



Co-developing beliefs and social influence networks—towards understanding socio-cognitive processes like Brexit

Bruce Edmonds¹ 

© The Author(s) 2019

Abstract

A model of mutual influence is presented where the structure of individual's beliefs and the social structure both matter. The model thus combines processes of belief change base on Thagard's (Behav Brain Sci 12:435–467, 1989) theory of mental coherence with plausible processes of social network change. This combination of cognitive and social processes has outcomes that are qualitatively different from either only cognitive or only social processes, which shows the importance of studying these together. An illustration that moves towards representing the processes involved in Brexit is also exhibited to show the potential of this kind of simulation. Whilst only conceived of as an illustration of a kind of model, it is consistent with a number of observed patterns in opinion poll data, with some social and cognitive theories and only consists of plausible processes. This kind of model could also be used to relate and integrate different kinds of evidence into a coherent framework in the shape of more developed simulations.

Keywords Socio-cognitive system · Agent-based simulation · Politics · Social influence · Explanatory coherence · Thagard · Brexit · Opinion polls · Social intelligence · Opinions · Beliefs

1 Introduction

Many in the UK were surprised at the outcome of the Brexit vote. Thus, it is natural to want to understand the kind of processes that led to this, and similar processes of collective opinion change and formation. This paper describes some steps in that direction in the form of a model of mutual influence, where the structure of the agent beliefs and the social structure co-evolve. Such a model combines the cognitive with the social and requires a cross-validation methodology. It is consistent with a call to apply complexity science to the study of ideology (Homer-Dixon et al. 2013).

✉ Bruce Edmonds
bruce@edmonds.name

¹ Centre for Policy Modelling, Manchester Metropolitan University, All Saints Campus, Oxford Road, Manchester M15 6BH, UK

2 Motivation: some target macro patterns

In this section, we look at some of the patterns coming out of Brexit. The data comprises of all the public opinion polls aimed to be representative of the UK between 2011 and the referendum in 2016.¹ Firstly, some of the clues about the kind of processes that might be involved. Figure 1 (left) shows the raw %Remain—%Leave difference in polls leading up to the final vote. However, this graph is somewhat difficult to read. Figure 1 (right) shows the same data, but (a) adjusted for known biases between telephone and online polls² and (b) exponentially smoothed using a factor of 25%. These show definite medium-term trends, with periods of relative stability, but also what seem to be “turning points”. There is also a lot of smaller, short-term variation that maybe be measurement ‘noise’.

One of the factors that graphs, such as the above, leave out is that a significant proportion of those polled are undecided. These people are important, because votes are often won by persuading the supporters of the other side to stay at home and not vote, and by persuading the undecided to vote for yourself. It is rare for someone to change sides from definitely one side to the other (especially on such a divisive issue as Brexit).

Figure 2 (left) shows each poll plotted in terms of the proportion for leaving against the proportion undecided. These differ greatly in terms of the proportion undecided because each poll is designed very differently in terms of how easy it is to choose undecided as an option. Some polls strongly encourage users to make a decision (e.g. by not offering it as an option). Figure 2 (right) is the equivalent graph but for remaining against undecided. In the actual Brexit poll there were a relatively large number of people that did not vote—not all of these would be due to being undecided or conflicted, but at least some for this reason. One can see that the final result is consistent with the polls if one takes the undecided into account (and interprets not voting as being undecided). If there were no undecided in the referendum, this data suggests that the result might have been different, but not voting is as much a choice as any other. Anecdotal evidence suggests that whilst some voters wavered between remain and undecided or between leave and undecided, there were very few who swapped between remain and leave.

Turning to the structure of social influence, there is no doubt that the social network along which influence can occur is clustered into those with similar beliefs. However, direct evidence for the structure of this is hard to come by and we have to be satisfied with indirect indications. Krasodomski-Jones of Demos (2016) selected 2500 at random from a larger population of political twitter accounts, divided equally into Labour, SNP, Tory and UKIP supporters (as well as a control group that I do not include here). It then analysed over a million tweets from these users from May–August 2016. One of these analyses was to see who re-tweeted tweets from whom. Table 1 shows a summary of this analysed by the party they support. They then visualised the re-tweet network in a similar manner to Adamic and Glance (2005) (Fig. 3).

The picture that emerges is that these actors are very much sorted by (a) their own party and (b) on a roughly linear political spectrum. That is most re-tweets were within their own party. There was some re-tweeting between: UKIP & Tory, Tory & Labour, Labour & SNP, but very low levels of re-tweeting between SNP and either UKIP or Tory or between UKIP and either SNP or Labour. This is only a subset of the links of

¹ As collected by the financial times at: <https://ig.ft.com/sites/brexit-polling/>. Accessed: 2019-04-17. (Archived by WebCite® at <http://www.webcitation.org/77hwRbe3Y>)—a copy of the data in CSV format is available at: <http://cfpm.org/data/> in the file called “list of individual EU referendum polls.csv”.

² Internet polls consistently showed more support for leaving the EU than telephone polls.

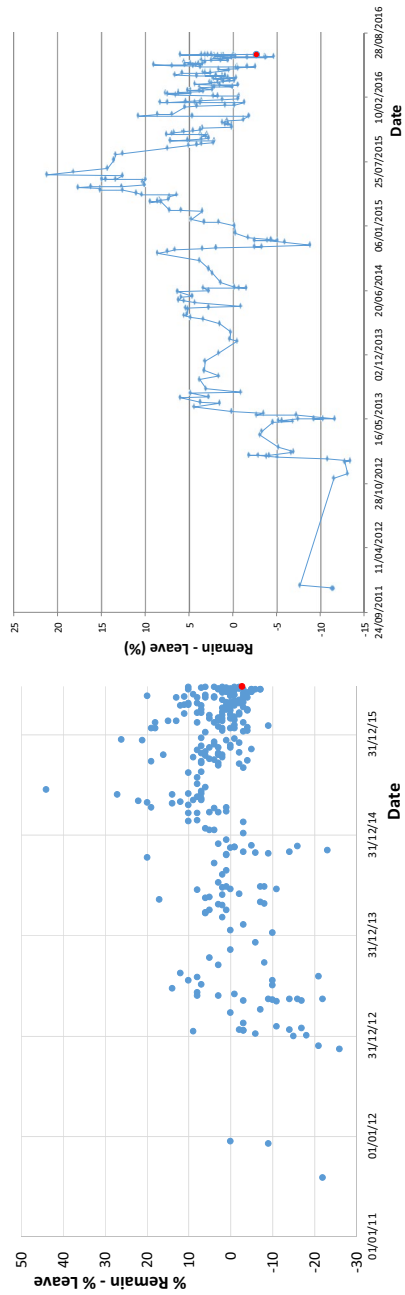


Fig. 1 (Left) %Remain—%Leave in polls over time, leading up to Brexit vote (red dot is final result). (Right) Adjusted, smoothed poll differences between Remain and Leave (%) during run-up to Brexit vote, adjusted for known biases in online and telephone polls (red dot is final result). (Color figure online)

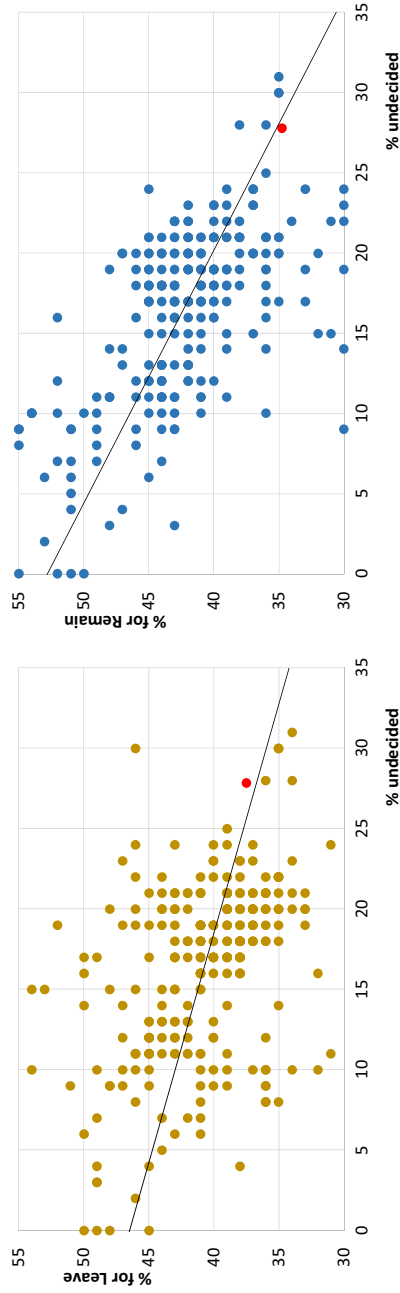
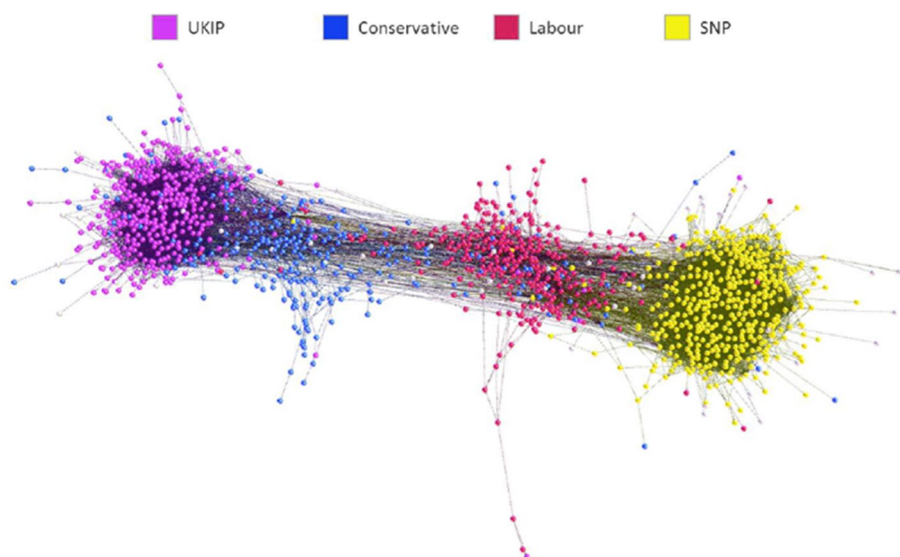


Fig. 2 (Left) % for leaving versus % undecided for opinion polls during run-up to Brexit vote, with linear regression line. (Right) % for remaining versus % undecided for opinion polls during run-up to Brexit vote, with linear regression line. (In both red dot is the final vote, mapping non-voters to undecided). (Color figure online)

Table 1 Percentage of users re-tweeted by user group, from Krasodonski-Jones (2016)

	Users re-tweeted			
	Labour (%)	SNP (%)	Tory (%)	UKIP (%)
Users re-tweeting				
Labour	65	<i>12</i>	<i>14</i>	6
SNP	<i>18</i>	78	8	3
Tory	<i>12</i>	5	46	<i>18</i>
UKIP	6	4	<i>32</i>	73

> 40% in Bold, ≤ 40% and > 10% in Italic

**Fig. 3** The re-tweet network between 4 groups of 500 supporters of four UK parties, from Krasodonski-Jones (2016)

those existing in the larger population, those between actors that are far more politically involved than the average citizen and thus is likely to be more polarised than the general population. Furthermore, Twitter users seem to be disproportionately biased towards those interested in political matters. However, it does give us a striking pattern of polarisation.

The patterns I take from these are:

1. opinion formation is noisy and not smooth,
2. there are periods of relative stability but some turning points,
3. undecided actors matter almost as much as those of decided or fixed opinion, and
4. people communicate with (and hence tend to influence) those similar to themselves, and in particular have a tendency not to interact with those with dissimilar beliefs.

These are the some of the patterns that I aimed for in the development of the presented model.

3 Theoretical background

In order to further constrain the model here, we have based the model upon a number of theoretical considerations, which I now discuss.

The “Social Intelligence Hypothesis” (SIH) (Kummer et al. 1997) states that the *crucial* evolutionary advantages that human intelligence gives are due to the social abilities it allows. This explains specific abilities such as: imitation, language, social norms, lying, alliances, gossip, politics, group identification etc. Under this view, social intelligence is *not* a result of general intelligence being applied to social matters, but at the core of human intelligence. One might even go as far as saying that so-called “general intelligence” is a side-effect of social intelligence (e.g. linguistic ability and the ability to learn beliefs from others).

Consistent with SIH is the following evolutionary story. Social intelligence allows humans to develop their own (sub) cultures of knowledge, technologies, norms etc. (Boyd and Richerson 1985) within their groups. These allow the groups *with their culture* to inhabit a variety of ecological niches (e.g. the Kalahari, Polynesia) (Reader 1980). Thus humans, as a species, are able to survive catastrophes that effect different niches in different ways (specialisation) since it is unlikely that all inhabited niches will be wiped out.

This means that different “cultures” of knowledge, skill, habits, norms, narratives etc. are significant (including how to socially organise, behave, coordinate etc.), and that these will relate to each other as a complete “package” to a significant extent. Under this view human cognition is (at least partly) evolved to make this group survival work, including the capacity of maintaining a complex set of beliefs that are coherent with others in the group, whilst retaining flexibility. Under this view humans are not evolved as primarily rational beings, but social beings, so we should expect their cognition to allow for some variation, but to maintain the set of beliefs that make up a complete culture. The model described herein exhibits such a combination of local coherency and global variety corresponding to the social structure.

Granovetter (1985) contrasted both under- and over-socialised models of behaviour of human behaviour—criticising sociologists in their picture of humans as culturally determined, and economists for their picture of agents characterised by self-interested behaviour. That is, that the particular patterns of social interactions between individuals matter. In other words, only looking at either individual behaviour or aggregate behaviour misses crucial aspects. Under this view, to understand the behaviour of individuals one has to understand the complex detail and dynamics of the interactions between them. Both individual cognition and how this relates to those of the society it is embedded into matter. Agent-based simulations are one of the few techniques that allow for such embedded behaviour to be represented and understood.

In particular, it allows for the integration of cognition and social processes, going beyond emergence and immergence to allow for a relatively ‘tight’ loop between the processes that go on in the head of an agent, the interactions between agents and their social structure. In agent-based simulations, these three things can be co-evolving, no aspect taking priority.

A theory of cognition that fits these perspectives is that of cognitive coherence (Thagard 1989). This is a theory that goes some way to explain how we choose what to believe and what not to. It is not based on any idea of logic or rationality, but rather on which mental entities are coherent with each other. Under this theory, new beliefs will usually only be accepted if they are coherent with the existing set of beliefs. As one matures the set of beliefs that one

has grows more complex, meaning that it might be more selective in which other beliefs are likely to be accepted. Many of these beliefs will be learnt from an early age, characterising the culture it comes from. Questioning and even changing core beliefs is hard, because it might mean introducing dissonance with many other existing beliefs which might then also need to be changed.

This theory explains some of our cognitive features. The well known “confirmation bias” can be seen as a corollary, since evidence that disconfirms our beliefs will be less palatable than evidence that confirms them. Given that active perception also relies on beliefs about the world (e.g. what is worth attending to), this can lead to us not noticing or summarily rejecting evidence that disagrees with our beliefs as well as why different scientific paradigms can be incommensurable (Kuhn 1962). For example, even after President Obama made his Birth Certificate accessible to public scrutiny, many in the US would not change their belief that he was not born in the US (Nyhan 2012). This is thus one mechanism whereby “Directionally motivated reasoning” can occur. As (Taber and Lodge 2006) says:

Directionally motivated reasoning leads people to seek out information that reinforces their preferences (i.e., confirmation bias), counterargue information that contradicts their preferences (i.e., disconfirmation bias), and view proattitudinal information as more convincing than counterattitudinal information (i.e., prior attitude effect)... (p. 757).

A later extension of this model (Thagard 2006) includes goals and emotions as entities that need to cohere. Thus one might not believe in climate change because you like driving your car, or might reject evidence about your performance because it would make you feel bad to accept it (Kruger and Dunning 1999). The model described here will be based on Thagard’s theory in order to incorporate some of the above features of human cognition.

To summarise this section I am aiming at a model:

- (a) Where agents do sometimes accept beliefs that are suggested to them
- (b) Where agents are socially embedded with a tight loop between cognition and social structure
- (c) In which there naturally develops a combination of some coherency within emergent groups but also with variety
- (d) Where “opinions” do not directly act on each other but emerge from a meaningful engagement of different beliefs
- (e) Where the coherency of the belief set drives belief change in individuals

4 ‘Linear’ opinion dynamic models

There has been a stream of models that aim to directly model the evolution of opinions within a group of interacting agents. These are the “opinion dynamics” (OD) models (e.g. Hegselmann and Krause 2002; Deffuant et al. 2002). In these the opinions of agents lie on a continuous line³ and these opinions directly effect each other when agents are sufficiently similar/near to each other on this scale. Thus the term ‘linear’ here denotes that

³ There are variations on this, a set of discrete values rather than continuous, a vector of binary values, more dimensions, but this does not change the resultant behavior much (Flache et al. 2017).

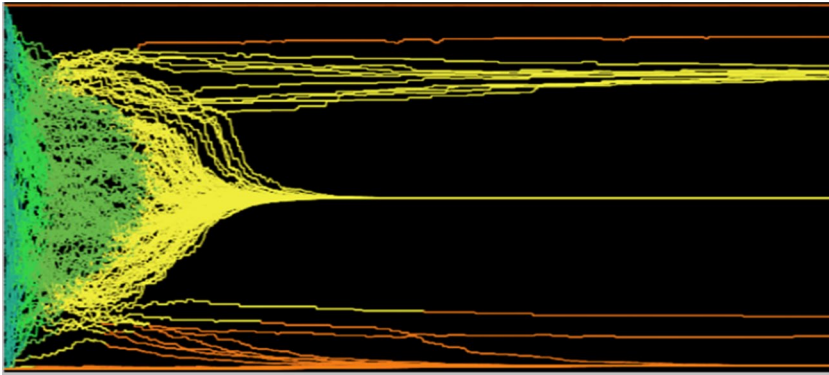


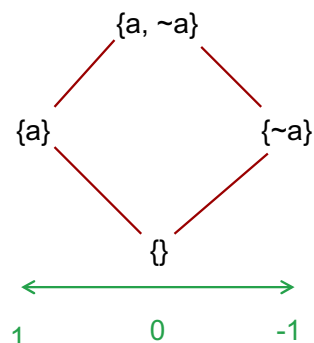
Fig. 4 A typical evolution of opinions from an opinion dynamics model. Horizontal axis shows the strength of the opinions (from -1 = completely disagree, to 1 = completely agree). Time is along the x-axis. The colour shows the “certainty” of the agent from red = certain to blue = uncertain. (Color figure online)

the interaction occurs on a line of values (the opinions)—there is no structure in the opinions—rather than that the model displays simple dynamics in the sense of “non-linear systems”. An example run of such a model is shown in Fig. 4.

However, such models exhibit resultant behaviour that is very different from that we observe (for example as exhibited during the Brexit vote as we saw previously). Firstly, the behaviour in OD models results in a number of stable groups that subsequently do not change. Secondly, typically all agents are influenced towards increased certainty as time progresses, so there are fewer (sometimes no) uncertain agents left at the end. Thirdly, after a while, all opinions settle down to stable values.

But there is a more fundamental reason to object to OD models. They represent opinions in terms of a linear scale (usually one dimensional and continuous) and further assume that influence occurs directly between agents via these values. In particular OD models do distinguish between agents who are conflicted and those who have no opinion. To take a simple example, consider the case of a single belief, “ a ” or its opposite “ $\sim a$ ”. This gives rise to four possibilities: believing neither a or $\sim a$, believing a only, believing $\sim a$ only and believing both a and $\sim a$ (as illustrated in the top half of Fig. 5). I have a feeling that including such dissonant sets of beliefs (as in a and $\sim a$) is particularly important for modelling political discussions, as this is a common experience of many individuals involved.

Fig. 5 The four possibilities of belief in a proposition “ a ” or its negation “ $\sim a$ ” (following Belnap 1977), compared to a linear scale



A ‘linear’ representation of opinions (as illustrated in the bottom half of Fig. 5) would represent these both as neutral in terms of belief. Whilst believing both a and $\sim a$ might seem daft at first glance, if you look at belief as rooted in evidence (or other process), then one may well have some evidence for a and some evidence for $\sim a$, and that would be a different situation from having no evidence for either (Belnap 1977). What all the OD models have in common is that opinion is essentially a matter of degree: degree to which opinions are held and, in many models, the degree of certainty of that opinion. Agents influence each other by their degrees maybe shifting the degrees of those they interact with.

What OD models miss out is any *structural* interaction with other beliefs. One method of doing this would be to have a set of beliefs that agents could hold (or not) with some kind of logical inference occurring between them. However, here I adopt an alternative theory, that of Thagard (1989, 2002). This is essentially a coherence model, that different beliefs are more or less coherent with each other so that a set of beliefs is more or less stable, due to the structure of coherencies between the individual beliefs. People seem quite good at tolerating low levels of incoherency but overall, people will tend to adapt their beliefs (dropping beliefs or adopting new beliefs) so as to increase the overall coherence.

To take a simple example, I might believe $MB = \{\text{Manchester is the best city in the world}\}$, $RM = \{\text{It rains quite a lot in Manchester}\}$ ⁴ and $LM = \{\text{I want to live in Manchester}\}$. Here a strong coherence between MB and LM might keep my belief in LM , despite LM being somewhat dissonant with my other belief, RM . If a new belief or experience weakens the strong coherence between MB and LM (e.g. I have kids and realise that although Manchester is the best city for adults it is not so good for kids), then the dissonance between RM and LM might lead me to drop LM , even though I retain a version of MB .

In this paper I wish to go beyond ‘linear’ models of opinions, towards a more belief-based representation, where the structure of individual’s belief sets matter. This is because I suspect that linear models of influence will not provide a good basis for *explaining* many observed collective processes.

5 The model

This model is an *illustration* of an idea—an idea of how cognitive and social structures might co-evolve to produce a collective phenomenon of opinion change in a population. Because the aim here is illustration rather than empirical, I have kept the model fairly simple. It aims to reproduce patterns, 1–4, listed at the end of the first section, the theoretical commitments, a–e, listed in the previous section and be based on plausible cognitive and social processes. It shows one way in which belief change and influence based upon the coherency of underlying beliefs can be implemented within a model.

This is an agent-based simulation. That is the individuals and their interactions are individually represented in the simulation, allowing agents to have different characteristics, for the social network to be dynamic and the outcomes to emerge from the

⁴ Of course, it does rain in Manchester more than many places, but this stereotype has been much exaggerated.

interactions between the agents. See (Squazzoni et al. 2014) for an introduction and survey of agent-based simulation in the social sciences.

In this model there are a fixed number of agents. These are connected via a set of links that represent the social network, which determines which other agents any agent can interact with. However, the network can change—any agent can drop an existing link or make a new link depending on what happens to it.

There are a fixed number of distinct ‘foreground’ beliefs in the model. Other beliefs that the agents hold are not explicitly represented here. Each agent can either hold or not hold each of these beliefs, so if there are n explicitly represented beliefs in a version of this model, then there are 2^n possible belief sets that each agent in that model could have. Thus this model could be conceived of as each agent having a binary belief vector, with each position having a 1 for belief or 0 for a lack of belief for each one. It is assumed that sets of agents share a set of ‘background’ beliefs that do not change but are not explicitly represented. For many purposes it would be impractical to try and represent *all* the beliefs of those involved—we want to understand and track a relatively small number of contested beliefs.

In this model beliefs are not independent of each other. That is, if an agent believes A this will affect whether it also believes B. This is achieved here using a function of belief coherency, which is a generalisation of Thagard’s pairwise (in)coherence (as discussed above). It gives a measure of the extent to which the whole set of current beliefs is coherent. This allows for great flexibility in how coherent different belief sets are (compared to a pairwise dissonance/consonance between beliefs or a measure based on logical inference). For example we could have the coherency evaluations: $\{A\} \rightarrow 0.3$ and $\{B\} \rightarrow \{0.7\}$ but also $\{A, B\} \rightarrow -0.4$ if beliefs A and B are mutually inconsistent, but individually coherent (against the background beliefs). Thus, the structure of this model does not put any constraint upon how belief sets relate to their coherency, this is up to the programmer.

There are two important ways that the beliefs of agents can change. One agent may ‘suggest’ one of its beliefs to another agent it is connected with. The receiving agent may then adopt that belief, thus adding it to those it already has. In the other direction, an agent can unilaterally drop one of its beliefs. The tendency of gaining a new belief from another or dropping an existing belief in this model is monotonically dependent on whether it increases or decreases the coherency of the node’s belief set. This is a strong assumption behind this model, but a plausible one—all it says is that people tend towards more coherent belief sets. Here we implement this tendency as a probability of change within two processes.

There is another link in the chain needed here, a mapping from the extent of any change in belief coherency to the probability of the belief change occurring—a monotone function from changes in coherency to such probabilities. Depending on the shape of this, the strength of tendency towards more coherent belief sets can be determined. A relatively ‘flat’ function would mean that coherency was a soft influence on belief change, so that belief change is more random, e.g. allowing temporary decreases in coherency. A ‘steep’ function would have the effect that only belief changes that increased coherency would be likely to occur. Thus this mapping function affects how strongly coherence impacts upon belief change in agents: from noisy and tolerant of incoherence to unidirectional and ‘logical’. Let us call this the “scaling function”.

In parallel to this process of belief change are changes in the social network. The model without network change is described and explored in Edmonds (2012). There are two main processes whereby links can change: dropping and adding links. The first is where an agent drops a link to another agent. Here we are following the principle of homophily (Lazarsfeld and Merton 1954), so that agents with incompatible beliefs will

tend to interact less. The way this is implemented here is that if a belief that is suggested by an agent is not adopted then there is a probability the link to the suggesting agent is deleted. Given the processes of belief adoption the reasons for non-adoption are likely to be that the suggested belief was incoherent with the existing belief set of the receiving agent, corresponding to incompatible views. The other process is one of link creation. With a given probability, an agent will make a new link—to a ‘friend of a friend’ if there is one not already linked to, otherwise to a random other agent. This process can be seen as one of exploring and making new connections—of course, if their beliefs turn out to be incompatible then the link might be dropped again via the first process. There is no ‘magic’ comparison of the contents of agents’ heads here, both link dropping and creation are via socially plausible processes.

To help explore and illustrate what happens when the above processes happen together, there are different types of agents. Each type represents a set of agents with a shared set of background beliefs, represented here by the same coherency function, and the susceptibility to coherence, respected here by having same scaling function. In the examples below different types and their proportions are specified then randomly connected, supplied with random beliefs and the results of the above processes explored.

A more detailed description of the model can be found in the “[Appendix](#)”.

6 The co-development of individual belief and social structure

Two examples will be given to show some of the properties of this model and why it might be interesting to develop. The first, described in this section, is to show the essential co-development of individuals’ beliefs and social structure—why the combination of belief change and network change can result in qualitatively different outcomes than either belief change or network change on its own.

For purposes of exposition, beliefs are arbitrarily assigned colours and types of agents shapes. In the example in this (and the following section):

- There are 50 agents
- Simulations last 1000 simulation ticks
- There are, on average 3 arcs per node
- The copy-rate parameter is 0.3 (the probability that one belief is considered for being copied along any link)
- The drop-rate parameter is 0.075 (the probability that an agent considers dropping one belief in a tick)
- Probability of dropping a link is 0.2 (if there is a suitable candidate where a copy has been rejected)
- Probability of a random belief change is 0.001

If not otherwise mentioned parameter values for the runs below are those listed in the “[Appendix](#)”. In order to see the resultant possible patterns in terms of opinion change (as might be seen with repeated opinion polls such as in Fig. 1) We need some way of recovering an opinion from agents. Here (and in the following example) this is done in a simple manner whereby the belief sets: ({blue}, {}, {blue, yellow}, {yellow}) is mapped to (1, 0, 0, -1). The derived opinions thus reduce to the number of “blue” beliefs minus the number of “yellow” beliefs divided by the population.

This example helps give a flavour of the model. In this:

- There are 3 tracked beliefs: “yellow”, “blue” and “red”
- The probability of a new link is 0.01

The two kinds of agent are as follows.

- 20% of agents (stars) are such that the ‘yellow’ beliefs are attractive and the ‘blue’ ones unattractive (due to coherence with background beliefs). Here (yellow, neither, both, blue) are mapped to $(1, 0, 0, -1)$ respectively (red make no difference). They are also ‘strong minded’ in the sense that they only change their mind if it increases their coherence (a strong mapping function similar to the bottom graph of Fig. 21).
- 80% of agents (circles) are such that the ‘blue’ beliefs are attractive and the ‘yellow’ ones unattractive. Here (yellow, neither, both, blue) are mapped to $(-1, 0, 0, 1)$ respectively. They are also ‘weak minded’ in the sense that they only have a tendency to change their mind if it increases their coherence (more probabilistic in their belief change with a mapping similar to the top graph in Fig. 21).

Both kinds change their links (or not) similarly and both are agnostic with respect to the ‘red’ belief. Runs are initialised with random beliefs and network.

Here we try 10 runs of each variant: with no belief change and no link change; with belief change only; with link change only; and with both belief and link change. Output is shown in terms of some typical runs and some summary graphs. In snapshots of the runs, agents shown in colours indicating the mixture of beliefs held (or if none, grey)—so blue if they only hold the Blue belief, green if they hold yellow and blue, etc. The shapes indicate what coherency (and scaling function) the agents have, which are (at least initially) independent of their beliefs which are randomly initialised at the start.

When only beliefs are allowed to change, there is some sorting of agents, so agents who are connected are more likely to hold similar beliefs, but the fixed network structure limits the extent to which this can occur. There has been a shift in the balance of beliefs—a shift towards blue (Fig. 6).

When only links are allowed to change, then some structure can evolve that partially separates agents with different beliefs. The balance between blue and yellow beliefs is largely unchanged (remember there is some random change) (Fig. 7).

When both beliefs and network can change we find that a considerable sorting of kinds of agents has occurred, with the “stars” being marginalised. There has been a marked shift in overall opinion towards the blue (remember the purple are agents that hold the blue belief but also the red etc.) (Fig. 8).

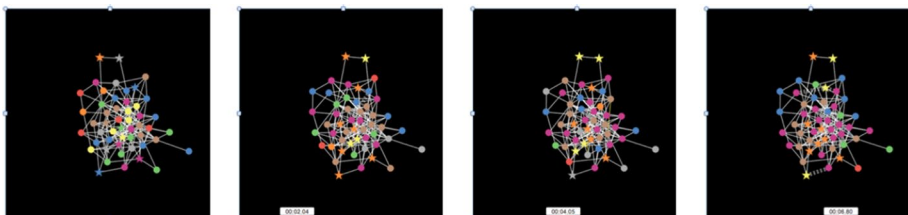


Fig. 6 Four snapshots of the run where only beliefs are able to change. (Color figure online)

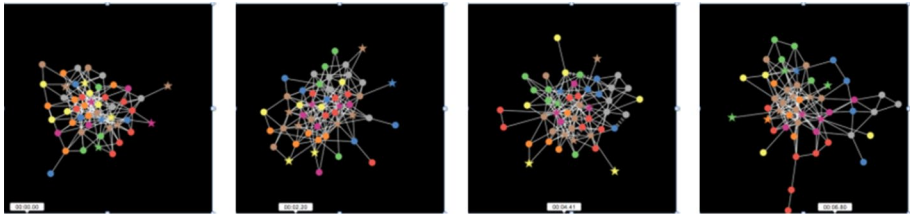


Fig. 7 4 snapshots of the run only allowing links to change

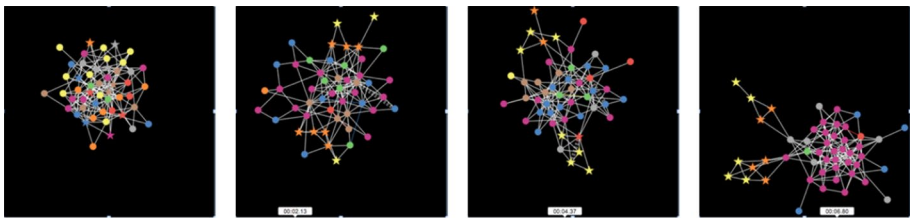


Fig. 8 4 snapshots of a run where both beliefs *and* links change

Fig. 9 Proportion of same kinds together at end (av. over 10 runs, “drop & add”=link change, “bel ch”=belief change)

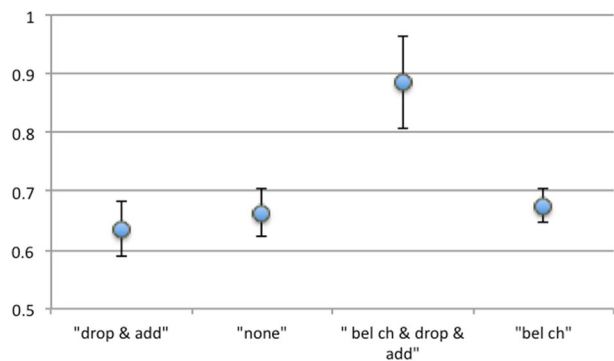
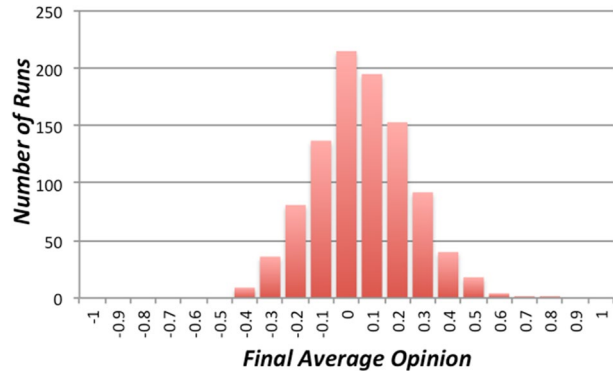


Figure 9 the proportion of links with the same kind of agent at its ends by the end of the simulations variants, averaged over 10 runs. This shows that the combination of belief change (a cognitive process) and link change (a social process) produces *qualitatively different results* from either belief change only or link change only. Sometimes one has to study cognitive and social processes together if one is to understand how they might work in practice.

7 Towards a Brexit example

In the second set of results we move towards the kind of situation that occurred in the Brexit referendum. Here we have three groups: floaters (most voters), yellows (Leave campaigners) and blues (Remain campaigners). The different types reflect anecdotal

Fig. 10 Distribution of final average opinions over 1000 runs



observation that whilst leavers argued using fierce emotive language and seemed largely immune to corrections of fact or argument, the remaining arguments were more equivocal, e.g. “Yes, there is a lot wrong with the EU, but on balance we are better off remaining”. On the other hand those arguing actively to remain outnumbered those for leave. Thus, the model is composed of:

- 70% Floaters (circles), these are weakly positive towards having either yellow or blue beliefs, but not both. So (yellow, neither, both, blue) are mapped to (1, 0, -1, 1) respectively. They have a weak scaling function so they are more open to change and more tolerant of temporarily tolerating moves to lower coherence (so similar to top curve in Fig. 21).
- 10% Leavers (stars) are for yellow and against blue with a strong scaling function (leave) as in the last example.
- 20% Remainers (triangles) are for blue and against yellow (as in the last example), with a medium scaling function (remain) (with a curve somewhere between those in Fig. 21).

Groups start separate (to allow for self-reinforcement), so that initially nodes are only linked with others of their type. They are initialised with random beliefs.

There are two beliefs: blue and yellow, the opinion function is the same as before (blue–yellow). The probability of a new link is 0.025 in these runs. There are no variants of the run (other than having different random seeds). All runs have the same proportions of agents, kind of initialization and parameters.

When the simulation is run 1000 times, we get the distribution of final average opinions as shown in Fig. 10. One can see that, given this set-up there is a slight overall bias towards blue (remain) outcomes.

Below I will merely display 4 specific runs to show the kinds of dynamics that this model can display. For each of these I will show 4 snapshots of the state of the model (as before) plus a graph that shows how the average opinion changed over time.

7.1 Example run 1

In this run the stars connect with floaters first, followed by the circles which has the effect of polarising the floaters, which then separate off into two groups. In each of these groups, the campaigners slowly convert the floaters to their own colour. One can see how the campaigners of both sides are now embedded within tightly formed groups (Fig. 11). The

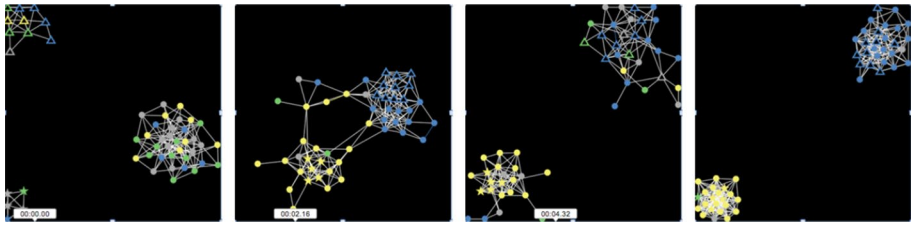


Fig. 11 Four snapshots of example run 1

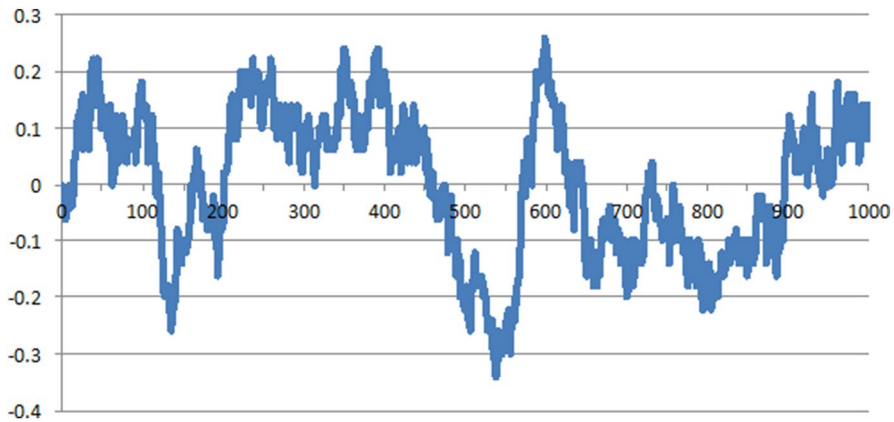


Fig. 12 The changing average opinion in example run 1

average opinion oscillates between blue and yellow over time but the blues gather the biggest group in the end and win (Fig. 12). This might correspond to a situation where there two separate groups formed, each lead by their own influencers.

7.2 Example run 2

In the second run a similar thing happens, two groups are established (earlier) which are stable for a while. Then some yellow mutate within the blue island, which then pulls apart. The yellow group then integrates into other and converts some more of the floaters (Fig. 13). Figure 14 shows the stability of the overall opinion for blue until towards the end where the yellow group manage to connect into the main group and shift some more floaters to their side, winning in the end. This could relate to situations where separate groups can join up and produce a less polarised range to opinion.

7.3 Example run 3

In this run, we have a ‘Westminster bubble’ of stars and triangles forming, separate from the floaters (Fig. 15). Free from influence from campaigners, floaters flip each other back and forth but in a random walk, which happens to end with more yellow (Fig. 16). This

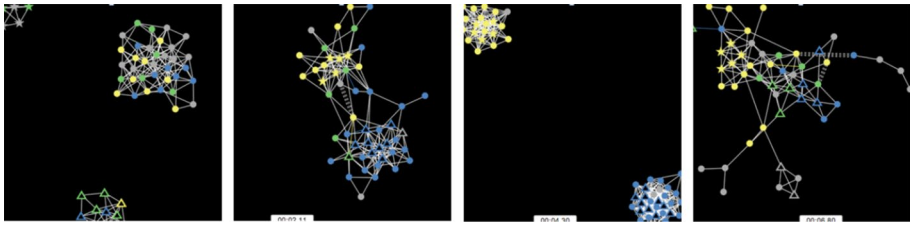


Fig. 13 Four snapshots of example run 2

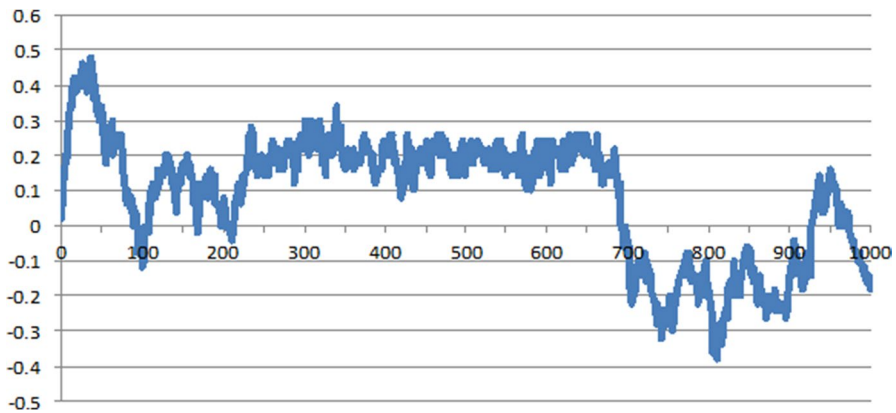


Fig. 14 The changing average opinion in example run 2

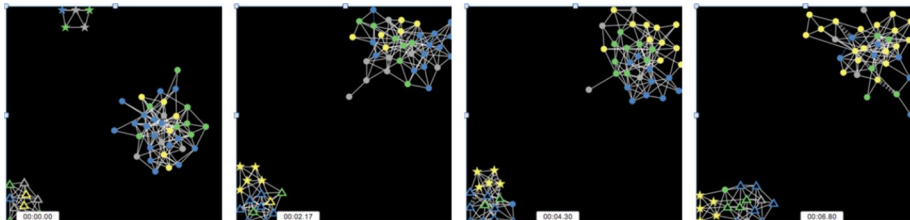


Fig. 15 Four snapshots of example run 3

could correspond to a situation where politicians focus on influencing each other, not connecting with the general public and hence allowing their opinion to drift.

7.4 Example run 4

In this run, both stars and triangles connect with floaters, but triangles more intimately, stars always on the peripheral. The yellow stars connect repeatedly to the main group but are isolated again each time, so they do not ever gain enough influence to convert many floaters (Fig. 17). The average opinion oscillates but always on the blue side (Fig. 18).

If one measures the end diversity of beliefs in linked nodes and the homogeneity of types (circles etc.) linked over 1000 runs one ends up with two distinct clusters of

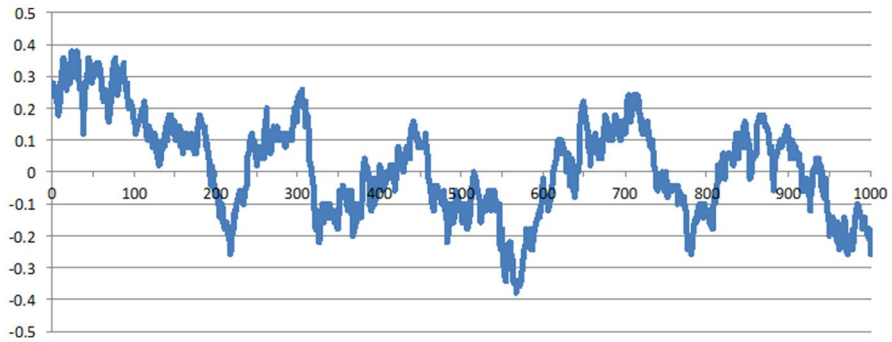


Fig. 16 The changing average opinion in example run 3

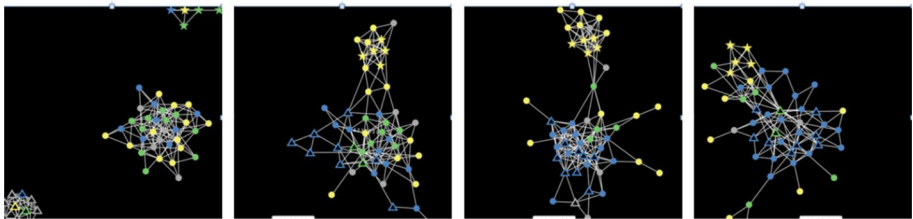


Fig. 17 Four snapshots of example run 4

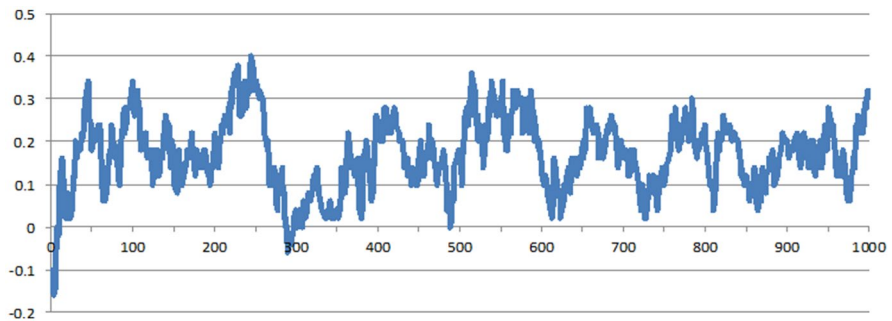


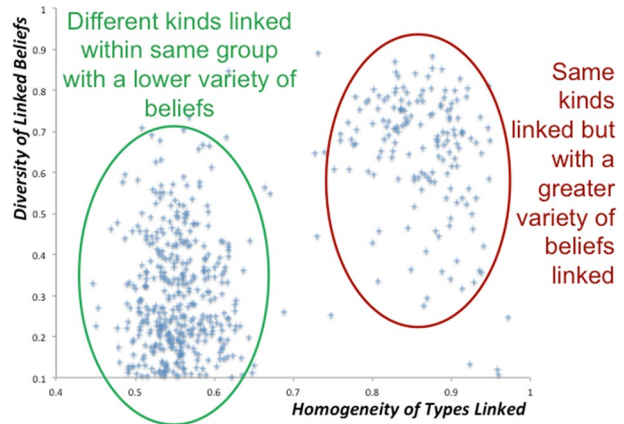
Fig. 18 The changing average opinion in example run 4. (Color figure online)

outcome (Fig. 19). These correspond to a situation where all agents are linked and hence a range of beliefs as they moderate each other (the green cluster in Fig. 19) and where clusters of more similar kinds are linked together, but beliefs are more polarised.

8 Model limitations and extensions

This model is highly stylised with many simplifications made consistent with its aim of illustration. These include that: there is a fixed number of agents, only a limited number of foreground beliefs are represented, beliefs are individually identifiable and discrete, belief

Fig. 19 A scatter graph of the diversity of beliefs between linked nodes and the homogeneity of types of agents linked at the end of 1000 runs of 1000 ticks. (Color figure online)



propagation is only done agent to agent (there is nothing corresponding to broadcast media here), dropped beliefs are not remembered but have to be re-suggested by a link neighbour, the coherency and scaling functions are unchanging for each agent, and that a social network adequately represents the communication possibilities between agents.

However, many of the limitations imposed here are not necessary. The agents need not be fixed but agents could drop out or and join the simulation. Super agents that suggest beliefs to many others could be added to represent broadcast media (the process of dropping links would naturally lead to agents selecting the media that suited their view). The scaling functions here are fixed for each type of agent, but these could be allowed to develop along with other aspects of the simulation, but a simple accumulation of beliefs can already have the effect of ‘locking in’ complicated belief structures once they have accumulated. Entities representing agent goals and emotions could be added. More complicated algorithms for agent social structure could be added, such as forming links through the institutions one participates in as in (Fieldhouse et al. 2016). Beliefs could be given more internal structure allowing some inference along with the coherence. However, for the purpose of illustration of the key ideas, these needlessly complicate the model, so have not been adopted here. If a more empirical model was attempted one might add some of these.

This kind of model does have more options and parameters than a standard OD model. Furthermore OD models are easier to compare to evidence from such as experiments—one can simply measure the opinions of participants before and after interaction, and do not have to delve into the murky world of beliefs etc. This reflects the difference in purpose—whilst some intend for OD models to be predictive of (average) opinion change, given an initial mix of opinions and certainties, the model presented here is more descriptive in intent—a wish to represent and understand how opinion change may occur. I would like to be able to explain the observed outcomes and patterns in opinion change come about from basic mechanisms of belief change, but this is still someway off.

9 Related models

There are two models that do look at the interaction of different beliefs. The first is (Jager and Amblard 2005a, b) which looks at a two dimensional model of opinions, based on “Social Judgement Theory (Sherif and Hovland 1961). That is there are three zones of interaction:

one when attitudes are sufficiently similar (in which case they converge a bit), one when they differ sufficiently to have the opposite effect and a neutral zone. This is combined with the “Elaboration Likelihood Model” of (Petty and Cacioppo 1986), which identifies two routes to attitude change: a central (conscious) and a peripheral (unconscious) route to attitude change. Although the two dimensions are basically independent, they explore different processes for agent–agent interaction within a 2D grid that link the two, for example closeness in the central dimension affecting the attitudes in the peripheral dimension. They conclude that that peripheral processing can be “responsible for the emergence of a correlation between originally unrelated issues”. This model does not deal with attitudes that are meaningfully related for the agents (the relatedness emerges as a result of the social process) but might explain effects such as described in (Macy et al. 2015).

Friedkin et al. (2016) model situations with a fixed weighted influence network between nodes and a universal matrix of co-certainty between beliefs. Each time click, beliefs are propagated to others within the agent via the co-certainty matrix, then linearly influenced by the corresponding beliefs of other agents via the weighted network, but this is moderated by the self-weights of nodes that ‘anchor’ the resulting beliefs to the nodes original belief (depending on its self-weight). They illustrate this model with a few examples, showing how the certainty of beliefs converge under group interaction in a similar, sometimes multimodal, but smooth manner shown in the OD models above. Whilst this is based on linear algebra and so can be rapidly calculated and analysed, but this does mean that the model tends to smoothly converge, but maybe to a series of clusters. They do not explore the model with any negative values in the influence weights or in the co-certainty matrix (which might allow for mutually incompatible beliefs). The influence of nodes on each other is the same for all their beliefs, and all nodes have the same co-certainty matrix.

10 The prospects for validation and integrating evidence on beliefs

In Moss and Edmonds (2005) we suggested an approach of “cross-validation”. That is, using qualitative evidence to inform the specification of an agent-based model, but then validating the outcomes against available statistical or time-series evidence. This is a way of checking a simulation that is used to explain the aggregate outcomes from the micro-level processes that are used to program agent behaviour. However, the measured outcomes from the simulation just described will match any independent data (such as those on Brexit) in terms of exact value—this is not its purpose. Rather what is possible is that a future version of this model could simultaneously match the patterns of such processes. That is, it would be a pattern-oriented modelling approach (Grimm et al. 2006), where many different patterns are matched at the same time. Matching many patterns simultaneously is difficult to fudge with simple model calibration and (given enough patterns) can be as effective at constraining model possibilities as matching up a couple of graphs in terms of simple value. Roughly, in order not to be deceiving ourselves, the constraint from evidence (of all kinds) needs to be greater than the ability of the model to fit these constraints by adjusting arbitrary elements in the model (both processes and free parameters). This can be addressed from both sides: by increasing the number of patterns that the model has to reproduce, but also limiting the elements in the specification of a model that are not supported by evidence—the ‘KIDS’ approach (Moss and Edmonds 2005).

However, in this case this depends upon our ability to inform the specification of the simulation using the available evidence, including qualitative evidence (Lotzmann and

Neumann 2017). Ideally, for this model we would be able to identify the relevant ‘fore-ground’ beliefs (those that are subject to doubt and change) and (at least approximately) map their relative coherency with each other. However, this is exactly what (Boutyline and Vaisey 2017) do, they use survey data from a population of individuals to infer the strength of association between different beliefs. This kind of mapping could be used as a base line for identifying beliefs and their coherence; this could then be supplemented by surveys of participants on which of these beliefs they held.

In this model, although the processes are plausible and accord somewhat to some existing theory, they are not evidence-based. To put these on a better footing, qualitative accounts as to how and under what circumstances people change their minds (accepting new beliefs or dropping existing ones) or alter who they interact with could be used. This would take narrative accounts from people (either written themselves or transcribed from an interview) and analyse these to inform computational representations of these for specifying the rules for agents in a simulation (Edmonds 2015).

11 Concluding discussion

In this model, here are ‘competing’ processes of social influence (suggestion) versus internal coherence with existing set of beliefs; also between social influence versus social linking. Thus, an ‘extreme’ group may be good at convincing another group when well connected to that group but groups tend to disconnect from those with very different views to themselves.

How processes actually happen may matter a lot, so it may be that this model has these wrong. We just do not know what influences people’s change of links—do people have a ‘whitelist’ of those they are willing to allow to influence them? In addition, this model does not touch upon the development of people’s belief structures within their society during acculturation or youth.

However, this model does suggest possibilities. It may be that how we act collectively is not through a direct spread (imitation) of action but via a spread of beliefs, norms, stories, habits etc. from which directed action springs. This would allow for specialisation and diversity of action whilst maintaining coherency. In this picture culture (any pattern, knowledge, norms, technology passed down the generations etc.) is important, but the group-structure of society is dynamic and can be complex (almost fractal in structure).

It does vividly show that if one modelled only the belief change/influence processes or only the social network processes then one could be missing significant aspects of what might be happening during these complex socio-cognitive processes. Thus, the main conclusion might be that we sometimes need to model both social and cognitive processes together—mutually influencing each other—to capture some social phenomena.

Like a tethered goat (Fig. 20), individuals may find it hard to wander too far from the belief set of the group it is currently attached to,⁵ but we might be able to choose to which group we are tied.

⁵ It is always possible to question some of one’s beliefs, but not all of them.

Fig. 20 A tethered goat

Acknowledgements Many thanks to discussions with Guillaume Deffuant, Edmund Chattoe-Brown and Giangiacomo Bravo about opinion dynamic type models and their alternatives (including this one). Also to the participants of the Workshops on *Coherence-Based Approaches to Decision Making, Cognition, and Communication* in Berlin, the symposium on *Coherence-Based Approaches to Decision Making, Cognition, and Communication* in Odense, the *Interdisciplinary Workshop on Opinion Dynamics and Collective Decision* in Bremen, and the *Social Simulation* conference in Dublin.

Funding Funding was provided by European Commission (Grant No. 822337) under the Horizon 2020 programme.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: more about the model

This model is an extension of the one discussed in Edmonds (2012), but with network change processes added. The full model has quite a number of options and extensions not discussed here. The code and full documentation can be accessed from the CoMSES.net archive (Edmonds 2016).

Basic entities and processes

In this model:

- There is a network of a fixed set of nodes and arcs (that can change)
- There are, n , different beliefs $\{A, B, \dots\}$ circulating between nodes
- Beliefs are copied along links or dropped by nodes according to the change in coherency of the node's belief set that this would result in
- Links can be randomly made
- Links are dropped when beliefs are rejected for copy between nodes

Node properties

Each node has:

- A (possibly empty) set of these “beliefs” that it holds
- A fixed “coherency” function from possible sets of beliefs to $[-1, 1]$ where 1 is completely coherent, 0 is neutral and -1 is maximum incoherency.
- A fixed scaling function that maps changes in coherency to the probability of a change in beliefs
- A record of the last node it “rejected” a belief from

Initialisation

Beliefs and social structure are randomly initialized at the start according to some global parameters. In the variants explored here there can be up to 3 types of agent, which are only distinguished by their coherency and scaling functions.

Coherency function

The coherency function is a function from all possible sets of the possible beliefs to real values in the range $[-1, 1]$. This is the key measure on beliefs—changes in this determines the probabilities of belief change by agents.

Belief change processes

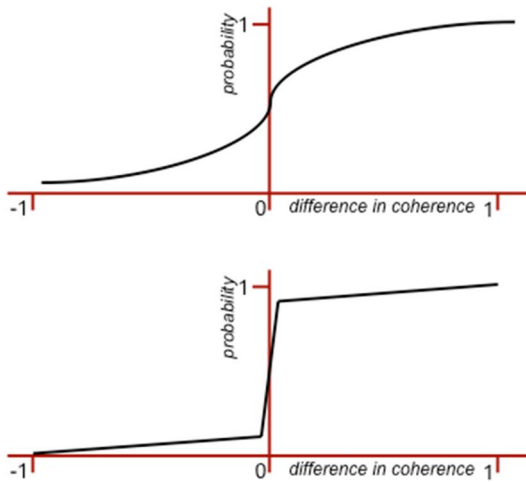
Each iteration the following occurs:

- *Copying*: each arc is selected; a source end and destination end selected; a belief at the source is randomly selected; then copied to the destination with a *probability* related to the change in coherency it would cause (due to the scaling function described next).
- *Dropping*: each node is selected; a random belief is selected and then dropped with a *probability* related to the change in coherency it would cause

Scaling impact of coherency function

There is a variety of ways to map a change in coherence to a probability (of a change occurring). The function that maps from changes in coherency to probability could be any that: (a) is monotonic (b) such that a $-1 \rightarrow 1$ change has probability of 1 (b) a $1 \rightarrow -1$ change has probability of 0. Two example such functions are illustrated in Fig. 21.

The scaling function in this model thus affects how amenable an agent is to change and the extent to which it may change. E.g. whether only to increase coherency or if it can occasionally decrease. A more gradual function (such as the top graph in Fig. 21) means that the “pressure” towards coherency is relatively weak, so more often changes of belief might temporarily go in the direction of greater incoherence, for example accepting a belief



A '*weak*' mapping –
probably changes to
increase coherence

A '*strong*' mapping –
almost certainly only
changes to increase
coherence

Fig. 21 Two example mappings from a change in coherency to a probability (of either a “copy” or a “drop” of a belief)

that is not completely compatible with existing beliefs. A sharper function (such as the bottom graph in Fig. 21) means that only changes that increase coherency are likely to occur.

Network change processes

There are two processes for changing the influence network. Each iteration the following occurs for each agent:

- *Link drop*: with a probability: if a belief copy was *rejected* by the recipient, then drop that in-link.
- *New links*: with another probability, create a new link with a random other (with a friend of a friend if possible, otherwise any)

Recovering opinions from beliefs

The “opinion” of agents is derived from the belief state of the agents. There are a number of ways in which one could do this. One interesting way might be to say the opinion of an agent on an issue, X , is the change in coherence that would occur to the belief set of that agent if X were added to their beliefs. In the examples above I chose a simple function from the belief set to $[-1, 1]$, namely $\{\text{blue}\} \rightarrow 1$, $\{\} \rightarrow 0$, $\{\text{blue, yellow}\} \rightarrow 0$, $\{\text{yellow}\} \rightarrow -1$. The global opinion is an average of this function applied to each agent.

Other

In order to maintain the average link density I added the following ‘kludges’: If there are too many links (as set by arcs-per-node) slightly increase the rate of link drop, if there are

not enough, slightly reduce the rate of link drop. Also, nodes have to have at least one link, or one is added, to stop isolates forming.

Finally there is a small probability that a belief is randomly added or dropped, this adds a little bit of extrinsic noise into the system and stops beliefs disappearing (through chance) from the entire population (as discussed in Edmonds 2012). This is kept at a low level in all the examples discussed, 0.001 per node per tick.

Parameters

A full list of parameters with descriptions maybe be found in the documentation accompanying the model (Edmonds 2016). Important parameters for our purposes include the following (the default value is shown in brackets following the description).

- *num-agents*: number of agents in the simulation (50)
- *num-beliefs*: number of atomic beliefs around (2)
- *init-prob-belief*: probability that agents hold each of the atomic beliefs at the start (0.5)
- *copy-prob*: the probability that a (random) belief from one agent will be attempted to be copied to another during the copy process (0.3)
- *drop-rate*: the drop-rate is the probability that an individual will do the drop process once (per simulation tick) (0.075)
- *mut-prob-power*: this is the power of 10 of the probability that a random belief of an agent is flipped each time click (so -3 is a probability of 0.001)
- *arcs-per-node*: how many arcs lead into each node on average (3)
- *init-sep-prob*: the probability that types only link to their own kind at the start (1)
- *init-prob-drop-link*: the probability of dropping a link (0.2)
- *prob-new-link*: this is probability of adding a new random link (0.025)
- *Opinion-Fn-Name*: the function that is used for recovering the opinion from the beliefs of agents which is then averaged for the global Opinion (blue–yellow)

References

- Adamic, L.A., Glance, N.: The political blogosphere and the 2004 U.S. election: divided they blog. In: Proceedings of the 3rd International Workshop on Link Discovery (LinkKDD '05). pp. 36–43. ACM, New York, NY, USA (2005) <https://doi.org/10.1145/1134271.1134277>
- Belnap, N.D.: A useful four-valued logic. In: Modern uses of multiple-valued logic, pp. 5–37. Springer, Dordrecht (1977)
- Boutyline, A., Vaisey, S.: Belief network analysis: a relational approach to understanding the structure of attitudes. *Am. J. Sociol.* **122**(5), 1371–1447 (2017)
- Deffuant, G., Amblard, F., Weisbuch, G., Faure, T.: How can extremism prevail? A study based on the relative agreement interaction model. *J. Artif. Soc. Soc. Simul.* **5**(4), 1 (2002)
- Edmonds, B.: Modelling belief change in a population using explanatory coherence. *Adv. Complex Syst.* **15**(6), 1250085 (2012). <https://doi.org/10.1142/S0219525912500853>
- Edmonds, B.: Using qualitative evidence to inform the specification of agent-based models. *J. Artif. Soc. Soc. Simul.* **18**(1), 18 (2015)
- Edmonds, B.: A Model of Social and Cognitive Coherence. CoMSES Computational Model Library. <http://www.openabm.org/model/5116> (2016). Accessed 10 June 2019
- Fieldhouse, E., Lessard-Phillips, L., Edmonds, B.: Cascade or echo chamber? A complex agent-based simulation of voter turnout. *Party Polit.* **22**(2), 241–256 (2016)
- Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., Lorenz, J.: Models of social influence: towards the next frontiers. *J. Artif. Soc. Soc. Simul.* **20**(4), 2 (2017). <https://doi.org/10.18564/jasss.3521>

- Friedkin, N.E., Proskurnikov, A.V., Tempo, R., Parsegov, S.E.: Network science on belief system dynamics under logic constraints. *Science* **354**(6310), 321–326 (2016)
- Granovetter, M.: Economic action and social structure: the problem of embeddedness. *Am. J. Sociol.* **91**, 481–510 (1985)
- Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmannith, E., Rütger, N., Strand, E., Souissi, S., Stillman, R.A., Vabø, R., Visser, U., DeAngelis, D.L.: A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* **198**(1–2), 115–126 (2006)
- Hegselmann, R., Krause, U.: Opinion dynamics and bounded confidence: models, analysis and simulation. *J. Artif. Soc. Soc. Simul.* **5**(3), 2 (2002)
- Homer-Dixon, T., Maynard, J.L., Mildenberger, M., Milkoreit, M., Mock, S.J., Quilley, S., Schröder, T., Thagard, P.: A complex systems approach to the study of ideology: cognitive-affective structures and the dynamics of belief systems. *J. Soc. Polit. Psychol.* **1**(1), 337–363 (2013)
- Jager, W., Amblard, F.: Multiple attitude dynamics in large populations. In: Agent 2005 Conference on: Generative Social Processes, Models, and Mechanisms, Argonne National Laboratory, The University of Chicago (2005a)
- Jager, W., Amblard, F.: Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. *Comput. Math. Org. Theory* **10**(4), 295–303 (2005b)
- Krasodomski-Jones, A.: Political Debate Online and the Echo Chamber Effect. Demos. <http://www.demos.co.uk/project/talking-to-ourselves/> (2016). Accessed 10 June 2019
- Kruger, J., Dunning, D.: Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J. Pers. Soc. Psychol.* **77**(6), 1121 (1999)
- Kuhn, T.S.: The structure of scientific revolutions. University of Chicago press (2012)
- Kummer, H., Daston, L., Gigerenzer, G., Silk, J.: The social intelligence hypothesis. In: Weingart, P., et al. (eds.) *Human by Nature: Between Biology and the Social Sciences*, pp. 157–179. Lawrence Erlbaum Associates, Hillsdale (1997)
- Lazarsfeld, P.F., Merton, R.K.: Friendship as a social process: a substantive and methodological analysis. In: Berger, M., Abel, T., Page, C.H. (eds.) *Freedom and Control in Modern Society*, pp. 18–66. Van Nostrand, New York (1954)
- Lotzmann, U., Neumann, M.: Simulation for interpretation: a methodology for growing virtual cultures. *J. Artif. Soc. Soc. Simul.* **20**(3), 13 (2017). <https://doi.org/10.18564/jasss.3451>
- Macy, M., DellaPosta, D., Shi, Y.: Why do liberals drink lattes? *Am. J. Sociol.* **120**(5), 1473–1511 (2015)
- Moss, S., Edmonds, B.: Sociology and simulation: statistical and qualitative cross-validation. *Am. J. Sociol.* **110**(4), 1095–1131 (2005)
- Nyhan, B.: New surveys show the persistence of misperceptions. *Huffington Post Pollster*. http://www.huffingtonpost.com/brendan-nyhan/new-surveys-show-the-pers_b_1718794.html (2012). Accessed 10 June 2019
- Petty, R.E., Cacioppo, J.T.: The elaboration likelihood model of persuasion. In: *Communication and persuasion*, pp. 1–24. Springer, New York (1986)
- Sherif, M., Hovland, C.I.: *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. Yale University Press, New Haven (1961)
- Squazzoni, F., Jager, W., Edmonds, B.: Social simulation in the social sciences: a brief overview. *Soc. Sci. Comput. Rev.* **32**(3), 279–294 (2014)
- Taber, C.S., Lodge, M.: Motivated skepticism in the evaluation of political beliefs. *Am. J. Polit. Sci.* **50**(3), 755–769 (2006)
- Thagard, P.: Explanatory coherence. *Behav. Brain Sci.* **12**, 435–467 (1989)
- Thagard, P.: *Coherence in thought and action*. MIT press (2002)
- Thagard, P.: *Hot Thought: Mechanisms and Applications of Emotional Cognition*. MIT Press, Cambridge (2006)