Philosophical Analysis in Modeling Polarization: Notes from a Work in Progress

Patrick Grim, Aaron Bramson, Daniel J. Singer, Steven Fisher, Carissa Flocken and William Berger

Introduction

Computational modeling and computer simulation have quickly established themselves not merely as useful add-ons but as core tools across the range of the sciences. We consider computational modeling to be a promising approach to a range of philosophical questions as well, and to questions that sit on the border between philosophy and other disciplines (Burkholder 1992; Bynum & Moor 1998; Holyoak & Thagard, 1997; Grim, Mar, & St. Denis, 1998; Grim 2004). Questions regarding the transference of belief, social networks, and opinion polarization fall in the latter category, bridging epistemology, social philosophy, sociology, political science, network studies and complex systems. These are the focus of some of our current research.

Our purpose here is not to sing the praises of computational modeling as a new philosophical technique. Our purpose is rather to emphasize the continuity of computational model-building with the long philosophical tradition of conceptual analysis (Hanna 2000; Sandin 2006; Beaney 2009). With reflections from the process of building a specific model, we want to emphasize two points: that 1) the work of constructing a computational model can serve the philosophical ends of conceptual understanding, in part because (2) attempts at computational modeling often require clarification of the core concepts at issue.

I. Computational Modeling and Philosophical Analysis

In their final form, papers in scientific computational modeling always look perfect: they appear to be the work of a rational investigator who thought things through step by step in advance: from methods, to results, to discussion and conclusion. It's good that these papers look that way—good for brevity, evaluation, and use in future work. That is how we want our work on belief networks and polarization to look eventually.

But, of course, the polished published form of a paper can give an entirely misleading impression of the research trajectory—the impression that both the conceptual work at issue and the path of design and programming were neat, tidy, and fore-ordained. Almost inevitably, they were not. We will use our current work in progress as an example. Here, unlike its future final form, we will lay out the research in something more like real time, complete with fits, starts, and second thoughts. A key point is that those fits, starts, and second thoughts often indicate the need for philosophical analysis in a fully traditional sense. Computational modeling calls for and enforces a full and explicit conceptual understanding of what it is one is trying to model. To employ computational techniques, one must have a full and explicit understanding of what it is one is trying to find out, within what parameters, with what background assumptions, and why (Pollock 1998).

We offer our current work on belief polarization as a case in point. The history of this project is one in which we have repeatedly had to ask what abstract representations of social information contact were plausible. We've had to ask and ask again whether certain modeling assumptions were realistic portrayals of belief and trust, and whether it matters the extent to which they clearly were not. The history of the project is one in which we have repeatedly had to return to questions of how to define and measure the phenomenon we were after, and even whether there was just *one* phenomenon at issue. This exploration, which is at the edge of various sciences, has repeatedly demanded far more than computational resources. Flying under the colors of updating algorithm design and definition of quantitative measures, for example, we repeatedly found ourselves doing just good old-fashioned conceptual analysis in a new-fangled computational terminology.

II. Understanding Polarization: Initial Motivations

What we wanted to know about was polarization of beliefs in society. We started with the impression that the increased polarization of America was an agreed and established sociological fact. Everybody talks about it and a range of books are written about it (McCarthy, Poole & Rosenthal 2006; Brownstein 2007; Hetherington & Weiler 2009; Fiorina, Abrams & Pope 2010), so we thought it must be real.

The idea was to use the tools of agent-based modeling to try to understand that polarization better—to understand the factors that influence polarization: factors necessary for polarization, perhaps a handful of factors sufficient for polarization, and perhaps even social measures that could be used to reduce polarization.

At the beginning, we had a hunch that increased polarization in America might have something to do with the structure of media sources. The core idea was the following: We seem to have been less polarized when there was essentially one source from which everyone got their news: the Evening News on ABC, NBC, and CBS. The news coverage on the three major networks was essentially interchangeable—all a version of Walter Cronkite. All followed a journalistic code that insisted that editorializing be kept strictly separate from reporting.

News is no longer like that. Fox News and MSNBC have obvious political slants, are positioned at rival ends of the political spectrum, and do not seem to care where journalism leaves off and the editorial begins. Perhaps the change in where we get our news has something to do with why America is so polarized.

That was the initial motivating hunch. Could a model illustrate that belief dynamics? Could it show us whether split news media was an easy route to polarization, or even a possible route? Could it give us hints as to what kinds of factors might ameliorate or reduce polarization?

We had worked previously with networks of artificial agents whose beliefs were modeled as numbers between 0 and 1 and who updated those beliefs in terms of the other agents with whom they had contact. We had used that abstraction in the context of investigating infection, belief transference, and genetic crossover as alternative modes of information diffusion on networks. All that work saw final presentation in polished form (Grim, Reade, Singer, Fisher, & Majewicz 2010; Grim, Singer, Reade, & Fisher 2011; Grim, Singer, Reade, & Fisher 2012).

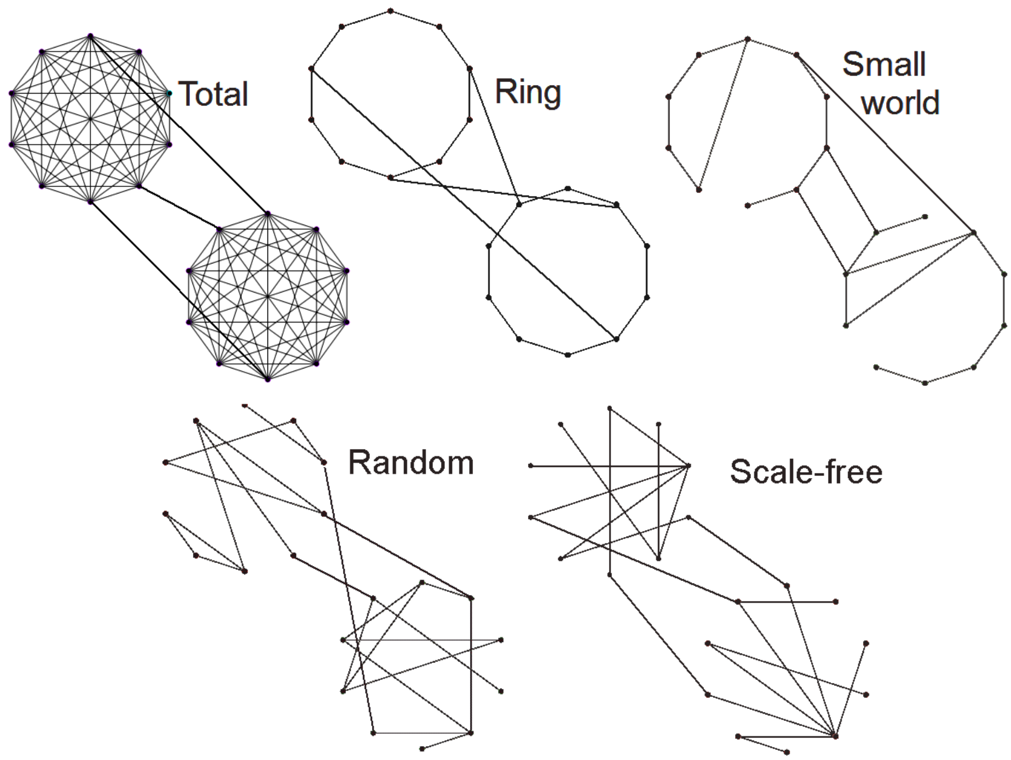


Fig. 1 Types of linked sub-networks used in previous work on belief and infection dynamics (Grim, Reade, Singer, Fisher, & Majewicz 2010; Grim, Singer, Reade, & Fisher 2011).

We had used a more complicated version of that kind of belief updating in building models of information networks for Black and White communities, based on data in the Greater Pittsburgh Random Household Health Survey. In this latter model we had also used data on trust: what kind of trust do members of each community put in information they receive from the government, for example, from their friends and family, from their church or religious leaders (Figs 2, 3)?

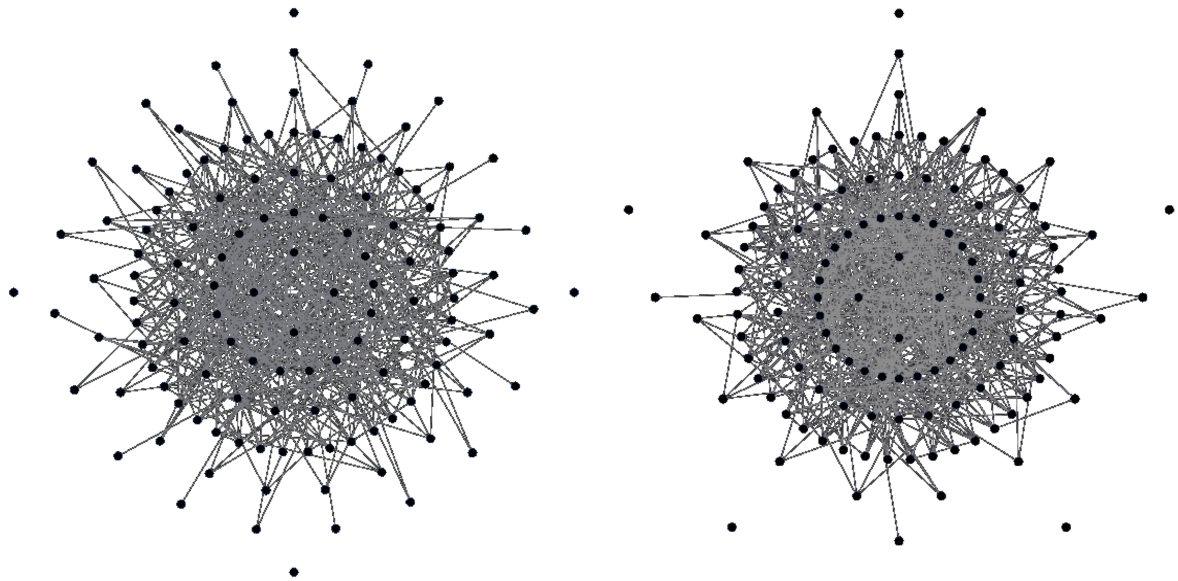
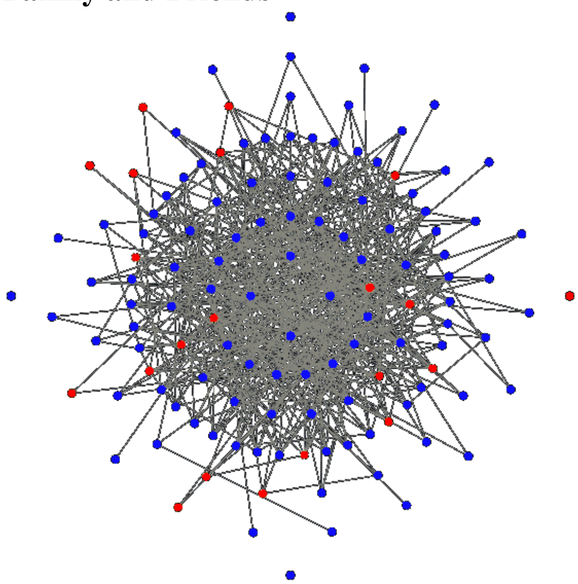


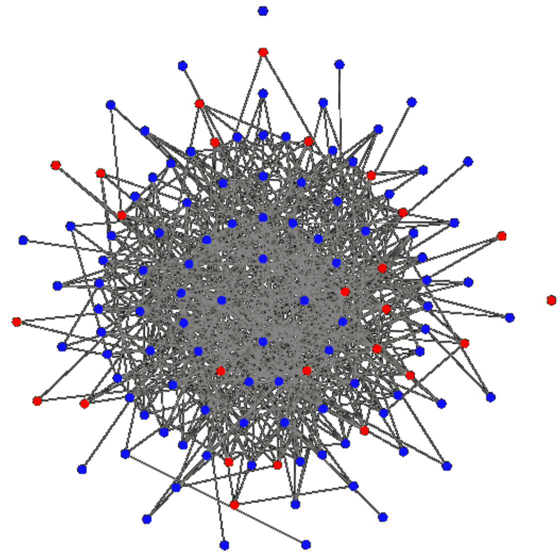
Fig. 2 Histograms and networks constructed to match degree distributions drawn from data within the Black and White communities, Pittsburgh Random Household Health Survey (Grim, Thomas, Fisher, Reade, Singer, Garza, Fryer, & Chatman, 2012a, 2012b)

Trust Levels - Black community

Family and Friends



Government



Religion

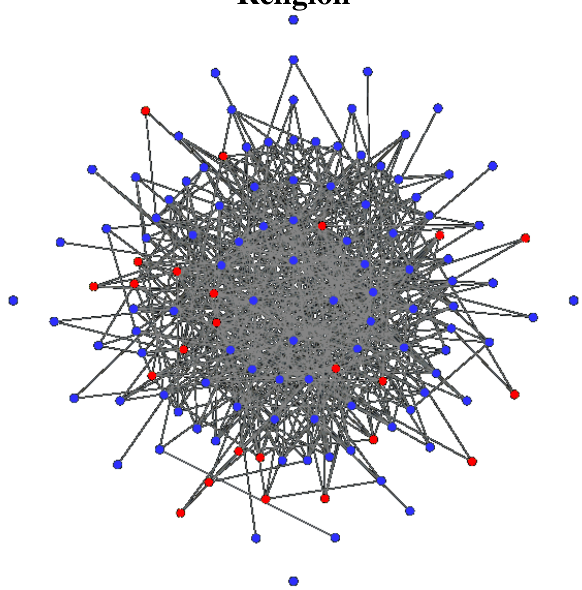


Fig. 3 Trust levels in the Black community correlated with network position, Pittsburgh Random Household Health Survey. Red nodes indicate low trust; blue nodes indicate high trust (Grim, Thomas, Fisher, Reade, Singer, Garza, Fryer, & Chatman, 2012a, 2012b)

This last piece of work had shown patterns of belief polarization in the two communities given conflicting input from, for example, governmental and religious sources. Why not apply the computational techniques developed in this earlier work, geared to belief change on networks and the effect of trust, in order to try to understand opinion polarization more generally?

III. The First Models

Our initial models were built along the following lines. Model individuals are connected via a communication network. They start with randomized 'beliefs' modeled as numbers between 0 and 1. They update their beliefs based on the beliefs of their neighbors on the network. The idea is simply that we are influenced by the beliefs of those around us. If my friends all confirm my beliefs, those beliefs will be reinforced. If my contacts all seem to believe something different than I do, my beliefs can be expected to shift in that direction over time (Visser & Cooper 2003).

In practice we made belief updating a weighted averaging of an agent's previous belief and the beliefs of other agents with whom he had informational links in the network. Is this artificial? Certainly. Implausible? Not as a rough approximation, perhaps. Precedented in the literature? Numerous times (French 1956; Harary 1959; DeGroot 1974; Golub & Jackson, 2010, forthcoming). What we were after was an explanatory model; as modeling assumptions go, that representation of belief reinforcement seemed a promising start.

From the beginning, however, we also wanted to build in issues of trust. Here again, the goal was to start with something simple. The simple assumption we started with was that widely divergent opinions can strain bonds of trust (Lord, Ross & Lepper 1979). If the views expressed by a particular source are views I consider radically incorrect, wrong, or misguided, then *ceteris paribus* I can be expected to discount information from that source.

Our first models therefore had two forms of updating running in tandem: a belief updating in terms of a weighted averaging of my network contacts, and a trust updating based on belief distance that is reflected in those weights. The hypothesis was that we can more fully understand the dynamics of belief polarization in terms of the interplay between (a) belief revised in terms of trust and (b) trust revised in terms of belief.

Perhaps the fact that people discount information from contrary sources is enough to explain polarization. Perhaps a single media source—Walter Cronkite, CBS, NBC, ABC—would tend to counteract that force toward polarization. Perhaps multiple media sources—Fox and MSNBC, or the infinite number of sources one can find online to reinforce any chosen—would tend to make polarization worse.

III. Conceptual Questions from Computational Models

It was at this point in model development, however, that things started getting messy. They got messy both in the model results and in the conceptualization of the model itself.

In the first models we built, given our initial updating assumptions for belief and trust, we kept getting convergence rather than polarization. Polarization didn't seem easy to produce, even with contrasting media sources. We therefore had a wonderful model illustrating the fact that everyone is always destined to come to the same view on everything—a model that explained perfectly something that we knew didn't really happen.

From another direction, and independently, we began to worry about conceptual foundations. A major issue was trust. As one of the research group repeatedly reminded us, trust can be of various forms, from various sources. Bob has great trust in the thinking of his friend Alice. He takes Alice's views seriously and pays close attention to Alice's arguments and evidence, despite the fact that they are often in wide disagreement.

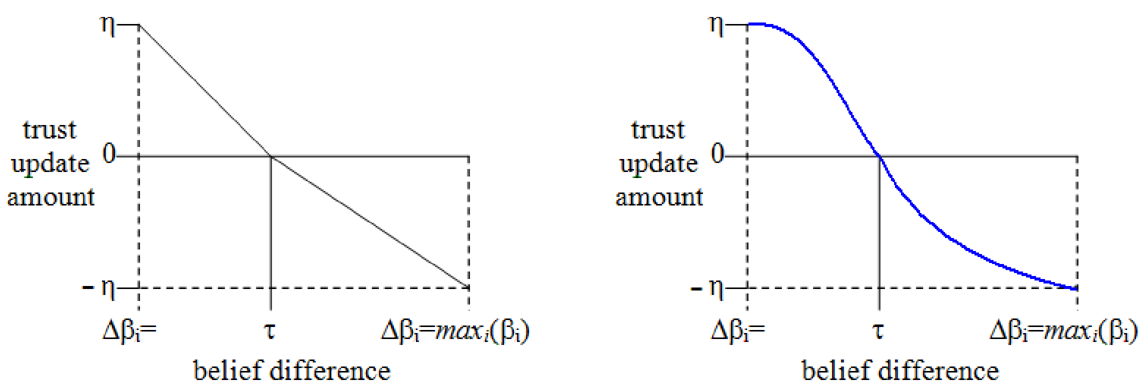
That is a classical philosophical counter-example. It shows, quite legitimately, that trust does not correlate with belief distance alone. We have clearly over-simplified. But is that over-simplification one that can be tolerated for purposes of modeling? Is it a modeling assumption that *could* be used *ceteris paribus*? Might we build a model in which we tracked the effect of that factor as if it *were* the only one, drawing conclusions of the explicitly hypothetical form 'were trust a matter simply of belief distance…?' Or is that clear over-simplification a modeling assumption that goes too far, losing track of the phenomena with which we are really concerned?

We worried that belief was single-issue and one-dimensional in our model, and that trust followed suit. Our real beliefs are multiple, and our disagreements often reflect that. I may come to trust you on one issue in one hundred, despite initial disagreement, if I have learned to trust your judgment in the other ninety-nine.

All of these are conceptual issues of a type that should be familiar to philosophers: conceptual issues regarding what belief and trust are and how they change. Here those issues arise in terms of the interpretation of a computational model: are belief and trust enough like their 'representations' in the formal model to allow us to draw useful conclusions from that model, or have we sacrificed so much in the course of model simplification that we have disqualified ourselves from genuine conclusions regarding the dynamics of belief?

Goals of simplicity play a significant role in evaluating models. A model is useful only if it is simpler and easier to understand than the reality it is meant to capture, but is also useful only to the extent that it matches its target in those respects relevant to the purposes of design. Whether a model has adequately captured the relevant respects, and captured them in relevantly significant degree, is always an open question (Miller & Page 2007; Grim, Rosenberger, Anderson, Rosenfeld, and Eason, 2011; Rescher 2011, 2012).

Even waiving those interpretational concerns in the name of model simplicity, however, we faced an issue regarding trust updating that had to be resolved in order to build the model at all. If I do discount information from those who hold views opposed to mine, precisely how much should our model discount those views? Should trust updating be modeled linearly, as in Figure 4a, or more like in Figure 4b? In the latter case, what precisely should our curve of trust-discounting look like?



Figures 4a and 4b. Two ways of graphing trust updating. In each case an agent increases trust as shown in an agent with a belief less than τ in distance from his own, and decreases trust as shown in an agent with a belief greater than τ from his.

In both cases, τ is the distance from an agent's belief at which there is a shift from increased trust to decreased trust. Call that trust watershed the τ-point. What should the τ-point be in our model for trust updating? Moreover, what field of comparison should we use for such a calculation? Should we increase and decrease trust on a *local* scale, with the scope of our trust updating calibrated to each individual's immediate contacts? That would mean that our individuals discount the beliefs of those *among their network contacts* most distant from them. Or should we discount on a global scale, in the sense that an individual distrusts those who would be most distant from him *across the full field of beliefs*, whether or not he has immediate contact with agents widely differing in belief?

IV. Exploring the Impact of Alternatives

If we were to wait for psychologists to tell us how people update trust in terms of belief differences, whether in accord with Fig. 4a or 4b and whether against a local or global standard of comparison, we would have a long wait indeed. The truth is undoubtedly that trust updating does not occur in terms of single beliefs, is not solely in terms of belief distance, and varies in terms of update function and background comparison depending on the people and the issue involved.

That means that a predictive model of precisely what the belief dynamics will be in a particular community and a particular case is beyond us, and perhaps beyond social science generally. But prediction is not the only purpose behind computational modeling, and perhaps not the primary purpose. Explanation of general phenomena through an understanding of general mechanisms is of value even where point prediction is possible—and may indeed tell us that there will be many cases in which point prediction is *not* possible. Understanding potential dynamics in a range of cases can be as important, or even more important, than offering a specific prediction in a particular case. Understanding what factors can be expected to carry particular weight, individually or in combination, can be as important as any specific prediction based on a specific set of values for those factors.

As modelers, therefore, an alternative course of action is entirely appropriate. Our goal need not be to build some single set of realistic psychological assumptions into some specifically predictive model. What psychological assumptions are realistic may vary from person to person, from belief topic to belief topic, from community to community, and from case to case. In the attempt to understand belief dynamics in general, it is entirely appropriate to ask what the impact of alternative assumptions regarding trust will be for belief dynamics across a community and for belief polarization, for example. In that case, we are not attempting to peg the 'right' value of potential factors for any particular case. In that case we are attempting to figure out the relative importance of those potential factors across a range of cases, real, hypothetical, and counterfactual.

For purposes of point prediction, the level of abstraction at which we are building computational models would be a detriment; the variations in variables we are considering would simply represent a confession of ignorance. For purposes of a more general understanding of a phenomenon, the level of abstraction of models like ours can be a positive gain. With the abstract unreality of distance from the specifics that would be required for prediction in a specific case comes the power of generality. Aspects of dynamics observable in a wide range of general abstract models will be good candidates for aspects of dynamics that will hold across not just one but a range of specifiable cases. We can come to know where results change with changes in our variables.

Without being able to answer some of the questions our initial models raised, we began to make models with which we could explore what happened on some of the various options available. In some of the models we were building at this stage, polarization still refused to appear. But the scale on which trust updating was applied—the scale on which beliefs were discounted—did seem to make an important difference.

Figure 5 shows a typical evolution of beliefs in a network that starts with a random connection between agents of different beliefs and in which trust in other agents is discounted in terms of belief distance on a global scale. This is the evolution of beliefs in a community in which agents discount those far from their own beliefs, but far from their own beliefs in terms of the entire spectrum of opinion in the community. The result is convergence.

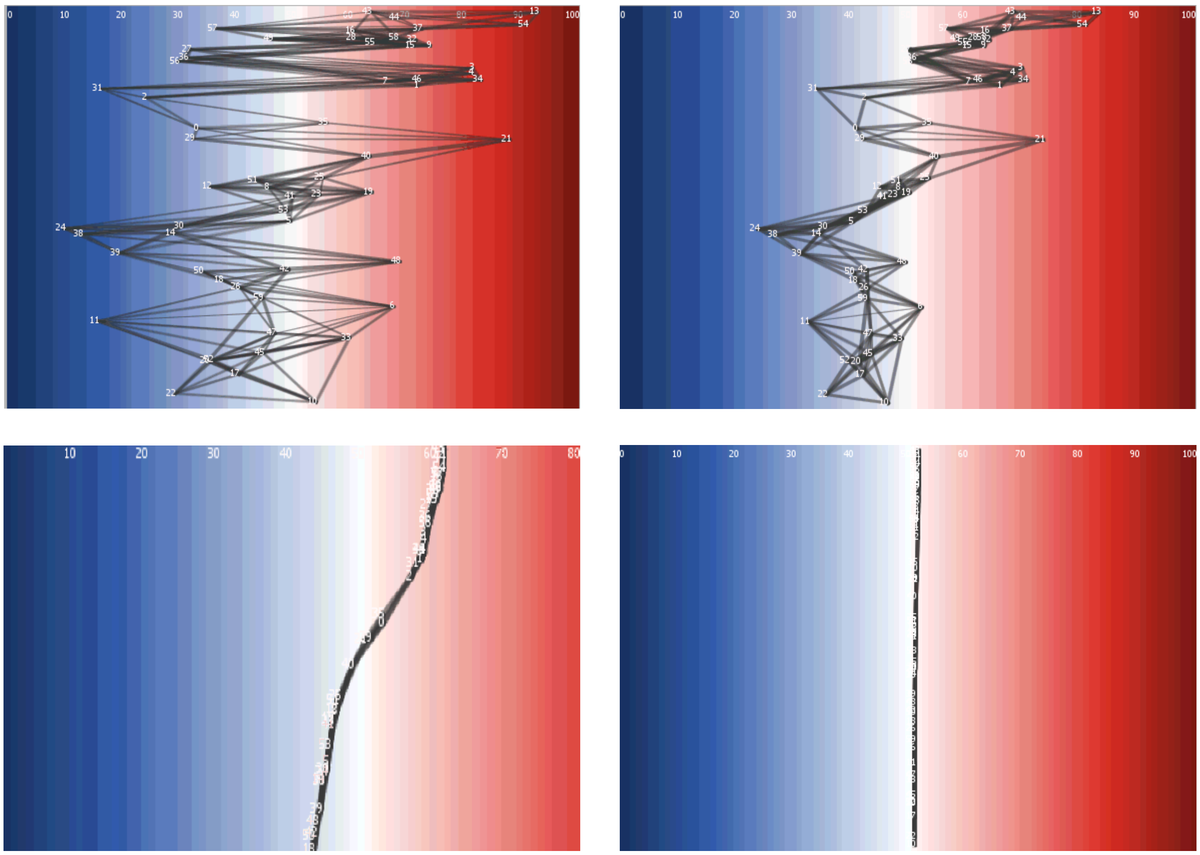


Fig. 5 Horizontal location represents belief. Snapshots show a typical evolution of random network with global trust updating. Generations 5, 15, 25 and 30 shown.

Figure 6, in contrast, shows a typical evolution of beliefs in a similar random network but in which trust is discounted in terms of belief distance on a local scale. This is the evolution of beliefs in a community in which agents discount those far from their own beliefs in their own network of immediate contacts. The result starts to look more like polarization.

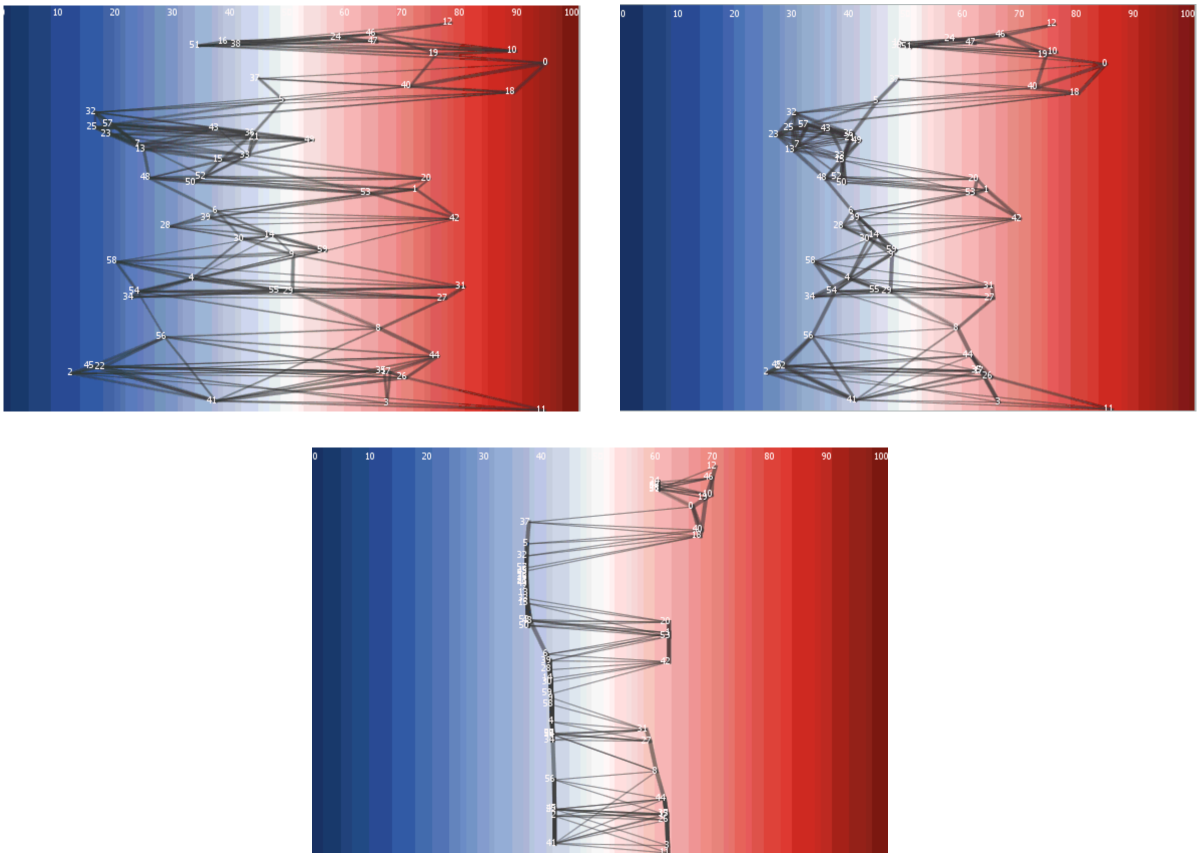


Fig. 6 5 Horizontal location represents belief. Snapshots show a typical evolution of random network with local trust updating. Generations 5, 15, and 30 shown.

V. Philosophical Analysis in Computational Modeling: The Case of Polarization

At this point, we had the essentials of a more promising model. With networks of agents, belief updating by weighted averaging, and a range of possibilities for trust updating, we could start to measure various factors and their influence on polarization. What difference to polarization does the type of network make—a random network of connections, for example, or a scale-free network more like many real social networks? What difference to polarization does the shape of trust-updating make? We are currently working with the linear graph because it's easier to handle. But even given that shape, what difference does a shift in τ make? What polarization difference does it make if I discount those .5 distant from my current view, .4 distant, or .3 distant?

The exploration of those parameters form the core of our work in progress. That work is currently qualitative, eyeballing the belief distributions that those parameter differences make, just as we invited you to eyeball them in the figures above.

What we would like in the end, however, is something more: a quantitative take on questions of belief dynamics, network structure, media effects, and the issue at hand. Within a range of abstract model assumptions, we'd like to know just how much each of these factors can be seen to contribute to polarization. For that we need a quantitative measure of polarization. But, there *another* conceptual difficulty arose.

As indicated in introduction, we started with the impression that the increasing polarization of America was an agreed and established sociological fact. Everybody talks about it, a range of books are written about it, so it must be real, we thought.

Has polarization in American increased? What exactly do people mean when they talk of polarization? Is there just one thing they mean, or are there various senses of the term? How are we to measure them? If you try to build a model, however simple, in which you measure polarization, that kind of abstract conceptual question becomes immediate and pressing.

A major task we have faced is simply to tease out different senses of 'polarization' which appear at various points in literature of sociology and political science but which are not clearly distinguished in that literature. Often entire articles appear on the topic of polarization, but with little attempt to make it clear what precisely is meant by the term. A real understanding of the phenomena at issue demands that we do better. The methodology of computational modeling strengthens that demand.

Without claim to completeness, the following is a brief catalog of senses of the term in the literature that we have found it necessary to distinguish, and which we intend to pursue in quantitative form in further modeling:

Polarization type 1: Spread

Polarization is measured in terms of the range of opinions. One might therefore ask: How far apart are the extremes? In one of the best sociological pieces on the issue, DiMaggio, Evans, and Bryson 1996 call this "dispersion": "The event that opinions are diverse, 'far apart' in content." They also outline a dispersion principle: "Other things being equal, the more dispersed opinion becomes, the more difficult it will be for the political system to establish and maintain centrist political consensus" (694).

In our model, we can measure polarization in the sense of spread as the belief level of the agent with the highest belief value minus the belief level of the agent with the lowest belief value. Polarization in this sense, however, does not consider whether the agents with minimum and maximum beliefs are extreme case outliers or the edges of large clusters. Spread is also independent of any measure in terms of groups; even if the minimum and maximum agents are representative of groups at the ends, the measure will ignore any groups in between. Although polarization in the sense of spread is important, it is also clear that we will want to measure other aspects of the phenomenon as well.

Polarization type 2: Distinctness

If we can identify different belief or attitude groups—clusters along a scale, for example—how distinct are these factions? Unlike polarization in the sense of spread, polarization in the sense of distinctness is a measure explicitly defined in terms of groups. What matters here is how clearly distinct those groups are, regardless of the distance between them. DiMaggio and his co-authors call this 'bimodality.' People are polarized in this second sense "insofar as people with different positions on an issue cluster into separate camps, with locations between the two modal positions sparsely occupied" (DiMaggio, Evans & Bryson 1996, 694).

One way to measure distinctness would be to rank the groups in order of their mean belief values and then perform pair-wise comparisons of the distributions using the Kolgomorov-Smirnov (KS) two-sample test (Kaner, Mohanty & Lyons 1980; Wilcox 1997). This nonparametric method examines two sets of data and determines the probability that they were drawn from the same distribution (without making any assumptions about what those distributions might be). The resulting p-values for their being separate distributions act as measures for how distinct the groups' beliefs are. A related N-sample test or Bayesian method can extend that approach for any number of groups.

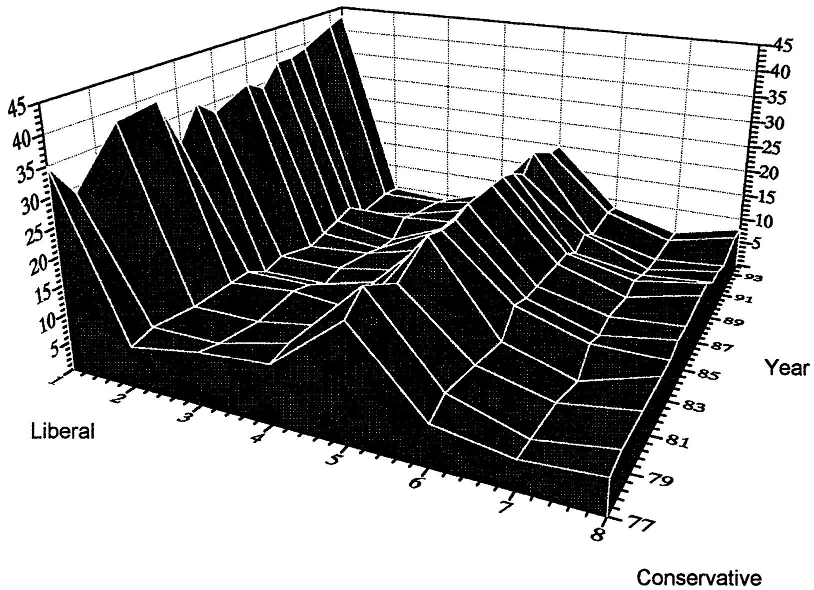


Fig. 7 Attitudes toward abortion, distribution by year, from the full sample General Social Survey 1997-1994. (DiMaggio, Evans & Bryson 1996, p. 709)

There is no necessary connection between polarization in sense 1 and 2; between spread and distinctness. A population might have a very diverse set of views on an issue without particular clusters emerging around any particular view. But there is no necessary disconnection, either. Attitudes toward abortion between 1970 and 1990 show both a great spread and distinctness, for instance (Fig. 7). In their words, "If attitude polarization entails increased variance, increased bimodality, and increased opinion constraint, then only attitudes towards abortion [amongst those considered in the article] have come more polarized in the past twenty years, both in the public at large and within most subgroups" (DiMaggio, Evans & Bryson 1996, 738). "No issues represents contemporary social conflict as vividly as does abortion, the struggle over which has become symbolic of the so-called culture wards (Hunter 1994)… Americans have become more divided in their attitudes towards abortion and, less dramatically, in their feelings toward the poor. The fact that division on these latter issues has increased without large directional change in central tendencies confirms the importance of inspecting change in distributions as well as in means" (DiMaggio, Evans & Bryson 1996, 715).

In other sociological work, Bartels 2000 argues that voting behavior shows increased distinctness between political groups since the 1950s. Bartels demonstrates that party identification increased sharply in the 1990s, with both strong and weak identifiers increasing along with a corresponding down-tick in the number of voters that identify as independents (Bartels 2000, 36-7). The trend identifies a growing distinctness of the political parties along with the diminishment of independent, non-affiliated voters in the middle. The impact of distinctness on presidential and congressional races has been greater than at any time since the mid-sixties (Bartels 2000, 42).

Polarization type 3: Uniformity within Groups

How diverse are opinions *within* each group? In contrast to distinctness, this measure looks at uniformity within, rather than between, groups. The more single-minded or unanimous views are within distinct groups, the greater this sense of polarization between them. A suggestive measure is absolute deviation. The smaller the variance within distinct groups, the greater this sense of polarization across the population.

Increased uniformity as a measure of polarization is clear in the Congressional voting records of the major parties. Between 1969 and 1976—the Nixon and Ford years—the rate at which Republicans voted along party lines was about 65% in both the House and the Senate. The same was true of Democrats. Between 2001 and 2004, under George W. Bush, Republicans voted with their party 90% of the time. Democrats voted with their party 85% of the time (McCarthy, Poole, & Rosenthal 2006).

Baldassarri and Gelman (2000) also find increasing party polarization. They write, "Looking separately at trends among Republican and Democratic voters … we find clear evidence of increasing constraint within issue domains, especially among Republicans. In fact, Republicans have become more consistent on economic and civil rights issues, while Democrats have lost constraint on these issues and become a bit more coherent in their moral views. In both groups of voters, the constraint is growing faster than in the populace as a whole" (p. 436). On numerous accounts, the Democratic and Republican parties have become more internally uniform.

Polarization type 4: Size disparity

A society that has one dominant opinion group with a few small minority outliers seems less polarized than one with a small number of comparably sized competing groups. Groups are more polarized in this sense if the different beliefs are held by equal numbers of people. . Using the notation that G is the set of groups, and γ_i is the size of group i, size disparities can be measured by calculating the absolute deviation: $1/(2N) \times \Sigma \mid \gamma_i - \mu_G \mid$. This is just the normalized sum of distances from the mean community size; it equals zero when all the groups are the same size and increases the more groups differ from the mean size. It maxes out at 1 as the number of groups and size differences go to infinity, making it a nice measure for comparison across different configurations.

Views on women's role in public life are no longer as polarized in this sense as they once were, even there are small groups who continue hold anti-feminist views that were once much more common. In the past, major portions of the population once fought racial integration vociferously. Even if the views represented there are still held by some, polarization on the issue of racial integration has clearly decreased.

Polarization type 5: Coverage

We think of polarized societies as having a few tightly packed sets of beliefs. The inverse of this, a broad spectrum of beliefs, can be captured in a variety of ways. One example is the proportion of the belief spectrum held by members of society. The larger the areas of unoccupied belief space, the more polarized the society. The more focused and less diverse the beliefs in a society are, the more polarized it is.

A simple way to envisage the measure in a discrete instantiation is think of the spectrum of possible beliefs between 0 and 1 as divided into small bins of size d (e.g., $d = 0.01$ or normalized by setting d to 1 / the number of agents). We can then measure coverage in terms of the proportion of bins filled. Alternatively, we might want a continuous measure over the belief space. This can be done by summing the amount covered by d-diameter halos around each agent; i.e., any portion of the belief space that is within d of an agent is considered covered; the rest is uncovered.

Polarization in the sense of coverage is related to dispersion, but does not include the shape of the belief dispersion. We might therefore think of coverage as a sub-measure of global dispersion, measuring how much dispersion there is without measuring its location.

Polarization type 6: Regionalization

While polarization in the sense of coverage represents how much belief dispersion there is without accounting for where beliefs are dispersed, we might also want to measure certain aspects of belief regionalization without attending to the belief area covered over all. In considering small bins of possible belief, for example, we might mean by polarization not how few bins are filled but the extent to which there are regions of empty bins *between* regions of bins that *are* occupied.

With 100 bins, for example, there might be three different cases: (a) that in which bins 0-50 are the only bins filled, (c) the situation in which bins 0-25 and 30-55 are filled, and (c) the situation in which 5-bin regions are filled, separated by 5-bin holes: regions 0-5, 10-15, 20-25, 30-35… are the only ones filled. Each of these will be equally polarized in the sense of polarization as coverage. Counting the number of empty regions between regions of occupied spaces, however, gives us a measure of polarization in which (c) is more polarized than (b), which is in turn more polarized than (a). Regionalization seems a further intuitive sense of polarization well worth quantifying.

It should be noted that regionalization per se does not distinguish between the case in which (b) bins 0-25 and 30-55 are filled, and (d) that in which 0-25 and 75-100 are filled. In terms of regionalization that may be exactly what we want: beliefs in the two cases are regionalized in precisely the same sense, though the groups are farther apart in the sense of spread.

Senses 1 through 6 of polarization can all be seen in terms of histograms of beliefs on a single issue across a population. But there are other senses of the term that are essentially (a) multiple-opinion or (b) network-based.

Polarization type 7: Multiple opinion convergence

Given polarized groups on issue A, are these same groups polarized on B, C, and D? The more interlocked rival beliefs are within rival groups, the greater the polarization across the community. Fiorina and Abrams 2008 note that intra-group polarization in this sense may increase even though population distributions on particular issues may not change. Bishop 2008 notes that individuals may move to "neighborhoods where others have similar political views, changing their partisan identifications to match their ideological and issue positions" (578).

Polarization type 8: Community fracturing

Sub-communities may be polarized simply in the sense that there is little or no communication between them. Even if two separated communities have identical and uniform beliefs, that uniformity may be coincidental and temporary.

In *Ethnic Conflict and Civil Life*, Varshney 2002 demonstrates how group interactions ameliorate levels of inter-group violence, and conversely, how group isolation increases the likelihood of violence. Varshney's central claim is that "pre-existing local networks of civic engagement between two communities stand out as the single most important proximate cause" for the difference between peace and violence (9). Put another way, cities with social networks that connected Hindus and Muslims through the same institutions were much less likely to see outbreaks of ethnic violence than cities in which Hindus and Muslims belonged to distinct civic institutions.

VI. First Results and Work in Progress

We think we have made progress, along the lines above, in the conceptual foundations necessary to model building with an eye to understanding polarization. Simple assumptions of a single belief scale and belief updating will remain, but with a range of variability to be explored in (a) trust updating functions with (b) different τ values against (c) local and global scales, with further variations in (d) social network structures and sizes, (e) initial configurations, and (f) media sources and effects. Our measures in exploring variations in those parameters will be measures of polarization in the distinct conceptual senses outlined above.

The following is a sample of the kinds of results we're headed for.

Begin with a random network of 50 agents, initially assigned beliefs between 0 and 1. Begin with a simple linear function for belief updating. That function is 'tune-able': it may be when a contact is within .2 of an agent's belief that his trust in that contact increases, and beyond .2 that he begins to discount input from that source. Or that τ-point may be wider: it may be a distance of .3 that marks the difference, or any other number.

Consider now two variations. In one, the τ-point is marked on a scale calibrated to the entire spread of beliefs across the population. In that case the belief spread of my particular contacts may not be as important. Relative to the range of opinions across the population, all of my friends may think pretty much like me. We will have a mutual opinion admiration society, increasing trust in each other and influence on each other based on trust. This first variation is a 'global' updating model. I tend to trust individuals with beliefs like mine, gauged against the whole spread of public opinion.

Consider a second variation that differs only in the scale on which trust updating is measured. In this case τ-points .2, .3, .4 aren't measures across the whole spread of beliefs within the population at large. They are measures across just the spread of beliefs of my immediate contacts. In this case it will be guaranteed that one of my contacts is the farthest out—and I will decrease trust in that individual no matter how close our beliefs on the 'objective' scale of the entire spread within the population. This is a 'local' updating model. I trust those among my contacts with beliefs like mine, gauged against the field of opinion among those with whom I am in contact.

Given the other particulars of the model—a random network of 50 agents and a linear updating function—that difference between global and local scaling makes a major difference in the emergence of polarization, in several senses.

Figure 8 shows a sample of what happens with a τ point of .25 and global updating. Figure 9 shows by contrast what happens with a τ point of .25 and local updating. More complete animations for each are available at www.pgrim.org/workinprogress.

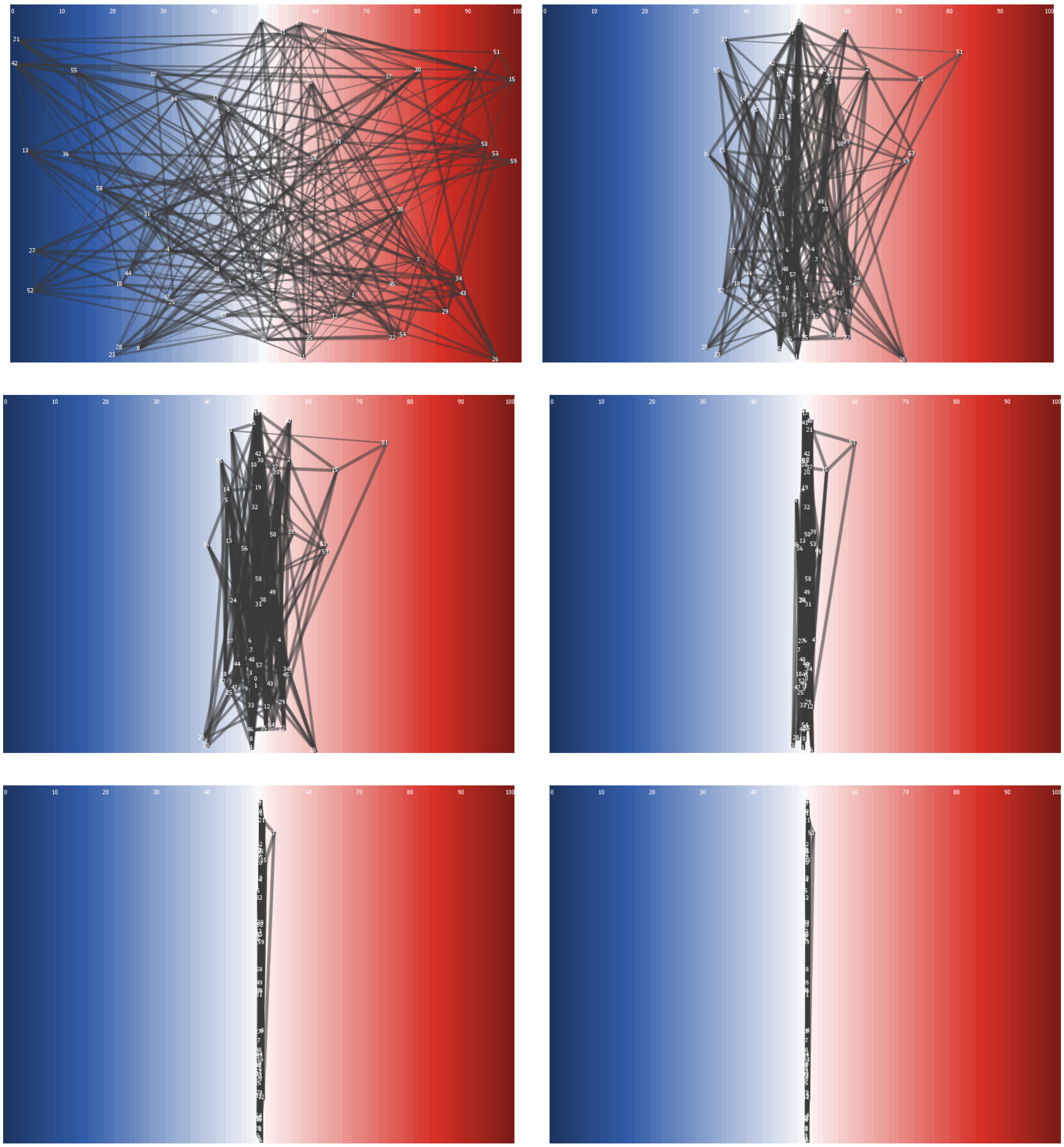


Fig. 8 Horizontal location represents belief. Representative slides from evolution of a random array with a τ point of .25 and global updating. Agents update trust positively in those closest to their beliefs, update trust negatively in those farthest away, with a transition point from positive to negative update at t = .25. See also www.pgrim.org/workinprogress.

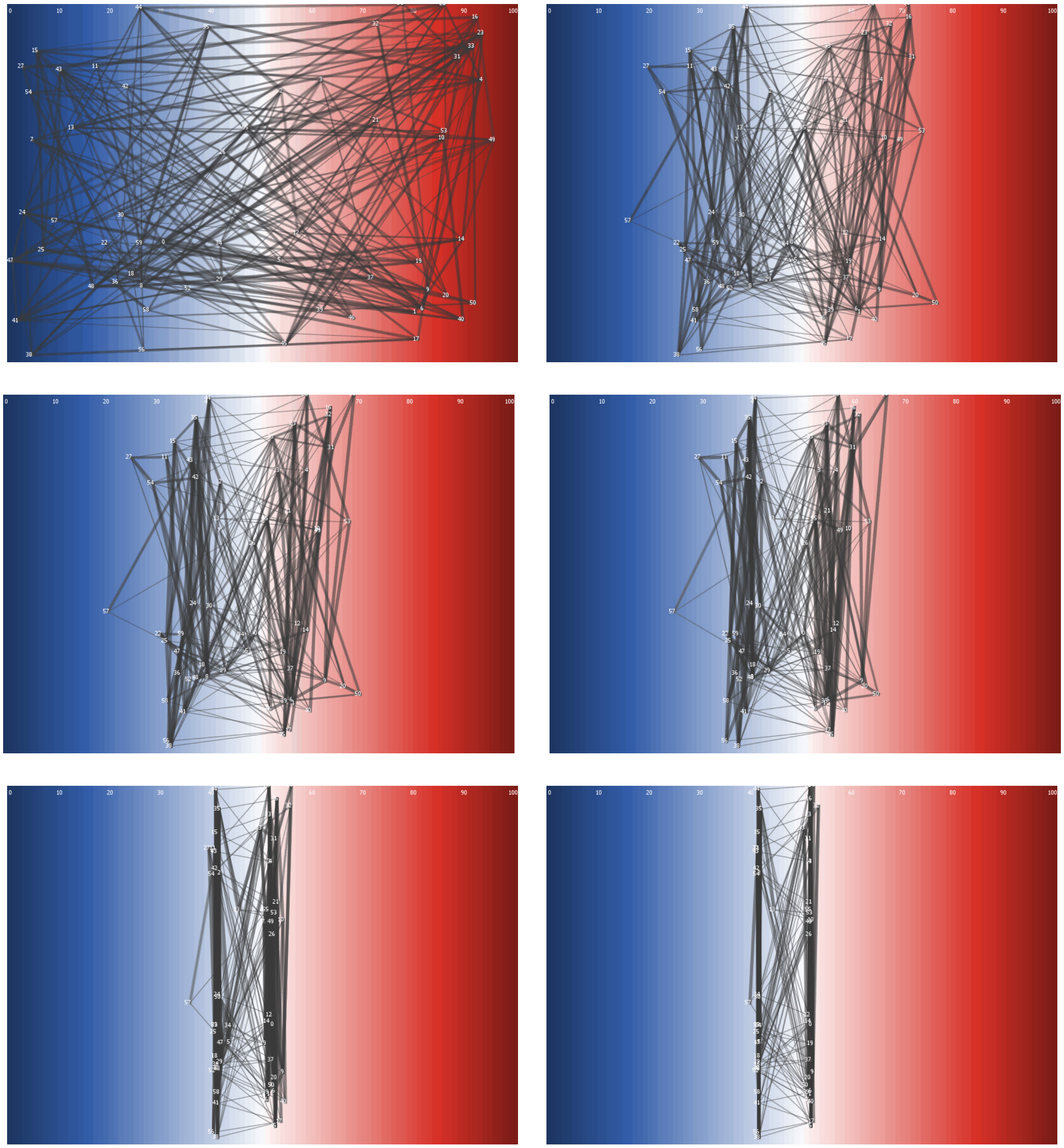
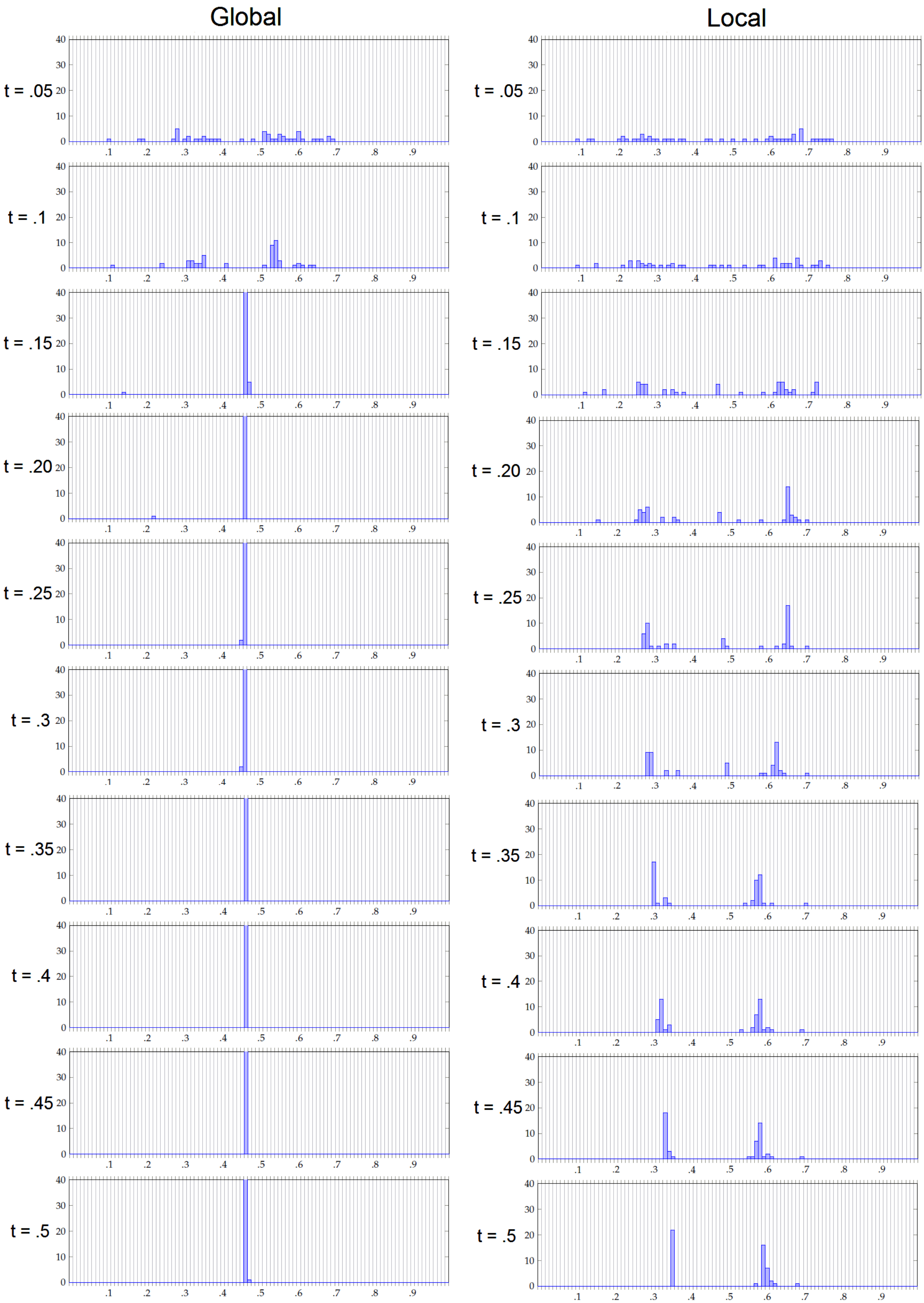


Fig. 9 Horizontal location represents belief. Representative slides from evolution of a random array with a τ point of .25 and local updating. Agents update trust positively in those closest to their beliefs, update trust negatively in those farthest away, with a transition point from positive to negative update at t = .25. See also www.pgrim.org/workinprogress.

Figure 10 shows results side by side for different τ points from .05 to .75 with the same initial random seed, so that the initial beliefs in the community are the same. On the left are results for global updating. On the right are results for local updating. Global updating, it turns out, goes to belief convergence with even a very small τ value. Local updating produces polarization all the way up to a τ value of .5.


Global
Local
t = .05
t = .1
t = .15
t = .20
t = .25
t = .3
t = .35
t = .4
t = .45
t = .5

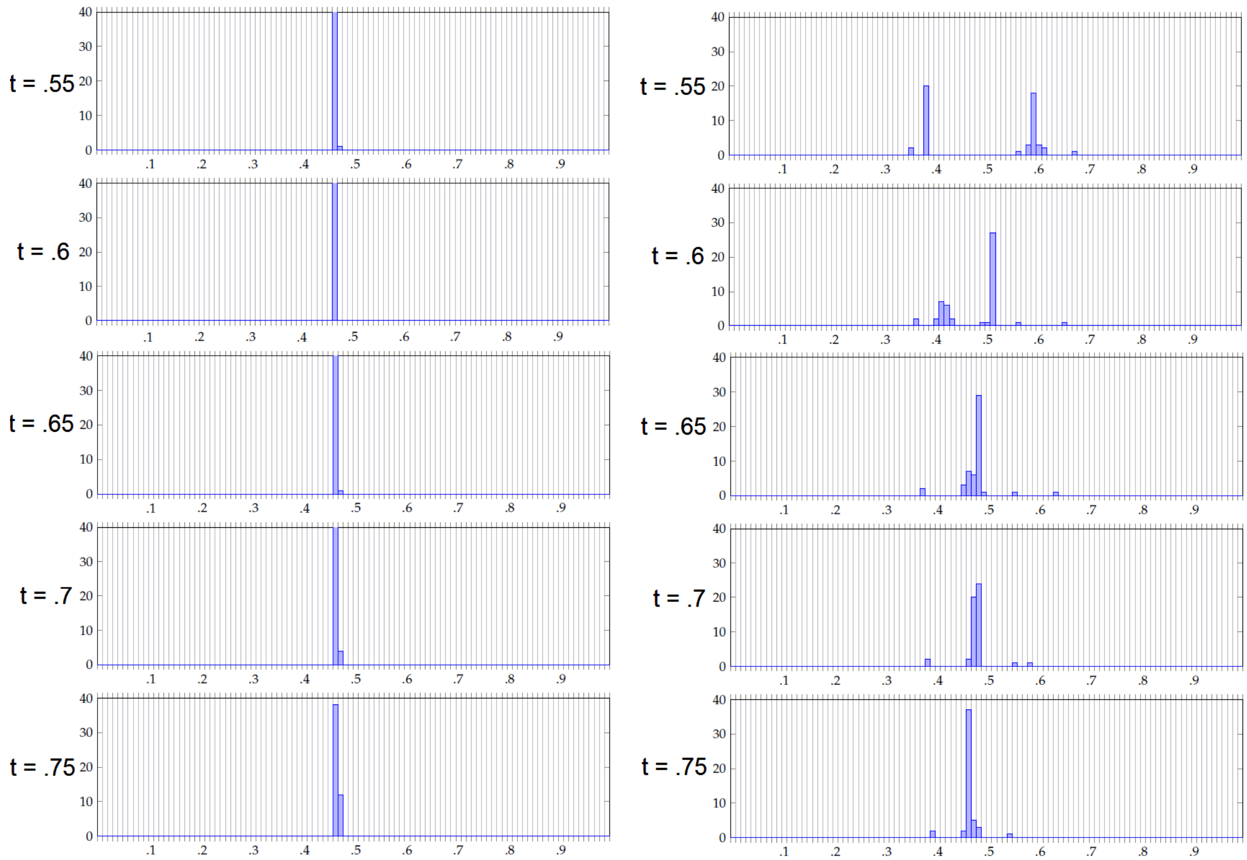


Fig. 10 Results for global (left) and local (right) scaling with the same trust updating function (as in Fig 4a) and different τ points from .05 to .75, using the same initial random seed throughout.

What these initial results indicate is that in looking for factors that favor polarization, local versus global updating can play a major role.

Note also that we can distinguish many different types of polarization mentioned above in these images. In the image for local updating with a τ of .5, polarization is high in a number of senses. We have two major groups and a smaller intermediate group that are clearly distinct—polarization sense 2. They vary in how sharply peaked they are—polarization sense 3. The two major units are fairly equal in size, at least in this run—polarization sense 4. If network links are broken when trust falls below a certain level, it's a good guess that the networks at issue are fractured in polarization sense 6.

It is worth emphasizing that those senses of polarization are conceptually distinct. There is nothing that says logically or conceptually that polarization in one sense need accompany polarization in others. As the work progresses, it will be interesting to see whether some of these senses nonetheless appear together in modeled network dynamics much as they often seem to go together in the social dynamics that are our ultimate target.

Note also how patterns of polarization change in trust updating on a local scaling with increases in the τ point. Consider for example the patterns of polarization with a τ points at .2, at .3, at .4, and so on. Several senses of polarization stay the same at those points. Distinctness does—polarization sense 2. Sharpness of peak on each side stays about the same—polarization sense 3. The major units remain comparative in size—polarization sense 4. The sense of polarization that changes with increasing τ is polarization sense 1—the distance of the extremes. With increasing τ points the objective position of the two groups comes closer together. Polarization in sense 1 *slowly* decreases. In the other senses it remains fairly uniform, without decrease, until the two groups actually meet. Polarization in the other senses disappears in this progression only when polarization in sense 1 does, and only because polarization in sense 1 does.

VII. Conclusion

All the work offered here is work in progress, with just a tease of initial results. We have found that polarization in all the senses outlined is a complex phenomenon, sensitive to initial conditions. Global trust updating uniformly gives us consensus. Local updating clearly does not, but the clarity, extent, and patterns of polarization differ widely across runs. In a random network of 50 agents, with a linear trust update, local and global scaling mark a major difference. But other factors are of importance as well. We know that the shift from a random to a scale-free network gives a different picture—one in which that difference between local and global scaling is not so pronounced. Even population size will be important.

It is a better appreciation for the role of different factors in the network dynamics, not of polarization, but of polarization*s* that is the wider area we want to explore.

What we have tried to indicate here is that an exploration of this form, though computationally instantiated, remains in large part conceptual in the sense that philosophical analysis has always been conceptual. We want our final results to be scientifically grounded. We hope they may offer some genuine social understanding. But in order to fill those goals they must also be philosophically sound, with a clear conceptual base.

We have also tried to make it clear that exploration of this kind often involves demands and openness to and opportunistic exploitation of the unexpected. We encounter conceptual problems we didn't anticipate, which force us to distinctions and tools we didn't have in advance, which lead us to build different models than we initially envisaged, which promise unanticipated results. We hope those results will tell us something genuinely new about the real social polarization we want to understand.

In the end, of course, when this is more than work in progress, we will write up our results in standard scientific fashion. We will make it look like we knew what we were doing all along, step by step, using a well-motivated methodology from a clear initial plan that produces a compelling compilation of results toward a tidy conclusion. In that final report, the crucial role of philosophical analysis in computational modeling may also go unmentioned.

Acknowledgments

We are grateful for comments on an earlier version of the paper presented at the Human Complexity 2012 conference at the University of North Carolina, Charlotte. That conference grew in turn from an NEH Institute for Advanced Topics in the Digital Humanities: Computer Simulations in the Humanities, hosted at the University of North Carolina in the summer of 2011. Research supported in part under a MIDAS grant NIH 1U54GM088491-01, "Computational Models of Infections Disease Threats," administered through the Graduate School of Public Health at the University of Pittsburgh.

References

Bartels, L. M. (2000). “Partisanship and Voting Behavior, 1952-1996.” *American Journal of Political Science* 44, 35-50.

Beaney, M. (2009). “Analysis.” *Stanford Encyclopedia of Philosophy* http://plato.stanford.edu/entries/analysis/

Bishop, B. (2008). *The Great Sort: Why the Clustering of Like-Minded America is Tearing Us Apart*, New York: Houghton Mifflin.

Brownstein, R. (2007). *The Second Civil War: How Extreme Partisanship has Paralyzed Washington and Polarized America*. New York: Penguin.

Burkholder, L. (1992). *Philosophy and the Computer*. Oxford: Westview Press.

Bynum, T. & J. Moor (1998). Eds, *The Digital Phoenix: How Computers Are Changing Philosophy*. Oxford: Blackwell.

DiMaggio, P., J. Evans, & B. Bryson (1996). “Have American's Social Attitudes Become More Polarized?” *American Journal of Sociology* 102, 690-755.

Fiorina, M. P., & S. J. Abrams (2008). ‘‘Political Polarization in the American Public.’’ In *Annual Review of Political Science 11*, ed. M Levi. Palo Alto: Annual Reviews.

Fiorina, M. P., S. J. Abrams, & J. C. Pope (2010). *Culture Wars? The Myth of a Polarized America*. New York: Longmans.

French, J. (1956). “A Formal Theory of Social Power,” *Psychological Review* 63, 191-194.

Golub, B., & M. O. Jackson (2010). “Naïve Learning in Social Networks: Convergence, Influence, and the Wisdom of Crowds,” *American Economic Journal: Microeconomics*, 2, 112-149.

Golub, B., & M. O. Jackson (forthcoming). “How homophily affects learning and diffusion in networks.” http://arxiv.org/pdf/0811.4013v2.pdf

Grim, P., C. Reade, D. J. Singer, S. Fisher, and S. Majewicz (2010). "What You Believe Travels Differently: Information and Infection Dynamics Across Sub-Networks," *Connections* 30, 50-63.

Grim, P., D. J. Singer, C. Reade, & S. Fisher (2011). "Information Dynamics across Linked Sub-Networks: Germs, Genes, and Memes" *Proceedings, AAAI Fall Symposium on Complex Adaptive Systems: Energy, Information and Intelligence*, FS-11-03, AAAI Press.

Grim, P., D. J. Singer, C. Reade & S. Fisher (2012). "Germs, Genes, and Memes: Function and Fitness Dynamics on Information Networks,” *ALife 13 Proceedings*, MIT Press.

Grim, P., S. B. Thomas, S. Fisher, C. Reade, D. J. Singer, M. A. Garza, C. S. Fryer, & J. Chatman (2012a). "Belief Polarization within the Black and White Communities: Information Dynamics in Data-based Networks MIDAS scientific presentations, MIDAS Network meeting, Boston, February 2012.

Grim, P., S. B. Thomas, S. Fisher, C. Reade, D. J. Singer, M. A. Garza, C. S. Fryer, & J. Chatman (2012b). "Polarization and Belief Dynamics in the Black and White Communities: An Agent-Based Network Model from the Data." *ALife 13 Proceedings*, MIT Press, 2012.

Grim, P., R. Rosenberger, B. Anderson, A. Rosenfeld, and R. E. Eason (2011). "How Simulations Fail" *Synthese*, forthcoming, advanced online version Springerlink Synthese 2011.

Grim, P., G. Mar, & P. St. Denis (1998). *The Philosophical Computer: Explorations in Philosophical Computer Modeling*. Cambridge MA: MIT Press.

Hanna, R. (2000). "Conceptual Analysis." *Concise Routledge Encyclopedia of Philosophy*. London: Routledge, p. 106.

Hetherington, M. & J. Weiler (2009). *Authoritarianism and Polarization in American Politics*. New York: Cambridge University Press.

Holyoak, K. & P. Thagard (1997). "The analogical mind." *American Psychologist* 52, 35-44.

Hunter, James Davison. (1987). *Evangelism: The Coming Generation.* Chicago: University of Chicago Press.

Kaner, H. C., S. G. Mohanty & J. C. Lyons (1980). "Critical Values of the Kolmogorov-Smirnov One-sample Tests." *Psychological Bulletin* 88, 498-501.

Lord, C. G., Ross, L., and Lepper, M. R. (1979). "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology* 37, 2098-2109.

McCarthy, N., K. T. Poole, & H. Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press.

Miller, J. H. & S. E. Page (2007). *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, Princeton, NJ: Princeton University Press.

Pollock, J. (1998). *How to Build a Person: A Prolegomenon*. Cambridge, MA: MIT Press.

Rescher, N. (2011, 2012). "How Modeling Can Go Wrong: Some Cautions and Caveats on the Use of Models," Epistemology of Modeling and Simulation, University of Pittsburgh, April 2011, and forthcoming, *Philosophy and Technology*.

Sandin, P. (2006). "Has Psychology Debunked Conceptual Analysis?" *Metaphilosophy* 37, 26-33.

Varshney, A. (2002). *Ethnic Conflict and Civil Life: Hindus and Muslims in India*. New Haven, CT: Yale Univ. Press.

Visser, P. S., & Cooper, J. (2003). "Attitude Change." In M. Hogg and J. Cooper, eds., *Sage Handbook of Social Psychology*. London: Sage Publications

Wilcox, R. R. (1997). "Some Practical Reasons for Reconsidering the Kolmogorov- Smirnov Test." *British Journal of Mathematical and Statistical Psychology* 50, 9-20.