to have weaker position-change bias than liberals. We do not know of any study directly testing this prediction. However, it is clearly consistent with recent findings that conservatives opinions tend to align with a range of both conservative and liberal viewpoints whereas liberals political beliefs show less variation and more consistent support for liberal stances on issues (Kesebir et al., 2013).

Related to our hypothesis, Day et al. (2014) recently suggested that "altering the evoked moral foundations may shape peoples subsequent attitudes, particularly if the moral foundations seem relevant" (p. 2). They also reported data that generally supported this suggestion. However, it should be noted that these data do not speak directly to our hypothesis. For instance, whereas we are interested in the effect of exposure to arguments made by others, participants in these studies came up with arguments themselves. Also, the focus of our hypothesis is moral opinion, whereas Day et al. (2014) studied more general policy issues (immigration, the environment, economic markets, social programs, and education).

3. A mathematical model of population-level change of moral opinions based on individual-level biases

An intriguing aspect of the position-change bias hypothesis is that it offers a potential explanation for the pattern described in the beginning. When exposed to a counter-argument to one's current position, the likelihood that a conservative will abandon a position typically favored by conservatives is expected to be greater than the likelihood that a liberal will abandon a position typically favored by liberals. It seems intuitive that, in the long term, this asymmetry in position-change bias could lead to a shift in the population towards the position typically favored by liberals.

However, position-change bias is not the only bias that will affect the cultural dynamics of moral opinions. A large body of social psychological research suggests a permeative presence of ingroup bias. We should expect conservatives to be much more likely to *be exposed to and care about* a moral argument when it is made by a fellow conservative than when made by a liberal, and vice versa. In order to explore the potency of position-change bias in the presence of group-exposure bias we developed a mathematical model of the cultural dynamics of moral opinions.

Mathematical models enable researchers to explore the consequences of various possible assumptions. Such models have been used for decades in mathematical sociology and cultural evolution research (Boyd and Richerson, 1988; Coleman, 1994; Schelling, 2006). Chong (1996) used a dynamic model based on assumptions partly related to ours to discuss a number of issues in political attitude change, including the development of consensus or polarization on an issue. We know of no previous models aimed specifically at exploring how empirical findings on individual moral cognition should influence our understanding of the dynamics of moral opinions.

3.1. Model assumptions

For maximal simplicity we shall assume an infinite population in which every individual is either Lib or Con, a simple dichotomy. Now consider an issue on which people hold either of two *positions*, either FOR or AGAINST. The FOR position is better aligned with the Lib profile of moral foundations, whereas the AGAINST position is better aligned with the Con profile. However, these differences in moral foundation profiles do not strictly determine the individual's position on the issue. An individual's position may change if the individual is exposed to an argument for the other position.

Three parameters govern the process by which positions are acquired. The first parameter is the *group-exposure bias*. This bias measures how much less likely it is for an individual to be exposed to (and care about) the arguments of someone from the same political orientation than of someone from the other political orientation (i.e., a Lib is more likely to listen to another Lib, etc.). The group-exposure bias for a given set of values will be defined as the ratio between two probabilities as follows:

$$G_{\text{l}} = \frac{\text{Prob(a Lib is exposed to a Con's argument)}}{\text{Prob(a Lib is exposed to a Lib's argument)}}$$

and, analogously,

$$G_{\text{c}} = \frac{\text{Prob(a Con is exposed to a Lib's argument)}}{\text{Prob(a Con is exposed to a Con's argument)}}.$$

Note that these probability ratios incorporate the influence of all factors on exposure. Thus, the group-exposure bias will reflect the total influence of psychological factors (such as preferences for the ingroup) and structural factors (such as the tendency for similar people to cluster together, the proportions of Libs and Cons in the population, and any differences between the groups in their power and efforts to expose others to their arguments). With all exposure probabilities assumed to be non-zero and a Lib assumed to be at least as likely to be exposed to a Lib argument than to a Con argument, etc., the group-exposure bias parameters are assumed to satisfy $0 < G_{\mathrm{l}} \leq 1$ and $0 < G_{\mathrm{c}} \leq 1$.

Our assumption is that the source of bias in exposure is group membership, not position. In other words, we assume the conditional probability of being exposed to an argument FOR, given that the individual making the argument is a Con, to be just the current proportion of FOR within the Con group (and similarly for alternative possibilities).

The second parameter is the *position-change bias*. This bias measures how the difference in moral foundation profiles between the political orientations makes it less likely for an individual to be swayed by arguments for one position than for the other position (i.e., a Lib is less easily swayed to the AGAINST position than to the FOR position, etc.). The position-change bias for a given set of values is defined as the ratio between two probabilities as follows:

$$P_{\mathrm{l}} = \frac{\text{Prob(a Lib who is FOR is swayed when exposed to an arg. AGAINST)}}{\text{Prob(a Lib who is AGAINST is swayed when exposed to an arg. FOR)}}$$

and, analogously,

$$P_{\mathrm{c}} = \frac{\text{Prob(a Con who is AGAINST is swayed when exposed to an arg. FOR)}}{\text{Prob(a Con who is FOR is swayed when exposed to an arg. AGAINST)}}.$$

Assuming all swaying probabilities to be non-zero and a Lib more likely to be swayed to FOR than AGAINST, etc., the position-change bias parameters must satisfy $0 < P_\text{l} \leq 1$ and $0 < P_\text{c} \leq 1$.

The third parameter is the *influentiability coefficient*. This is a measure of how often others are allowed to influence an individual's position and incorporates both absolute levels of exposure per time step and absolute levels of influence. Previous research has indicated that conservatives tend to exhibit less openness than liberals (Jost et al., 2003), which suggests that influentiability differs between political orientations. We therefore introduce two separate influentiability coefficients, $I_\text{l} > 0$ and $I_\text{c} > 0$.

3.2. The dynamical system

Denote the proportions of FOR and AGAINST in the Lib population at a certain time by q_l and $(1 - q_\text{l})$, respectively. The corresponding proportions in the Con population are q_c and $(1 - q_\text{c})$.

Change comes from individuals being exposed to and swayed by arguments for the other position. Let $\Delta q_\text{l}/\Delta t$ denote the change over a small time step Δt in the proportion of FOR in the Lib population. Four types of events contribute to change:

- A Lib who is currently AGAINST may be swayed from exposure to a Lib who is FOR. This event happens with rate $(1 - q_\text{l})q_\text{l}I_\text{l}$.
- A Lib who is currently AGAINST may be swayed from exposure to a Con who is FOR. Because of group-exposure bias, this happens only with a rate of $(1 - q_\text{l})q_\text{c}G_\text{l}I_\text{l}$.
- A Lib who is currently FOR may be swayed from exposure to a Lib who

is AGAINST. Because of position-change bias, this happens only with a rate of $q_l(1-q_l)P_lI_l$.

- A Lib who is currently FOR may be swayed from exposure to a Con who is AGAINST. Because of combination of group-exposure bias and position-change bias, this happens only with rate $q_l(1-q_c)G_lP_lI_l$.

Under the assumption that the population is infinite there will be no stochastic effects and we can just sum the above rates of change, with the appropriate signs, to obtain the following formula for change over a small time step Δt:

$$\frac{\Delta q_l}{\Delta t} = \left[+(1-q_l)(q_l+q_cG_l) - q_lP_l((1-q_l)+(1-q_c)G_l)\right] I_l \qquad (1)$$

In this equation, note how the three parameters occur in the right-hand side expression. First, the influentiability coefficient I_l occurs as a factor of the entire change expression. As we discuss in the Appendix, this means that the influentiability coefficient will influence the speed of change, but it will not influence the long-term outcome. Second, the position-change bias P_l multiplicatively decreases all swaying of Libs in the AGAINST direction. The swaying of Libs in the FOR direction is decreased by the group-exposure bias G_l, but *not* multiplicatively; group-exposure bias decreases only the swaying of Libs by Cons (and regardless of direction), whereas the swaying of Libs by Libs is unaffected. As we shall see in the Appendix, it is therefore difficult, and often impossible, for a change in the latter bias to compensate for a change in the former bias.

By analyzing the analogous four events for swaying of Cons we obtain a similar equation for the dynamics of q_c:

$$\frac{\Delta q_c}{\Delta t} = \left[-q_c((1-q_c)+(1-q_l)G_c) + (1-q_c)P_c(q_c+q_lG_c)\right] I_c \qquad (2)$$

Note that the parameters occur in this equation in the same manner as in the previous one. Thus, the influentiability coefficient I_c occurs as a factor of the entire change expression; the position-change bias P_c multiplicatively decreases all swaying of Cons in the FOR direction; the swaying of Cons in the AGAINST direction is decreased by the group-exposure bias G_c, but not multiplicatively.

3.3. Results

A mathematical analysis of this dynamical system is given in the Appendix. Here we present the main results in an accessible way. Recall that the model assumes that conservatives and liberals are biased toward different positions on an issue. The model predicts that:

1. if a position is present at all in the population it will be present in both groups but in different proportions;
2. the long-term proportions do not depend on the initial proportions, nor on the influentiability coefficients;
3. the long-term proportions are determined by the strength of the group-exposure bias and the position-change bias such that the position favored by the more biased group will tend to become the majority position in the population;
4. one group's biases influence the long-term proportions of positions in both groups;
5. position-change bias plays a greater role than group-exposure bias (i.e., it is difficult to compensate for a difference in position-change bias by a difference in group-exposure bias).

To illustrate these analytic results we show the outcomes of a series of computer simulations of the model. Each simulation tracks the change of the proportion of FOR among liberals and conservatives over 50 time steps. Our reference case will be Simulation A, in which there is no group-exposure bias ($G_{\mathrm{l}} = G_{\mathrm{c}} = 1$), both groups exhibit equally strong position-change bias ($P_{\mathrm{l}} = P_{\mathrm{c}} = 0.5$), and are equally influentiable ($I_{\mathrm{l}} = I_{\mathrm{c}} = 0.5$). Starting at low proportions of FOR, $q_{\mathrm{l}} = 0.3$ and $q_{\mathrm{c}} = 0.1$, we see in Figure 1 that proportions of FOR increase over time towards equilibrium levels at $q_{\mathrm{l}} = 2/3$ and $q_{\mathrm{c}} = 1/3$ (as predicted by plugging these parameter values into the formula (A.8) in the Appendix).

Figure 1 also illustrates that the same equilibrium is approached regardless if the start values are radically different: In Simulation B the FOR position is initially in majority in both groups, yet ends up approaching the same equilibrium levels ($q_{\mathrm{l}} = 2/3$ and $q_{\mathrm{c}} = 1/3$). In the same vein, Figure 2 illustrates that the same equilibrium is approached, only at a slower speed, if an influentiability coefficient is set at a lower value (Simulation C).

Now consider the effect of group-exposure bias. Figure 3 compares the reference case with Simulation D, in which the conservative group-exposure bias is stronger than in Simulation A (G_{c} is set to 0.5 instead of 1). Over time this conservative group-exposure bias results in a lower proportion of the FOR position not only among conservatives but also among liberals. For the population as a whole, this means a substantially decreased support for the FOR position due to conservative group-exposure bias.

Next we turn to position-change bias, the key concept of this paper. Figure 4 compares the reference case with Simulation E, in which the liberal position-change bias is stronger than in Simulation A (P_1 is set to 0.25 instead of 0.5). Over time this liberal position-change bias results in a greater proportion of the FOR position not only among liberals but also among conservatives. For the population as a whole, this means a substantially increased support for the FOR position due to liberal position-change bias.

Finally, consider the interaction of the two types of bias. Figure 5 compares the reference case with Simulation F, combining the conservative position-change bias and the liberal position-change bias of the two previous simulations. Over time this combination of biases results in a greater proportion of the FOR position among liberals but no change among conservatives. For the population as a whole, this means increased support for the FOR position. In other words, position-change bias played a greater role than group-exposure bias for the population as a whole.

3.4. Discussion of the model assumptions

Our simple model could be extended and refined in various ways to make it more realistic. Here we point out five assumptions that could be relaxed and discuss what would be the likely impact on results.

First, predictions of long-term behavior are based on the assumption that parameter values are constant over time. However, the model itself is based on a rule that updates proportions of FOR and AGAINST in each time-step. Thus, the model allows simulations of fluctuating parameter values. Such fluctuations will lead to fluctuating proportions of FOR and AGAINST. Long-term average proportions over time should still be predictable by long-term average parameter values.

Second, our model made the unrealistic assumption of an infinite population. The point of this assumption was to let us ignore stochastic effects and obtain a deterministic rule for change in the population in each time step. A finite population model must instead keep track of how each individual is subject to a sequence of random events in which the individual with some probability is exposed to another's argument and, if so, with some probability is swayed. The expected population change in one time step is the same as in the infinite population model. By chance, the change may become smaller or greater than expected. Because the real population (e.g., in America) is very large, the law of large numbers implies that such stochastic effects would typically entail only the addition of a minimal amount of noise to the prediction of the infinite population model. Thus, we conclude that taking finiteness of the population into account is very unlikely to have a large impact on the results.

Third, our model allows no other differences between individuals than those connected with the division in liberals and conservatives. In other words, all individuals within a group are assumed to be identical in their parameter values. We can think of our model as replacing all individuals by the group average. A more refined model would, around these group averages, incorporate random within-group variation of characteristics between individuals. Because our results only deal with the aggregate level (i.e., the proportion of the population that changes in a time step), it is likely that they mainly depend on the aggregated level of characteristics (i.e., group averages). Thus, we conclude that taking within-group variation in parameter values account is unlikely to have a large impact on the results.

Fourth, our model assumes that exposure within each group reflects the current proportions of FOR and AGAINST in the group. This assumption would automatically hold if every individual is equally likely to be a source of social influence. However, it will not necessarily hold in actual networks of social influence, in which some individuals wield much more influence than others do due to factors like status, connections, ability and interest. A more refined model would incorporate an influence-weight for each individual, governing the probability of others to be exposed to that individual's argument. Such a refined model is likely to behave approximately as the original model with parameters set to the influence-weighted average parameter values. This would have a qualitative impact on results only if highly influential liberals are more like conservatives and vice versa.

Fifth, our model assumes that people change their position on moral issues from exposure to moral arguments. In reality there might be other important mechanisms of change. Specifically, people might adopt positions without ever considering how they fit with their endorsed moral foundations (e.g., due to automatic conformity). Any sufficiently well-defined proposal for such mechanisms could be included in a model. However, for such additional mechanisms to have a qualitative impact on our results it seems that they would have to be stronger for one group than for the other. We know of no a priori reason to expect morally unmotivated change of positions on moral issues to vary with political orientation. Thus, as long as the basic mechanism underlying our model is correct, we expect the results of our model to make qualitatively correct predictions in realistic settings.

4. Predictions

Building on prior theoretical work on the importance of individual differences in receptivity to various framings of an issue (Chong, 1996), we have here defined the concept of individual differences in "position-change bias." Building on moral foundations theory (Haidt and Graham, 2007; Graham et al., 2009) we then developed a hypothesis about how position-change bias should relate to moral foundation endorsement. This hypothesis makes the following prediction, which should be testable in longitudinal studies or experiments on change of moral positions (see Keasey (1973) for an example).

Prediction 1: Consider any particular moral issue for which different kinds of moral foundations tend to support arguments for different positions on the issue. Individuals who equally strongly endorse moral foundations of both kinds should tend to exhibit less position-change bias than individuals who strongly endorse only one kind of moral foundations.

We also discussed the empirical finding that conservatives tend to give more equal endorsement of different moral foundations than liberals (Haidt and Graham, 2007; Graham et al., 2009; Koleva et al., 2012). This empirical finding matches the premise of Prediction 1, thus yielding a second interesting prediction.

Prediction 2: On moral issues for which different kinds of moral foundations typically support arguments for different positions, conservatives should tend to exhibit less position-change bias than liberals.

Because position-change bias measures how likely an individual is to change position from for to against compared to the opposite direction, the aggregate effect of the second prediction over time should be long-term movement of moral opinons in the direction given by the moral foundations (harm and fairness) favored by liberals. We used a mathematical model to investigate how sensitive this conclusion are to the presence of group-exposure bias. The results can be summarized as a third prediction.

Prediction 3: Unless conservative group-exposure bias is much stronger than liberal group-exposure bias, stronger position-change bias among liberals should lead to a tendency for liberal positions to become majority positions over time.

It seems unlikely that group-exposure bias in the psychological sense should differ much in strength between liberals and conservatives. However, group-exposure bias in the model is defined simply as the degree to which members of one group are more likely to be exposed to arguments from the own group than to arguments from the other group. This means that groups could differ in their group-exposure bias for purely institutional reasons. A conservative media monopoly and organized sanctions against the expression of liberal opinions should lead to generally higher exposure to conservative arguments than to liberal arguments. In our model this would be represented as stronger conservative group-exposure bias and weaker liberal group-exposure bias.

The suggested pathway to more liberal moral opinions should therefore apply only under conditions of media pluralism and free speech. These conditions are, on the whole, satisified in the United States. The pathway proposed in this paper may therefore be (at least part of) the explanation for the American trend noted in the beginning: moral opinons seem to become more liberal without a corresponding liberalization of values. Many other societies lack media pluralism and free speech. Variation in these societal feature has been studied in several academic disciplines. For instance, economists and political scientists are interested in the relation between press freedom and corruption. Some propose a causal effect such that freer press leads to lower corruption (Freille et al., 2007). Psychologists have found press freedom to covary with self-expression values and individualism (e.g., Van de Vliert (2011)). In the same vein as these findings, our theoretical argument yields a prediction about societal differences.

Prediction 4: Societal trends towards more liberal opinons on moral issues should be found mainly in societies with media pluralism and free speech.

This prediction should be of broad interest to psychologists, sociologists and political scientists.

5. Conclusion

Nobel laureate Thomas Schelling wrote a book based on the principle that macrobehavior can be derived from micromotives (Schelling, 2006). Within the realm of economic behavior this principle has been extensively explored in economic models based on assumptions of profit-maximizing actors. Outside the economic realm the micromotive of profit maximization is less generally applicable, but various preferences and biases may instead apply on a case-by-case basis. For instance, one theme in the abovementioned book by Schelling is how individual preferences for being with similar others could lead to the emergence of macroscale segregation. In the present paper we have similarly argued that the macro-dynamics of moral opinons might be derived from individual biases in receptivity to moral arguments.

We have argued elsewhere that a fundamental aspect of cultural change is that individuals change their cultural traits over life; this feature distinguishes cultural evolution from genetic evolution, genes being approximately constant over each individual's lifetime (Strimling et al., 2009). Specifically, a model of the cultural dynamics of moral opinons should be informed by knowledge about biases in the process whereby individuals change their moral judgments. Piecing together findings from moral foundations research with findings from research on attitude change, we formulated a hypothesis about a mechanism at the individual level. The corresponding Predictions 1 and 2 about micromotives could be tested in future psychological research. Assuming the validity of the individual level hypothesis we then modeled the social dynamics that should emerge. The corresponding Predictions 3 and 4 about macrobehavior could be tested using data and methods from sociology and political science.

The success of the micro-to-macro approach depends on research at the different levels fitting together. If we may offer a general conclusion, it would be that psychological research on preferences and biases — anything that could serve as "micromotives" — could become more useful for the study of dynamics of macrobehavior by focusing more on how individuals change.

6. Acknowledgement

The preparation of this article was supported by grants from the Swedish Research Council (2009-2390 and 2009-2678) to the authors. We are grateful to Brent Simpson for inspiring this work.

Appendix A. Analysis of equilibria of the dynamical system

Equilibria (i.e., fixed points) of the dynamical system are obtained when there is zero change, that is, when the formulas in (1) and (2) equal zero. The equilibrium equations can be written

$$(1 - q_l)(q_l + q_c G_l) = q_l P_l \left((1 - q_l) + (1 - q_c)G_l\right) \tag{A.1}$$

and

$$q_c \left((1 - q_c) + (1 - q_l)G_c\right) = (1 - q_c)P_c(q_c + q_l G_c). \tag{A.2}$$

Note that the equilibrium equations do not depend on the influentiability coefficients I_l and I_c. Thus, although the speed of the dynamical process is influenced by these coefficients they do not influence what are the equilibrium outcomes.

Appendix A.1. Pure equilibria

It is evident from the equilibrium equations that there are always two *pure equilibria*: (A.1) and (A.2) are satisfied both by $q_l = q_c = 1$ and $q_l = q_c = 0$. These solutions correspond to the entire population being FOR and AGAINST, respectively.

Appendix A.2. Mixed equilibria

Next we investigate the possibility of *mixed equilibria*. First, it is straightforward to see that the equilibrium equations (A.1) and (A.2) can never be satisfied when only one group is mixed, so a mixed equilibrium must have both $0 < q_l < 1$ and $0 < q_c < 1$. For any such mixed state we can define ratios of proportions:

$$\gamma := \frac{q_c}{q_l} > 0 \text{ and } \beta := \frac{1 - q_l}{1 - q_c} > 0.$$

Note that from these ratios the proportions can be retrieved: $q_l = (1 - \beta)/(1 - \gamma\beta)$ and $q_c = \gamma q_l$.

Appendix A.3. Trivial mixed equilibria

First consider the trivial case where nobody has any position-change bias (i.e., $P_l = P_c = 1$). Both equilibrium equations then reduce to $q_l = q_c$, that is, any proportion of FOR is an equilibrium as long as it is the same in both groups. All remaining ("nontrivial") mixed equilibria must satisfy both $P_l < 1$ and $P_c < 1$.

Appendix A.4. Nontrivial mixed equilibria

In order to look for parameters allowing non-trivial mixed equilibria, we will henceforth assume that $0 < q_l, q_c, P_l, P_c < 1$. If we divide the equilibrium equations by $q_l(1 - q_c)$, they can be expressed in terms of γ and β as

$$\beta(1 + \gamma G_l) = P_l(\beta + G_l) \tag{A.3}$$

and

$$\gamma(1 + \beta G_c) = P_c(\gamma + G_c). \tag{A.4}$$

Solving for β in (A.3) and substituting into (A.4) we obtain a quadratic equation in γ. On standard form:

$$\gamma^2 + \frac{\gamma[(1 - P_c)(1 - P_l) - G_l G_c(P_c - P_l)] - P_c G_c(1 - P_l)}{G_l(1 - P_c)} = 0 \tag{A.5}$$

Note that the constant term is negative, so there will be only one positive solution. Letting

$$R := [(1 - P_c)(1 - P_l) - G_l G_c(P_c - P_l)]^2 + 4 G_l G_c P_c (1 - P_l)(1 - P_c),$$

the non-trivial mixed equilibrium solution to (A.5) can be expressed as

$$\hat{\gamma} = \frac{-[(1-P_c)(1-P_l) - G_l G_c (P_c - P_l)] + \sqrt{R}}{2G_l(1-P_c)} > 0. \tag{A.6}$$

In equations (A.3) and (A.4) the roles of β and γ are symmetric with respect to swapping Lib and Con(whereas R is invariant under this swap). Hence, by swapping Lib and Con in (A.6) we obtain the non-trivial mixed equilibrium value for β:

$$\hat{\beta} = \frac{-[(1-P_c)(1-P_l) + G_l G_c (P_c - P_l)] + \sqrt{R}}{2G_c(1-P_l)} > 0. \tag{A.7}$$

Recall the identities $q_l = (1-\beta)/(1-\gamma\beta)$ and $q_c = \gamma q_l$. The values of q_l and q_c in a non-trivial mixed equilibrium can therefore be expressed as

$$\hat{q}_l = (1-\hat{\beta})/(1-\hat{\gamma}\hat{\beta}) \text{ and } \hat{q}_c = \hat{\gamma}\hat{q}_l, \tag{A.8}$$

where $\hat{\gamma}$ and $\hat{\beta}$ are given by (A.6) and (A.7). Of course, the non-trivial mixed equilibrium exists only if the equilibrium proportions satisfy $0 < \hat{q}_l < 1$ and $0 < \hat{q}_c < 1$. From (A.8) it follows that these conditions are satisifed if and only if $\hat{\gamma} < 1$ and $\hat{\beta} < 1$. Using (A.6) and the assumptions that all parameters lie between 0 and 1, the inequality $\hat{\gamma} < 1$ straightforwardly simplifies to

$$(1-P_c)(1-P_l) + G_l G_c (P_l - P_c) + G_l(1-P_c) - P_c G_c (1-P_l) > 0. \tag{A.9}$$

Symmetrically, $\hat{\beta} < 1$ implies that

$$(1-P_c)(1-P_l) - G_l G_c (P_l - P_c) + G_c(1-P_l) - P_l G_l (1-P_c) > 0. \tag{A.10}$$

Appendix A.5. Importance of parameter values in determining equilibrium proportions of FOR *and* AGAINST

Figure A.6 illustrates how parameter space is divided into three sectors by the inequalities (A.9 and A.10) for the existence of a mixed equilibrium.

The figure is drawn for $G_\mathrm{l} = G_\mathrm{c} = 1/2$ but looks qualitatively similar for other values of the group-exposure bias parameters. When P_l is much lower than 1 (i.e., liberals are strongly position-change biased) but P_c is close to 1 (i.e., conservatives are not very position-change biased), a pure equlibrium where everyone is FOR is obtained. Symmetrically, when P_c is much lower than 1 but P_l is close to 1, a pure equlibrium where everyone is AGAINST is obtained.

The mixed equilibrium exists in the intermediate sector where both inequalities (A.9) and (A.10) are satisfied. In this sector the pure equilbria are unstable. It is well-known that an equilibrium of a dynamical system is unstable if the Jacobian of the dynamical system, evaluated at the equilibrium, has an eigenvalue with absolute value greater than one. It is straightforward to calculate the Jacobian, evaluate it at either of the two pure equilibria, and verify that the inequalities (A.9) and (A.10), respectively, imply that the largest eigenvalue is greater than 1. We omit the details.

For a comparison of the effects of position bias and group-exposure bias we refer to Figure A.7. This figure shows, for various combinations of P_l and G_l, whether FOR or AGAINST is in majority in the mixed equilibrium (under the simplifying assumption that the liberal and conservative subpopulations are of equal size). The axes are log scaled, that is, a constant distance along an axis corresponds to multiplication of the bias with a constant factor. A log scale is the correct scale for comparisons as biases are multiplicative. It is clear from Figure A.7 that the majority position is mainly determined by the position-change bias P_l, whereas the value of G_l has much less influence. In other words, a change in P_l can only be compensated for (if at all) by a much larger change in G_l. Figure A.7 is drawn for $P_c = G_c = 1/2$ but looks qualitatively similar for other values of these bias parameters. The intuitive explanation was mentioned in the main text: The position-change bias decreases all swaying in the direction that is atypical for the group, whereas the group-exposure bias applies only to a subset of the swaying in both directions.

Bloom, P., 2010. How do morals change? Nature 464 (7288), 490–490.

Boyd, R., Richerson, P. J., 1988. Culture and the evolutionary process. University of Chicago Press.

Brewer, P. R., 2002. Framing, value words, and citizens' explanations of their issue opinions. Political Communication 19 (3), 303–316.

Chong, D., 1996. Creating common frames of reference on political issues. In: Mutz, D. C., Sniderman, P. M., Brody, R. A. (Eds.), Political Persuasion and Attitude Change. Ann Arbor: University of Michigan Press, pp. 195–224.

Chong, D., Druckman, J. N., 2007. Framing theory. Annual Review of Political Science 10, 103–126.

Coleman, J. S., 1994. Foundations of social theory. Harvard University Press.

Day, M. V., Fiske, S. T., Downing, E. L., Trail, T. E., 2014. Shifting liberal and conservative attitudes using moral foundations theory. Personality and Social Psychology Bulletin 40 (12), 1559–1573.

Eiser, J. R., White, C. J. M., 1974. The persuasiveness of labels: Attitude change produced through definition of the attitude continuum. European Journal of Social Psychology 4 (1), 89–92.

Freille, S., Haque, M. E., Kneller, R., 2007. A contribution to the empirics of press freedom and corruption. European Journal of Political Economy 23 (4), 838–862.

Graham, J., Haidt, J., Nosek, B. A., 2009. Liberals and conservatives rely on different sets of moral foundations. Journal of Personality and Social Psychology 96 (5), 1029–1046.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., Ditto, P. H., 2011. Mapping the moral domain. Journal of Personality and Social Psychology 101 (2), 366–385.

Grossmann, M., Hopkins, D. A., 2015. Ideological republicans and group interest democrats: The asymmetry of american party politics. Perspectives on Politics 13 (01), 119–139.

Haidt, J., Graham, J., 2007. When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. Social Justice Research 20 (1), 98–116.

Haidt, J., Joseph, C., 2004. Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. Daedalus 133 (4), 55–66.

Jost, J. T., Glaser, J., Kruglanski, A. W., Sulloway, F. J., 2003. Political conservatism as motivated social cognition. Psychological Bulletin 129 (3), 339–375.

Keasey, C. B., 1973. Experimentally induced changes in moral opinions and reasoning. Journal of Personality and Social Psychology 26 (1), 30–38.

Kesebir, P., Phillips, E., Anson, J., Pyszczynski, T., Motyl, M., 2013. Ideological consistency across the political spectrum: Liberals are more consistent but conservatives become more consistent when coping with existential threat. Available at SSRN 2215306.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., Haidt, J., 2012. Tracing the threads: How five moral concerns (especially purity) help explain culture war attitudes. Journal of Research in Personality 46 (2), 184–194.

Lindström, P., 1995. Falsifiability of political opinions. Scandinavian Journal of Psychology 36 (3), 295–305.

Lindström, P., 1997. Retrospective versus prospective explanations of changes in political opinion. Scandinavian Journal of Psychology 38 (2), 157–163.

Maxwell, B., Narvaez, D., 2013. Moral foundations theory and moral development and education. Journal of Moral Education 42 (3), 271–280.

Schelling, T. C., 2006. Micromotives and macrobehavior. WW Norton & Company.

Strimling, P., Enquist, M., Eriksson, K., 2009. Repeated learning makes cultural evolution unique. Proceedings of the National Academy of Sciences 106 (33), 13870–13874.

Van de Vliert, E., 2011. Bullying the media: Cultural and climato-economic readings of press repression versus press freedom. Applied Psychology 60 (3), 354–376.

Wright, J. C., Baril, G., 2011. The role of cognitive resources in determining our moral intuitions: Are we all liberals at heart? Journal of Experimental Social Psychology 47 (5), 1007–1012.

Wright, J. C., Baril, G. L., 2013. Understanding the role of dispositional and situational threat sensitivity in our moral judgments. Journal of Moral Education 42 (3), 383–397.

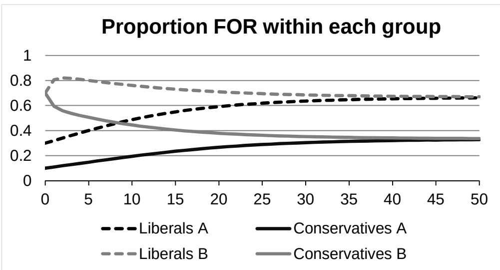

Figure 1. Simulations A and B have the same parameter values ($G_\mathrm{l} = G_\mathrm{c} = 1$, $P_\mathrm{l} = P_\mathrm{c} = 0.5$, $I_\mathrm{l} = I_\mathrm{c} = 0.5$) but different start values ($\mathrm{A} : q_\mathrm{l} = 0.3, q_\mathrm{c} = 0.1$ vs. $\mathrm{B} : q_\mathrm{l} = 0.7, q_\mathrm{c} = 0.7$.

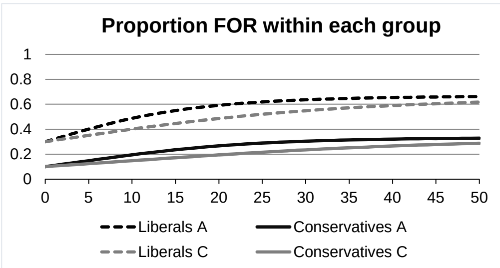

Figure 2. Simulations A and C have the same parameter values except the *liberal influentiability coefficient* is only half as large in C ($I_l = 0.25, I_c = 0.5$) as in A.

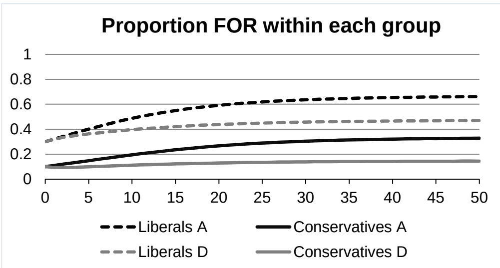

Figure 3. Simulations A and D have the same parameter values except the *conservative group-exposure bias* is twice as strong in D ($G_{\mathrm{l}} = 1, G_{\mathrm{c}} = 0.5$) as in A.

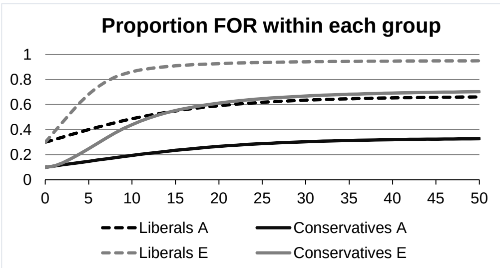

Figure 4. Simulations A and E have the same parameter values except the *liberal position-change bias* is twice as strong in E ($P_l = 0.25, P_c = 0.5$) as in A.

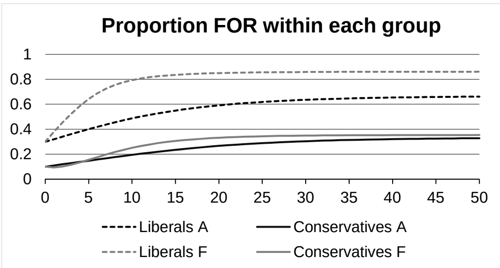

Figure 5. Simulations A and F have the same parameter values except both the *conservative group-exposure bias* and the *liberal position-change bias* are twice as strong in F ($G_{\text{l}} = 1, G_{\text{c}} = 0.5,\ P_{\text{l}} = 0.25, P_{\text{c}} = 0.5$) as in A.

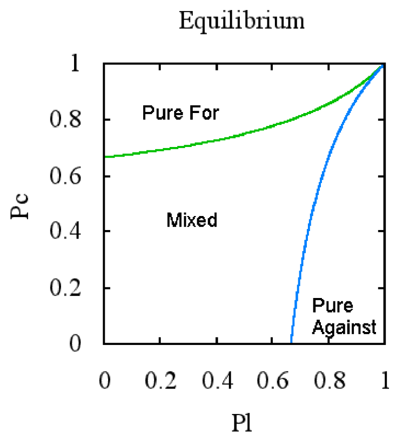

Figure 6. For fixed $G_{\mathrm{l}} = G_{\mathrm{c}} = 1/2$, the figure shows the combinations of values of the parameters P_{l} and P_{c} for which the equilibrium is pure or mixed.

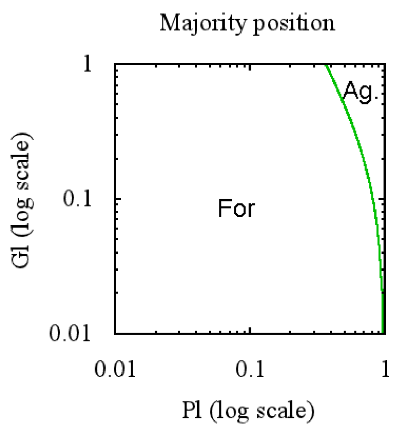

Figure 7. For fixed $P_{\mathrm{c}} = G_{\mathrm{c}} = 1/2$, the figure shows the combinations of values of the parameters P_{l} and G_{l} for which $(\hat{q}_{\mathrm{l}} + \hat{q}_{\mathrm{c}})/2 > 1/2$ (i.e., the majority is FOR in the mixed equilibrium) or $(\hat{q}_{\mathrm{l}} + \hat{q}_{\mathrm{c}})/2 < 1/2$ (i.e., the majority is AGAINST in th emixed equilibrium).