[This is a pre-submission draft. The correct citation (final version) is: Chattoe-Brown, Edmund (2014) 'Using Agent Based Modelling to Integrate Data on Attitude Change', *Sociological Research Online*, **19**(1), article 16, February. doi:0.5153/sro.3315]

Using Agent Based Modelling to Integrate Data on Attitude Change

Edmund Chattoe-Brown

School of Media, Communication and Sociology

Bankfield House,

132, New Walk, Leicester, LE1 7JA, UK

ecb18@le.ac.uk

Abstract

This article has two goals. Firstly, it shows how a relatively novel technique (Agent Based Modelling, hereafter ABM) can integrate different data types that are often used only in separate strands of research (interviews, experiments and surveys). It does this by comparing a well-known ABM of attitude dynamics with an alternative model using data from surveys and experiments. Secondly, the article explains ABM methodology and why it is important to the distinctiveness of ABM as a research method. In particular, the ramifications of differing approaches to ABM calibration and validation are discussed using the two different ABM as examples. The article concludes by showing how ABM might provide a progressive research strategy for integrating different data types and thus different disciplines in attitude research.

Keywords: Attitude Change, Agent-Based Modelling, Social Influence, Relative Agreement Model, Deffuant, NetLogo, Mixed Methods.

Introduction

It is a commonplace in sociology that research tends to be conducted either qualitatively or quantitatively and that each approach finds it hard to engage directly with the other.[i] This article demonstrates why and how ABM might be able to integrate qualitative and quantitative data into unified theories. To do this, however, it is necessary to introduce the relatively unfamiliar ABM approach and explain its distinctive methodology. As a case study to structure the discussion the article compares two attitude dynamics ABM. The first is the well-known Relative Agreement Interaction (hereafter RAI) model developed by Deffuant and others.[ii] The second ABM was developed to show improved fit with quantitative attitude data by incorporating other kinds of social science research (like experimental results on social influence). These examples also provide an opportunity to show concretely how ABM work and why the methodological issues raised bear on the actual conduct of effective ABM research. It is hoped that this example led approach will make the article both accessible and germane to readers with no previous ABM experience. The comparison of particular ABM also provides a starting point for the concluding discussion of ABM as a research strategy that can progressively integrate diverse forms of data and thus promote genuinely interdisciplinary research.[iii]

The structure of the article is as follows. The next section provides a basic introduction to ABM and its methodology. The following section illustrates points made in this introduction by presenting the RAI model as a typical example of a well-regarded ABM. The fourth section considers different possible relations between ABM and data and their implications. The fifth section looks at real attitude data from the British Social Attitude Survey and experimental data on social influence and considers its implications for the RAI model. The sixth section considers social processes whose absence from the RAI model might be expected to have a significant effect on its behaviour (particularly the role of the news media). The diversity of these processes also illustrates the need for a research method (like ABM) that can integrate diverse interdisciplinary data. The seventh section presents an ABM based on data from social influence experiments and incorporating a simple description of the role played by the news media, showing that this produces improved fit with the BSAS data. The concluding section sums up the role of data and ABM methodology in progressively bridging the gap between our (largely qualitative) knowledge of small-scale social interaction and our (frequently quantitative) knowledge of aggregate social attitudes.

A Very Brief Introduction to ABM and Its Methodology

The distinctiveness of ABM can best be presented in terms of two related aspects, both of which are effectively illustrated by contrast with methods of research already widely used in sociology (but also in other social sciences). Broadly speaking, quantitative research uses numbers in its data collection, analysis and theory building. To take a simple example, a regression analysis involves finding associations between numerical values collected using surveys where non-numerical data will be converted into a numerical form.[iv] Success in finding meaningful associations is also presented numerically (in terms of the size and sign of model parameters, significance tests, R-squared and so on.) To present a relevant example, we might have survey data on variables like gender, ethnicity and education and also on attitudes. Using simple regression, we might find that education had the most important effect on attitudes regarding abortion but was qualified (not surprisingly) by both gender and ethnicity.[v] By contrast, qualitative research operates on narratives or texts (interviews, documents and field notes) but also *argues* narratively from these to generalisations.[vi] For example, Siraj (2009) analyses interview data to show how Muslims use arguments drawing on theology and traditional gender roles to justify and maintain negative attitudes to homosexuality. On the purely descriptive level then, we might say that the difference between simulation and existing quantitative and qualitative approaches is that simulation involves representing accounts of social processes as computer programmes rather than equations or narratives. (For more detail on exactly what this entails – and further arguments supporting the claims made in this brief introduction generally – the reader is referred to the much less compressed exposition in Chattoe-Brown 2013.) However, matters are slightly more complicated than this and the implications turn out to be rather important. The distinctive contribution of ABM also relies on not confusing it with older and perhaps better-known simulation approaches. This is because these approaches do not really represent social processes in a distinctive way. Instead, they just *instantiate* existing kinds of theory differently. The most widely known example (which is also the easiest to explain) is probably System Dynamics (Forrester 1971). In this approach, a computer is used to establish the consequences of a set of dynamic equations.[vii] However, for social scientists sceptical that equations are adequately rich and flexible to represent human behaviour, a computer programme consisting of such equations is no more convincing as a theory than the same equations on paper. All that differs is how they are processed. Thus System Dynamics may resolve technical issues with establishing how systems of equations behave but it doesn't address the epistemological and methodological challenges of representing human behaviour in terms of equations in the first place.[viii]

This point leads us towards the second aspect of ABM distinctiveness. ABM do not simply translate existing theories (whether equation or narrative based) into computer programmes but start with the idea of representing social actors directly (rather than in terms of quantitative relationships or theoretical constructs) as they interact with each other and with their environment. Each agent is represented as a separate element of a computer programme that may have its own distinctive knowledge, point of view, thinking processes and capabilities. The simulated environment can respond to agents according to their properties or actions (a small agent can climb through a narrow window but a big one cannot, if you push at a rock it may fall on you) and agents can respond to each other on the basis of both their own mental processes and the properties and behaviour of others. (I am rude to my friends and polite to strangers. He is rude to strangers and polite to friends. She is polite to everyone except those who are rude to her.) The series of interactions between agents and between agents and the environment instantiated by an ABM corresponds to a relatively intuitive process based specification of sociality that is often used informally by social scientists. For example, first a job is advertised, then people may see the advert and apply (or hear about it through their social networks), then candidates are short listed, then they are interviewed, then an offer is made to the preferred candidate. (But only the least promising candidates may actually arrive for interview or the preferred candidate may not accept so the job may have to be re-advertised or it may be necessary for the employer to consider how far down the list of runner up candidates they are prepared to make a job offer.) To say that ABM represents social processes directly might seem philosophically and epistemologically contentious but what it means for my purposes is just that the computer programme represents relatively unproblematic (though not necessarily true) claims about social behaviour (the applicant considers job offers and selects the best based on wage, the forager wanders randomly looking for food) as opposed to representing theorised relationships – “suicide rate this year is associated in fixed proportions with suicide rate last year and average temperature this year" or theoretical constructs which may or may not be measurable – "individuals always act to reduce cognitive dissonance".[ix]

Because ABM directly represents social processes as sequences of interactions (rather than just instantiating existing theories) it gives rise to a distinctive methodology. When a regression analysis is performed, the object is to find the line that best fits the data according to statistical criteria. In this context, it makes little sense to distinguish between individual properties (the data) and aggregate ones (the slope of the line). The slope of the line just *is* the best summary and could not be otherwise for that data given the prevailing statistical criteria. By contrast, because an ABM represents a social process, it makes sense to ask separately: How do agents behave and what are the aggregate consequences of that behaviour?[x] For any particular specification of agent behaviour and environment, the dynamic process represented by the ABM gives rise to aggregate patterns (for example, the fraction of the population who report strongly opposing legal abortion at a particular time) but either the aggregate patterns or the individual behaviour (though obviously not both, linked as they are by the process specification) could *independently* and *meaningfully* be otherwise. (While it is perfectly possible to specify a regression line with arbitrary parameters not corresponding to those indicated by the data, what would be the point?) Real individuals might or might not make decisions on a rational basis and this might or might not have a large impact on the dynamics of the social system as a whole.[xi]

This two level specification (social behaviour by individuals in an environment and aggregate statistics – very loosely micro and macro) gives rise to different ways of using ABM. With no data at all, the operation of the ABM is rather like abstract mathematics: If we assume these things about individuals and their environment, then this is what happens in aggregate. Using data at either level, the ABM can also be used in an exploratory way with respect to the other level. (Under what social circumstances can hyperinflation occur? If everyone is strictly rational can aggregate cooperation be sustained in a Repeated Prisoner's Dilemma?) However, it is using data at *both* levels (as will be demonstrated in this article) that reveals the full distinctiveness of the ABM methodology and thus drives the argument that ABM can integrate different kinds of data. If empirically plausible individual behaviour is represented by an ABM and gives rise to simulated aggregate data which (in a sense that needs to be established) resembles real data, it is possible that we have not merely summarised or found an association in data but may really have *explained* it. The hypothesis is that the reason the aggregate simulated data looks like the real aggregate data is *because* the real social process unfolds in key ways *like* the simulated social process.[xii] Of course, there are countless issues of practice that stand between this logic and its realisation in real research. For example, how similar do real and simulated aggregate data have to be before it reasonable to suppose that the explanation instantiated in the ABM is genuine rather than coincidental? This and other important issues of practice will be discussed in concrete terms throughout the rest of this article using examples of attitude dynamics ABM but I hope the reader is provisionally prepared to accept (based on the arguments so far) that, despite significant challenges to its realisation, the ABM approach is nonetheless distinctive (both in the way it represents theories and in its methodology) and not simply a variant of either statistical model fitting or principled narrative persuasion.

A Well Known ABM of Attitude Dynamics: The RAI Model

In this section, I shall move the foregoing discussion from the abstract to the concrete by describing the RAI model.[xiii] In the following section, I shall consider its implications. As described above, the RAI model is a typical ABM in specifying a social process that links different levels (that of the individual in its environment and that of aggregate statistical properties like the proportion of the population who would report being strongly opposed to legal abortion.) In terms of attitude dynamics, the RAI represents both attitudes as attributes of individuals (responses they might give to a survey at any point if they were perfectly self aware and honest) and the effect of social interactions on those attitudes. If I am opposed to abortion being legal and discuss it with someone who thinks it should be, one or both of us may come away from the discussion with a changed attitude which would then be picked up if we were surveyed at some later point.

One reason for beginning with a discussion of the RAI model (apart from its popularity and acceptance in the ABM community) is that the process of social interaction specified in the original model is fairly simple. Each agent has an attribute standing for its attitude (represented as a continuous variable between +1 and -1) and another attribute standing for its uncertainty (represented as a symmetrical range around the attitude, for example -0.6 to -0.8 for an attitude of -0.7). The rationale for the uncertainty attribute is that some of our attitudes are better founded than others (compare my attitudes to the British government and to the government of Armenia) and this will have some bearing on how likely we are to be influenced by others, particularly if their attitudes seem better founded than ours. For example, it is not implausible that the attitude of a gynaecologist to abortion may have more influence on my attitude than that of a carpenter. In the RAI model, this rationale is translated into a simulated social process by saying that when two agents meet, one cannot influence another unless its attitude is within the uncertainty range of the agent to be influenced. The impact of both attitude and uncertainty in one agent on attitude and uncertainty of another is determined by equations whose exact specification need not concern us here (Deffuant *et al*. 2002 section 2.4) but a key aspect of social dynamics in the RAI model is that agents have the potential to change not only each other's attitudes, but also each other's uncertainties. In terms of broad intuition rather than exact detail, one agent in the RAI model has no effect on the attitude of another if their attitudes are too different and/or held with too great a certainty. Otherwise, one or both agents may move towards the other in attitude with consequences for their respective certainties. This is intuitive. It is unlikely that a discussion between a Catholic and a libertarian on legalising abortion would have any effect on either of them except irritation. On the other hand, a discussion between two libertarians might well move their attitudes closer together.

One part of the RAI model analysis presented by Deffuant *et al*. (2002) involves a slight variant of the basic model where some of the agents can influence others but cannot be influenced themselves. (Again, it seems intuitive that there may be at least some people like this in the real world.) If these agents are also assumed to begin with attitudes at the extremes of the possible range (i. e. +1 or -1), then two possible aggregate outcomes are observed. In one, moderate attitudes prevail (the extreme attitudes of the agents who can't be influenced have no lasting effect). In the other, the attitude of the population converges to one or both extreme positions. (Examples of each outcome can be found in sections 3.8 to 3.11 of Deffuant *et al*. 2002.) Obviously, the possibility of a few individuals with extreme and non-negotiable attitudes creating a highly polarised society (which could lead to irresolvable conflict) has potential resonance in the real world.

Having discussed the assumptions of the RAI model, the intuitions behind these and a potentially interesting result produced by the model, I consider its methodological implications in the next section as part of a wider analysis of current ABM practice and its potential drawbacks.

Implications of Different Relations between Data and ABM

As I have already suggested above, a number of relationships between different kinds of data (micro and macro) and ABM are possible. An ABM can make use of data at neither level (ABM as abstract mathematics), at one level (exploratory ABM looking at the macro consequences of micro assumptions or the possible micro causes of observed macro patterns) or at both levels. However, it should not necessarily be assumed that all these approaches are equally useful and unproblematic. While it is unhelpful to try and prescribe any particular use of ABM as correct (and it is certainly the case that models in all three classes are published), the logic of ABM methodology has a bearing on the value of different approaches to data use and this must be taken into account.

Taking the most straightforward case first, ABM as abstract mathematics might seem impossible to criticise (any more that it would make sense to criticise number theory). However, on closer examination, the parallel is inexact. Pure mathematics deals solely with the properties of numbers (or other mathematical objects like groups). Those properties may turn out to be useful in real applications (Newtonian mechanics may get a rocket to the moon if it is treated as a point mass with a vector of motion) but mathematics is not developed on the presumption that the numbers/objects really stand for any particular thing. By contrast, data free simulations still make claims about a mapping between programme objects and social reality.[xiv] (This number or list or procedure stands for a pile of wheat and that one stands for a tiger.) If data free ABM did *not* make such claims, it is hard to see why we might be interested in such complicated and arbitrary computational structures.[xv]

Two unpalatable consequences follow from this observation. Firstly, given the complexity of the average model there is an impossibly large number of data free ABM we might create. Which ones should we bother with and why? Secondly, it does not seem that we can compare data free models by any scientifically useful criteria and this means they will simply proliferate rather than progress. We cannot talk about realism because there is no data and the meanings of programme elements are assigned arbitrarily. But we cannot even talk about elegance or concision because there is no point in these unless they have some bearing on the way the ABM is used. A concise statistical model is better because it is more rigorously tested on the same amount of data. We might say that a concise ABM is easier to understand but if it has no connection to reality beyond what the author asserts it represents, what value has that?

Finally, there is a more constructive reason why data free ABM should be treated with caution. Given that they *do* assert a mapping between programme elements and reality why, if social science data exists about objects in the mapping and their relationships, should these *not* be taken into account to narrow the immense space of possible ABM we might build?[xvi]

The concerns regarding ABM anchored by data at only one level are different but still potentially serious. The most serious also applies to statistical analysis (because that is also anchored by data at only one level). In order to give meaningful results, it is necessary that a statistical model not be too complicated relative to the amount of available data. With enough free parameters you can summarise (but *not* explain) anything.[xvii] ABM are typically much more complicated than statistical models. The concern is then how we can be confident that a one level ABM has *explained* a particular set of data rather than just summarised it. Of course, the more complicated and extensive the data, the less likely this seems but without any formal analysis linking model size to the amount of data needed to falsify it, we just have to hope and those outside the ABM community might not share our optimism. To try and put the issue in a nutshell, if you only intend to match data at one level (while ABM, according to the argument I present in this article, is distinctive for trying to match it at two) why not use a traditional quantitative approach which is no less likely to be able to fit arbitrary data but does so much more parsimoniously and with properly developed statistics?[xviii]

Again, the significance of these apparently abstract concerns is much easier to grasp when we consider their practical bearing on ABM in general and the RAI model in particular. Based on previous discussion, the RAI model appears to fall into the class of exploratory ABM which starts with data at the macro level (there are sometimes sharp polarisations of attitudes in society) and uses the technique to explore how this might have come about. (Polarisations can sometimes occur when individuals influence but are not influenced and when people who are too incompatible in uncertainty and attitude do not influence each other but otherwise converge in their attitudes.) In fact, however, closer examination suggests that the RAI model has to be interpreted as abstract mathematics with the worrying ramifications already discussed. This is because the claim about aggregate attitude polarisations (Deffuant *et al*. 2002, section 1.1) is supported in the following style:

“Several examples in the world history show that large communities can more or less suddenly switch globally to one extreme attitude, because of the influence of an initially small minority. Germany in the thirties is a particularly dramatic example of such a process. In the last decades, an initial minority of radical Islamists managed to convince large populations in Middle East countries."

None of this rationale (of which I have quoted only the first part) contains any references to research and in particular to data. To discuss only the first example proposed by Deffuant *et al.*, while it is true that the Nazi party went from 2.6% of the vote in May 1928 to 33.1% in November 1932,[xix] the 43.9% support of March 1933 post-dates Hitler appointment as chancellor (as a result of Presidential action not democratic process) and took place within a context of "unparalleled brutality and intimidation" (O'Lessker 1968, p. 68). Within two months, Hitler had banned trade unions and within five, all other political parties. Thus, in March 1933 (and quite possibly earlier) it would have been extremely unwise to treat voter behaviour as a reliable proxy for attitudes towards the Nazis. None of the other examples given seem to correspond to commonly collected data or to data that, if collected (like election results in states with weak civil liberties), can safely be taken at face value. We are left to conclude that while what Deffuant *et al.* claim may be correct (there may be societies where attitudes have rapidly polarised) and is certainly not implausible, they have not shown that their model is needed to explain empirical aggregate patterns of attitudes. It is perfectly possible that the RAI model explains something that does not actually occur (or occurs very differently from the way that Deffuant *et al.* propose.)

In the same vein, although many of the assumptions made about individual behaviour and social interaction in the RAI model are intuitive, there are no citations for psychological experiments on social influence (which are numerous). Nor does it seem to be the case that the RAI model draws on data from the research it cites. Without data, there may be many models that are equally intuitive but have very different aggregate outcomes. How can ABM research then progress?

Having looked at applications of ABM without data and using data at one level to explore the other level, it is now much clearer why using data at two levels has distinctive advantages.[xx] (The fact that ABM of the first two types predominate in published research may have contributed to the relatively slow acceptance of ABM in the more empirical areas of social science.) Firstly, unlike no data applications, claims that this element of the computer programme is an attitude and that element is a social interaction are backed up by appeals to data. The *reason* this part of the computer programme is claimed to be an attitude is that, in the context of social processes that are themselves empirically grounded, it behaves like attitudes do.[xxi] Secondly, unlike exploratory applications based on data at only one level, we have considerably restricted our capability to achieve a match between real and simulated aggregate data by simply having too little data relative to the size of our model (over fitting) or adjusting the model arbitrarily (fudge factors[xxii]). Processes need to be in the ABM because of evidence that they operate in the real world and they need to look like real world processes as far as we can achieve that based on existing data.[xxiii] We cannot simply add processes to the ABM because they make real and simulated aggregate data match. This requirement, as I shall show, dramatically reduces the space of ABM justifying our attention.

This leads to a third advantage of using data on two levels for ABM development, which forms the crux of this article. If we look at the kind of data that social science actually collects, we see that it already corresponds rather well to the two ABM levels.[xxiv] If we want to know how attitudes change cognitively or through social interaction, we can resort to observation, laboratory experiments and qualitative interviewing. ("How would you say your attitudes to abortion have changed since you were younger?” “How would you justify your attitude to abortion?”)[xxv] If we want to know what patterns of attitudes exist in society over time, we will almost certainly be best served by large-scale representative social surveys. As I shall discuss later in the article, not all the data we might need currently exists and not all kinds of data fall perfectly into the micro and macro categories that I have suggested here but there seems no reason to think that either of these issues undermine the ABM methodology altogether. (After all, no method ever has all the data it needs and no categorisation of research methods seems to sit equally comfortably with all of them.)

To sum up the ABM methodology in one sentence then: If we design the micro level of an ABM making the best use of available data that we can (calibration) and that ABM proves capable of producing simulated aggregate data which resembles real aggregate data (validation), then we have reason to believe that our ABM is not arbitrary (as no data ABM are) and doesn’t simply *match* the real aggregate data (as the exploratory uses of ABM may do), but that it might actually explain observed patterns *because of* the similarity between the real and simulated social processes in key respects. According to this view, we think we know how this particular aggregate pattern arose because we have been able to generate it in the ABM using only micro social processes for which we have at least some independent empirical support. For this reason, the emerging ABM methodology discussed here is often referred to as the generative approach (Gilbert and Troitzsch 2005, Epstein 2007).[xxvi]

As already mentioned in passing, it should be noted that even if ABM methodology is followed (as with all other research methods) considerable practical problems remain. Firstly, there needs to *be* suitable data for both micro and macro levels. This may not occur where a particular field is dominated by quantitative or qualitative research.[xxvii] Secondly, how closely real and simulated data need to correspond to constitute a result cannot established *a priori*.[xxviii] In practice, as I shall argue, this issue is best addressed by a *sequence* of ABM, each justified by improving on the last.[xxix] Thirdly, not all comparisons between real and simulated data are equally well developed technically. For example, it is much harder to decide whether two spatial distributions are alike than two distributions of a simple variable value. Fourthly, how do we actually use different kinds of data (particularly qualitative narratives) systematically to build ABM? Fifthly, how do we adjudicate between different sorts of evidence (for example a model based on a lot of anecdotal examples and one based on a single experiment?) Finally, even following the methodology to avoid the problems of no data or exploratory ABM, do we still have to worry that multiple ABM with very different individual behaviours (not rejected by data) could give rise to the same aggregate properties? (Or that a particular specification of individual behaviour could give rise to almost any aggregate pattern such that it could not be falsified against real data?) If these things are possible, can we design a progressive research strategy that will identify and resolve them? The best way to understand these issues is to construct an ABM using real data and see how they arise in concrete contexts. It is to this project that I now turn.

Available Data for Attitude Dynamics Models

In the last section of this article, I argued that although data free and exploratory ABM are widely published, attention to the ABM methodology suggests that they may suffer weaknesses affecting their usefulness and plausibility. I have also argued elsewhere (Chattoe-Brown 2013) that it is not possible to justify a failure to calibrate or validate ABM because technical issues or lack of data make this impossible given the current development of the field. What seems to have happened instead is that seminal research that *did* attempt to calibrate and validate ABM has been disregarded and (for whatever reason) journals do not currently enforce calibration and validation even though its feasibility has been demonstrated by research of considerable age (from the sixties and seventies).

In this section, I shall provide additional support for the argument that data free ABM are unnecessary (as well as being unhelpful) by looking at data straightforwardly comparable with the outputs of ABM like the RAI model. This data is rather freely available from the British Social Attitudes Survey (http://www.britsocat.com/Home). The three cases discussed here have been chosen partly for the availability of the longest possible data series (although even then they cover only about 25 years from 1983-2011) and partly because they involve controversial issues that seem most likely to demonstrate the polarisation justifying interest in the explanation offered by the RAI model.

Figure 1. Changing UK Attitudes about Homosexual Sexual Relations

Figure 1 shows the percentage responses (for sample sizes varying between about 1000 and 3300 in different years) agreeing and disagreeing with the statement that homosexual sexual relations are wrong.[xxx] I have coded a negative attitude for always and mostly wrong, a positive attitude for not at all or rarely wrong and a neutral attitude for sometimes wrong or depends/varies. Although it is not clear what error bars would apply to this data, neutral attitudes appear to remain fairly constant while negative and positive attitudes are changing places. In addition, the data shows a reasonably clear turning point around 1987/1988 and there is arguably another in 2006/2007.[xxxi] Thus, this graph displays neither polarisation (decline of neutral attitudes at the expense of positive and/or negative ones) nor equilibrium at any particular percentage of positive and negative attitudes.

Figure 2. UK Attitudes about the Death Penalty for Some Crimes

Figure 2 reports similar data for attitudes about the death penalty. While this plot (unlike the other two) arguably displays convergence to equilibrium values for the distribution of attitudes (after about 2000), it does not display polarisation (positive and negative attitudes are not gaining ground at the expense of neutrality – in fact over the period as a whole positive attitudes are losing ground slightly to neutrality) and there is an even more plausible turning point in 1991/1992 than either of those perhaps shown in Figure 1.

Figure 3. UK Attitudes about the Reunification of Ireland

Finally, Figure 3 reports similar data regarding the reunification of Ireland. Here, we find neither equilibrium nor polarisation (in fact, neutrality appears to be gaining ground at the expense of positive attitudes while negative attitudes remain broadly constant) and there again appear to be some plausible turning points (1994, 2001-2003 and 2004).

Thus it is clear from even a small amount of real data that what ABM need to explain (and should be validated against) does not appear to be equilibrium (whether moderate or polarised) but a combination of non-trending (stable) and trending attitudes (and it may be the neutral attitude that changes its prevalence rather than the positive or negative ones) where there is, nevertheless, at least some evidence for turning points rather than simple trends. As suggested earlier, explaining something that doesn't seem to be occurring (at least in the BSAS data) may limit the usefulness of the RAI model even if it is intended to be interpreted as data free.

It is possible to make a similar argument regarding calibration. I have already made the point that the RAI model does not cite any data to back up its claims that those who have very different attitudes from our own do not influence us and that, provided agents influence each other at all, their attitudes will tend to converge. But in the same way that relevant attitude data is not being neglected because it does not exist, nor is this true of research on social influence. Unfortunately, a closer examination of this research area suggests that definitive understanding is lacking but this is not incompatible with the value of ABM at least being based on what *is* known. The difficulty is that, depending on the exact details of the experiments conducted (see White 1975), it has been found *both* that attitudes shift more the greater the discrepancy between the influencer and the influenced *and* that, if attitudes are very different, there may be no influence (or even a negative one). What seems to be lacking, despite many years of research, is a relatively concise theoretical statement about the empirical circumstances under which each effect may be found.[xxxii] Obviously, such a theoretical statement would be very convenient since it could be implemented directly within an ABM so agents behaved realistically in a range of circumstances. However, for the purposes of ABM methodology, empirically grounded modelling need not (and probably should not) be postponed until a definitive theory exists.

Consider, for example, experiments conducted by Hovland *et al*. (1957). In these, subjects who had expressed attitudes on an issue (whether alcohol should be available in a US state) were asked to respond to other expressions of attitudes (for ecological validity, these were constructed from real media and public debate on the topic which had recently occurred.) They were asked to indicate which expressions they found odious and which they were able to agree more or less strongly with. Using this data, Hovland *et al*. were able to systematically explore the effect on attitude change of individuals with any given attitude (in a defined range) being presented with messages (short narratives of attitude and evidence recorded using standardised voices) corresponding to different degrees of support or opposition to prohibition. For the purposes of my argument, a key result can be found in Table 3 on page 249. This shows, for example, that when someone strongly in favour of prohibition (a Dry) is exposed to a message strongly opposing it (a Very Wet message), they will change their attitude in the direction of Wetness 27.5% of the time, leave it unchanged 49.3% of the time and become more Dry 23.2% of the time. By contrast, when they are exposed to a message that is only Somewhat Wet, the inclination to move towards it increases slightly (31.6%) and to move away it decreases slightly (19.3%) but failure to be influenced remains almost unchanged (49.1%).

It is important to draw the right conclusions from my discussion of this research. It is only one study and it is very old. However, as with comparing real data on attitude distributions to simulated data, it would appear that even limited real data calls the RAI model significantly into question (and this is a general argument for breaking down the current culture of data free modelling in ABM research).[xxxiii] Unlike the assumptions of the RAI model, it seems to be the case that even a moderate opposing attitude can (far from influencing you towards consensus) make you hold your own attitude less moderately a significant proportion of the time. The implications of this finding for attitude dynamics are likely to be significant. As an extreme simplification, the RAI model works by the tendency of attitudes to move together (which is only qualified by ignoring those whose attitudes are too different from yours). Stochastic elements (which agents meet in what order) and the presence of agents who influence without being influenced have a bearing on *where* these attitudes converge (moderate consensus or one/two extreme groups) but do not change the basic dynamic of convergence. The Hovland *et al*. article shows that, reasonably frequently, one social actor may move *away from* and not towards the attitude of another. Under such circumstances, a gradually narrowing envelope of possible attitudes (leading to ultimate convergence) is avoided and continuing attitudinal heterogeneity (which we observe in the BSAS data) becomes a possibility.[xxxiv]

Thus ABM does not require a perfect theory of social influence to make progress on the integration of different kinds of data providing it follows the logic of its methodology. All it requires is that an ABM based qualitatively on the Hovland *et al.* results (and others like them) produce attitude distributions that look more like the real data from the BSAS than those from the RAI data do.[xxxv] In the section after next I shall investigate this possibility using an alternative ABM but before doing so I shall consider some other candidate social processes for explaining attitude dynamics.

Other Possible Social Processes in Attitude Dynamics

As well as the resemblance between real and simulated aggregate data (validation) and the empirical basis for micro assumptions (calibration), ABM also has to consider the scope of its models in particular domains in terms of different kinds of social processes likely to be relevant. Even before we establish exactly what is known about a particular process it makes sense to look at existing research to identify what we might call candidate processes. This approach can again be illustrated by reference to the assumptions of the RAI model. In particular, for everything that an ABM puts in, it is reasonable to ask what it leaves out and why because the innovative focus on explicit process based models of interaction may overshadow other important processes driving attitude dynamics. For example, how likely is it that all attitude change results from interaction between agents as the RAI model assumes? In particular, it seems highly probable that mass media accounts will be widely diffused and thus significant determinants of attitudes (St. George and Robinson-Weber 1983). Similarly, the RAI model effectively involves random mixing of agents when it seems at least plausible that social networks structure our interactions relevant to political attitudes (Huckfeldt *et al.* 2004). It is also clear that real social actors have more than one attitude and while modelling a single attitude may seem like a reasonable technical simplification, it disallows the possibility that individual attitudes may be stabilised relative to each other (by such cognitive processes as consistency checking or cognitive dissonance reduction) and form patterns (Fleishman 1986, Mullainathan and Washington 2009).[xxxvi] Finally the relationship between certain kinds of knowledge and attitudes may have a bearing on how likely they are to be modified[xxxvii] (Egan and Mullin 2012, Evans and Durant 1995).[xxxviii]

Thus we see, regardless of whether or not they actually turn out to improve the empirical performance of an ABM, there are simplifications in the RAI model that are surprising given our social science knowledge and processes we might reasonably expect to find in an effective attitude dynamics ABM. Furthermore, however we interpret the RAI model, this is a problem. If the RAI model is data free, why spend time studying an abstract model that doesn't match what we already know? If it is supposed to be exploratory, why does the phenomenon it intends to explain (polarisation) receive so little justification? If it supposed to be a two level ABM why isn't the interaction process empirically supported either? I shall now take these ideas of calibration, validation and scope (likely candidate processes to explain a particular phenomenon) and explore them in the context of a new model of attitude dynamics.

A New ABM of Attitude Dynamics

In this section, I present a new ABM of attitude dynamics as a concrete basis for considering the methodological and practical issues I have raised so far. This model differs from the RAI model in four important respects:

1) There is a very simple mass media process.[xxxix] At any time media outlets can be sending out messages corresponding to particular values for attitude. For example, one media outlet may be urging people to have an attitude of 27 on attitude 1, -55 on attitude 2 and -80 on attitude 3. (All attitudes are represented as having values between -100 and 100 with values greater than +/-76 be classified extreme, values between +25 and -25 classified neutral and all remaining attitudes classified moderate.) The intuition behind this is that newspapers regularly campaign on issues relevant to attitudes and they may be urging positive, negative or neutral positions on these. (A neutral campaign might be thought of as urging moderation.) There is a chance (with a default value of 3 in 1000 per tick – with each tick representing a day) that a media outlet will start a campaign (go from promoting a neutral attitude to promoting a positive or negative one).[xl] There is another chance (also with a default value of 3 in 1000 per tick) that an existing campaign will end and the media outlet will revert to an attitude in the neutral range. This means that campaigns vary in

length and are separated by periods when the media impact is neither positive nor negative. The intuition behind this is that a media outlet would lose credibility if they were campaigning flat out to get us out of Europe one day and keep us in the next. For the simulation runs reported here, there are three mass media outlets, agents attend to only one of these (selected randomly with equal probability) and a fraction of agents controlled by the user (set here at 55% to correspond roughly to UK media usage figures) do not attend to media at all.

2) There may be social networks that are static or dynamic.[xli] It is possible to run the ABM in a condition of random mixing (like the RAI model) or one in which, at initialisation, agents are placed in social networks. These networks are very simple and the only constraint on them is that the number of ties each agent has is broadly compatible with the number of close friends found in the political discussion literature (3-7). The user also has a choice about whether to run the ABM with this network fixed over the length of a simulation run or whether to allow ties to be broken and formed at a very low rate. The dynamic network is again extremely simplified in that ties are made and broken randomly. (In fact, it is likely that significant attitude differences may be one thing that breaks network ties but we disregard this possibility for now.) In both cases, there is no change in the number of interactions per tick. The only difference is whether these interactions are chosen randomly from the whole population or from a subset defined as friends.

3) Social interaction follows the Hovland *et al*. findings in that social influence can sometimes lead to convergent attitudes, sometimes to no change and sometimes to *divergent* attitudes. The user can control the gap between the attitudes of individuals that defines their attitudes as being far apart. If their attitudes are not far apart, the attitude of the influenced will move towards that of the influencer by a small random

range (default value 1-5 attitude points). If the attitudes are far apart, following Hovland *et al*., there is a 25% chance of the same kind of convergence, a 50% chance that the potentially influenced agent will actually not be influenced and a 25% chance that their attitude, instead of converging, will diverge by a similar amount (again 1-5 attitude points).

4) The model contains three attitudes. Depending on programme settings, these can be subject to media and individual interaction effects independently. However, if the user chooses they can add a process by which a set of archetypes exist in the world. (These are defined arbitrarily at present to cover a range of possible attitudes from all extreme to all moderate.) In this condition, agents have a small probability per tick to move each of their attitudes slightly towards the nearest archetype (measured as the simple sum of differences between the three archetype attitudes and the three attitudes currently held by the agent). The intuition behind this is that, unlike the RAI model, attitudes do not exist in a vacuum and some combinations of attitudes (represented as the archetypes) are more robust than others. (For example, it may make sense to support access to abortion, equal pay for women and gay civil partnerships simultaneously within a broader context of gender equality.) As such, as well as influences from the media and other agents, the archetypes also exert some influence through internal processes of reflection with the result that agents are not entirely at the mercy of external forces but have an internal structure to their beliefs which affects their ability to be influenced. (Of course, if the forces of external influence are great then an agent may come to be closer in their beliefs to another archetype and it is then that which will attract their attitudes.) In the discussion that follows, this process will be referred to as cognitive consistency.

It is important that the reasons for these assumptions are interpreted correctly. In each case there is evidence from the literature that the broad process specified (news, networks, multiple attitudes, cognitive consistency) occurs. All processes are completely absent in the RAI model (though some are discussed as possible extensions or modelled in earlier unpublished articles by some of the same authors). However, the assumptions regarding the exact operation of some process are clearly extremely simplified (and the associated parameter values more or less completely arbitrary). However, this is deliberate. (There is also the practical issue that describing and justifying a wide range of more precise assumptions could take up an entire article in itself and this is not mainly intended to be an article reporting a model.) The point at this stage is only to show that an ABM using selected processes and making at least some attempt to calibrate them (the number of close friends in political discussion networks, the possibility that attitudes may diverge through influence, approximate levels of media access) already makes a better job of resembling the BSAS data qualitatively than the RAI model does. I would be delighted if someone were to supersede my ABM with another that had better empirical grounding for more of its micro assumptions and resembled the BSAS data more closely still. (In other words, it did not only fit the BSAS data qualitatively but more or less quantitatively.) We would then have moved into the progressive research strategy that it was part of my objective to advocate in writing this article. Such an immediate refutation of my model by a better one would actually be a major step forward from the current situation where ABM are asserted to be anecdotally plausible and simply proliferate without data either being used in their construction or to adjudicate definitively between them. (Instead, there are futile – because irresolvable – arguments about what is more intuitive.)

Figure 4. Baseline simulation run with red meaning a negative attitude, green a positive one and grey neutrality: No news, no networks, no extremists, no disregard of discrepant attitudes and no cognitive consistency process.

Figure 4 shows very simple behaviour with individual interaction as the only source of attitude change.[xlii] Since all attitudes move towards each other in this case, extreme views are gradually eroded and (perhaps unsurprisingly) neutrality becomes the dominant attitude. (The other two attitudes for the same simulation run are virtually identical in profile.) Unless stated otherwise, in the simulation runs reported here, the user controls the initial populations of agents with different attitudes. The default settings (the same for all three attitudes) were 10% of agents strongly agree and 10% strongly disagree, 25% agree and 25% disagree and 30% are neutral. This suggests that, in the absence of any Hovland *et al*. divergence of attitudes what must drive the RAI result is the presence of extremists and the properties of the uncertainty process that prevent invariable convergence to neutrality. (In fact, the Deffuant *et al*. argument is somewhat odd on this point. Their model only converges to extremism sometimes. However, it is not clear how that could be reconciled with any particular observed instance of extremism.) It turns out that, selecting only single features of the model, the same pattern is observed for static and dynamic networks alone (although unanimous neutrality occurs at different rates). Cognitive consistency alone produces fairly stable non-extreme equilibrium values in all attitudes that differ from the initial distribution.

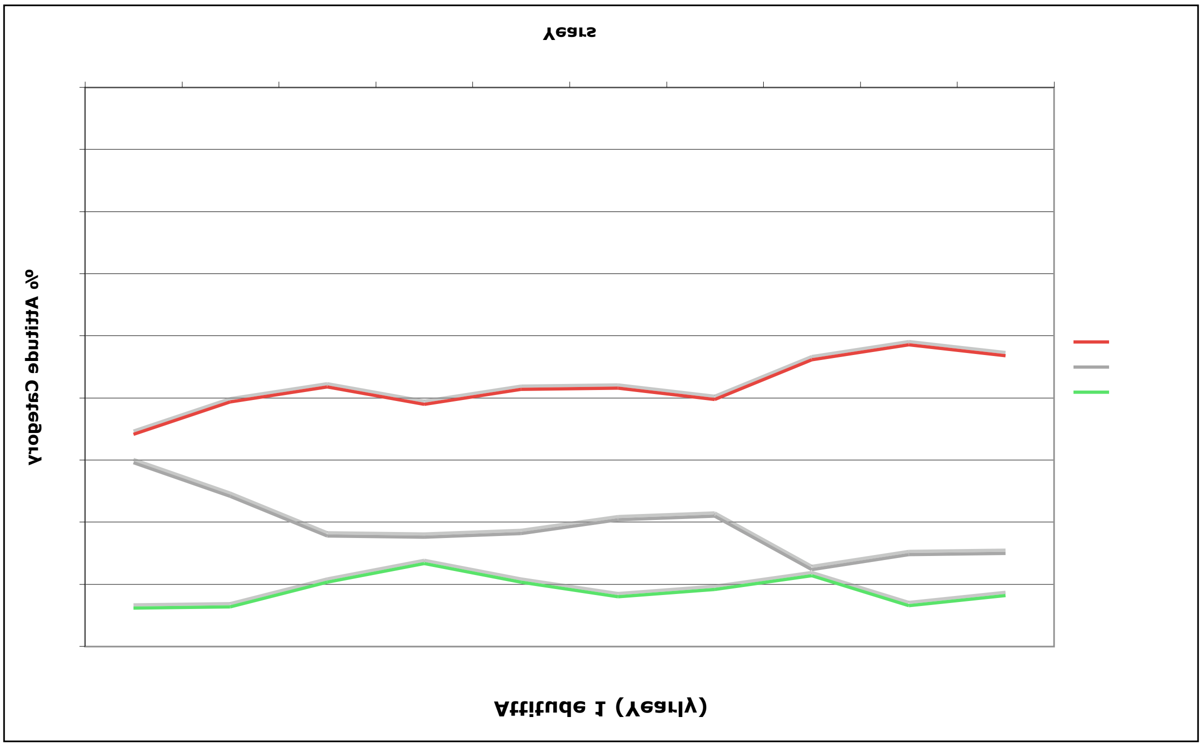

Figure 5. The effect on attitude dynamics of combining news, dynamic social networks, 15% extremists, cognitive consistency and disregard of discrepant attitudes

By contrast, Figure 5 shows the effect of combining the various processes already discussed (intended to be empirically plausible based on existing research). The raw data (tick by tick) has been sampled at annual intervals over 10 simulated years to mimic the form of the BSAS data. The gap beyond which attitudes are subject to divergence or being ignored (following Hovland *et al.*) is 50 attitude points. The extremists are automatically assigned to one of the archetypes with all its values in the extreme ranges (either greater than 75 or less than -75) and are immune to all influence (from media, cognitive consistency and social interaction).

This data shows at least qualitative similarity with the BSAS data, particularly when compared to the RAI model. There are both periods of stability (years 3-7 in neutral attitudes), relatively stable attitudes (positive attitudes over the whole period) and trends with recognisable turning points (negative attitudes over the whole period). Each attitude range (agreement, neutrality and disagreement) can independently show a trend or stability. (In this case, agreement is dominant with neutrality gaining ground at the expense of disagreement but other runs showed other combinations.)

Although the model is far too basic (and far too many of the parameters are arbitrary) for prediction to be a reasonable goal (in which case the measure of fit would simply be the gap between two time series), it is useful to consider how we might roughly compare real and simulated data here. One possibility is to compare the changes in attitudes over a time period. Table 1 shows that while the results of this comparison are hardly overwhelming, the real and simulated data are at least in the same value range. Furthermore, I neither tuned any parameter values from their initial arbitrary settings to achieve this result nor did I select the real data used for comparison on the basis of fit.

Attitude Category	Simulated Data (ten arbitrary years)[xliii]	Death Penalty Data (1998-2007)[xliv]
Positive	3.6%	6.3%
Neutral	8.6%	2.2%
Negative	7.2%	7.1%

Table 1. Comparison of changes in attitude categories over a ten-year period for real and simulated data

At the risk of labouring the point, it is important to be clear what this article attempts to show. It is not that the model presented here is particularly plausible. It is only to show that by attention to readily available data for calibration and validation (and reflection on what is already known in social science), it is fairly easy to improve the performance of ABM to produce at least qualitative similarity with real data (which the RAI model and its variants do not do). Obviously the next stage would be to progressively improve the *quantitative* similarity of the model by the methods outlined here (identification of broad social processes likely to be relevant, use of existing data on these processes for progressive calibration, more discriminating attempts to quantify similarity in validation in proportion to the performance of the simulation.) However, apart from limitations of space, that process adds nothing to my key argument. Attending to data *at all* can already produce a recognisable improvement in performance. There is no reason to suppose that such a performance improvement cannot be sustained by further attention to data following the methodology of ABM discussed here.

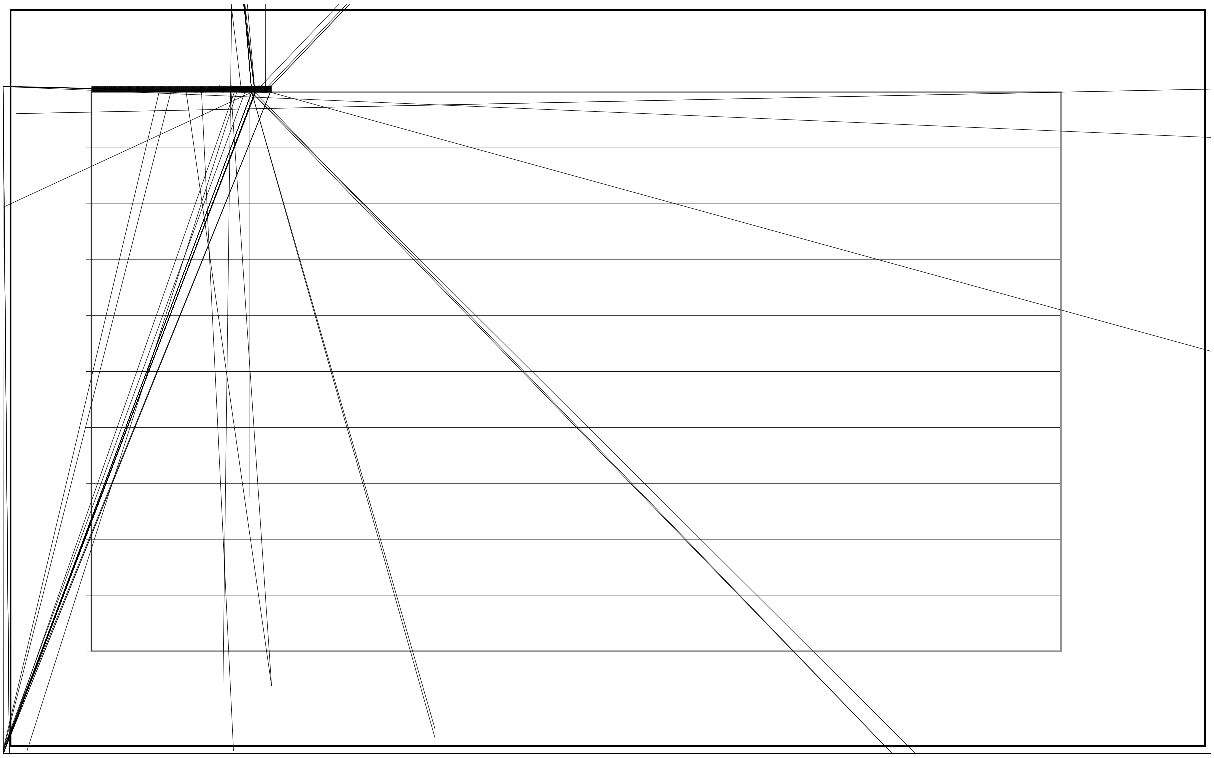

Figure 6. Daily attitude distributions for the same simulation run shown in Figure 5.

However, even this level of success raises a number of points that need consideration in the context of ABM methodology. Figure 6 shows the same simulation run as Figure 5 but with all the data plotted (rather than just data for one day at the same time each year to mimic the BSAS data). Firstly, despite its perceived quality relative to cross sectional data, it is clear that even large-scale longitudinal data collected annually loses a lot of information relative to the simulated data.[xlv] It is not unreasonable to think that attitude change takes place over time scales much shorter than those over which the data is collected but the exact nature of the changes that may arise from these more rapid processes are lost to sight. Secondly, following from this, we observe mirroring effects between attitude categories (in this run mostly between positive and neutral categories but also towards the end of the run between neutral and negative categories). By this I mean that the time series of one category follows rather closely the time series of the other (something that is largely washed out in the annualised data.) See, for example, the pattern between about day 936 and day 1684 for positive and neutral categories. On reflection, it is not so hard to see what is happening but it has some interesting implications for our view of attitude dynamics which might not have come to light had it not been for the explicit process specification offered by ABM. When the mirror effect occurs, it means that most of the change in the system is occurring across one of the boundaries (positive to neutral or neutral to negative) and not the other. If we imagine an influence (say a media campaign) that is pulling attitudes towards greater positive values then we would see an initially strongly negative value first crossing the negative/neutral boundary and (later) the neutral/positive boundary. (How much of this happens also depends on the distribution of attitudes at the start of the campaign. Only some of the population will have the negative attitudes that are likely to move across both boundaries.) Mirroring suggests that attitudes are going across one boundary but not crossing the second. For example, before the media campaign can give everyone a positive attitude, it changes to a neutral one. The net result is that a large proportion of agents spend their time in transit across the neutral attitude zone and this leads to dominant neutral attitudes. We observe that the least convincing of the comparisons between real and simulated attitude change in Table 1 is that for the neutral category. In the BSAS data, neutral attitudes are in a significant minority, while in the simulated data, neutrality is often the majority category. This suggests that part of the relative size of the categories may just reflect their relative width in terms of attitude points. (If neutrality was 10 to -10 and not 25 to -25, the zone would take much less time to cross and there would therefore be far fewer agents in transit across it at any time.) This also makes one wonder whether, in fact, social actors sometimes do move directly from positive to negative attitudes (or vice versa) without an intervening period of neutrality and that is something, at least in principle, that could be established empirically (by asking people who have significantly changed their attitudes exactly how it came about.) This would be another example of the claim that ABM helps us to think clearly about social processes, both when we specify them and when we confront them with data.[xlvi] Since this mirroring effect is barely visible in the BSAS data, it is unlikely that the need to explain it (and thus consider its causes to develop further provisional theory) would have arisen in a statistical approach. Finally, as another example of provisional theory building, we can see how this model might develop further. Figure 4 draws attention to a convergence tendency in attitudes. The RAI model shows us how a combination of extremists and uncertainty change may channel that convergence into extreme rather than neutral views. But this raises a wider (and probably more interesting) question generally about social processes that may not only maintain, but also increase attitude diversity. (It is fairly clear, as far as we can see anything from the BSAS data, that equilibrium in any sense is conspicuous by its absence.) The Hovland *et al*. divergence behaviour for discrepant attitudes is one such. Could another be the feedback between individual attitudes and media competition? Unlike the simple model presented here, it is likely that media outlets pay attention to what their consumers will like and tolerate and also to what other media outlets are doing. The ABM presented here allows all media outlets to agree on an attitude independently. I suspect that in practice media outlets often contradict each other more or less deliberately. One strategy for improving the performance of the attitude dynamics model presented here may simply be to tune the parameters.[xlvii] Another may be to look for the class of models that do not generate equilibria (a quest initiated by attention to the BSAS data) by reflecting on social processes likely to lead to increasing or variable attitude divergence rather than constant or decreasing divergence.[xlviii]

I have already argued that the data free and exploratory interpretations of ABM may not be as unproblematic and useful as is often implied. I have now shown that with only moderate attention to freely available data, it is possible to build a model that achieves at least qualitative similarity with the BSAS data (which the RAI model does not). I now turn to the wider implications of this argument.

Prospects for the Future

Despite general acceptance of the generative methodology in ABM, models that are validated and calibrated on real data remain in a significant minority. I have described problems that arise from this institutional practice but the main aim of the article has been to show that use of data does not have to be either labour intensive or inspired to produce improved model performance. Apart from the problems with no data and exploratory ABM, there seems to be no compelling reason *not* to attend to data that already exists. How then, does this article feed into the existing generative methodology? Basically, it suggests that there may be rewarding intermediate steps before direct comparisons between (for example) two sets of attitude time series. Firstly, it makes sense to survey the existing literature for widely accepted candidate processes that should occur in some form in the model (even if their actual specifications are rudimentary and the parameter settings speculative). It seems unlikely that the whole social network community is wrong about the relevance of social networks and, unless that is so, it seems unlikely that a model containing some network process (however oversimplified) won't perform better than one that one where networks are completely absent. The results presented here seem to support this view to some extent. On this basis, I added media, networks, attitude divergence and cognitive consistency to my model and despite simplifications and arbitrary parameters achieved some qualitative similarity with the BSAS data. Secondly, it is now possible to explore the parameter space of the model to see if the qualitative similarity can be turned into quantitative similarity (and if so of what quality). This might be achieved by simply searching the parameter space (though with the restriction that some parameter combinations – 85% extremists for example – are implausible in calibration terms.) It is rather likely, however, that parameter adjustment alone will only be able to achieve rather limited improvement in similarity. This is because not only the parameter settings but also the assumptions of the model are arbitrary. This challenge can be addressed in two ways. Firstly, the assumptions of the model can be better grounded in existing research (as I did with Hovland *et al*. and to a far lesser extent with media access and numbers of close friends.) This should simultaneously narrow the parameter space and, presumably, improve similarity between real and simulated data. However, what is also likely to happen is that in exploring the behaviour of the model, certain patterns will resist parameter adjustment (for example the dominance of the neutral attitude category when it is usually the minority in the BSAS data). This will inspire provisional hypotheses about what else might be needed in the model along with a search for corresponding data and more exact process knowledge. It is to be hoped that this combination of adding processes, exploring the parameter space and observing systematic discrepancies with the real data should gradually converge the set of possible models.

This set of strategies gives us a more nuanced view of how ABM methodology may actually be carried out in practice.[xlix] With proper attention to data (which the RAI model lacks), we can exclude certain combinations of social processes as candidate explanations for observed data relatively easily. This in turn leads to a steadily narrowing set of plausible combinations of included social processes and calibrated parameter values that could produce the data we actually observe. (It is one thing not to be sure whether people have an average of 5 friends or 3 but any model that only matches the real data on the assumption they have 500 is almost certainly wrong. Calibration, like validation, can be done incrementally as the overall quality of competing models improves.) As the requirements of similarity are tightened (not just some turning points but turning points of the right magnitude and frequency or even turning points at the correct moments in time) it would be reasonable to assume that further combinations of social processes and parameter values would drop away as candidate solutions.[1]

Although this process may be clear conceptually, it is far from trivial in practical terms. Relevant data may simply not exist (and may thus need to be collected.) Synthetic models will be harder to construct simply because relevant data may only be found far away from the home discipline of the researcher (and may be hidden by local quirks of terminology). The exploration of large numbers of parameter and process combinations (sensitivity analysis) remains computationally demanding. Nonetheless, I hope this article has shown clearly why there are advantages to this way of proceeding in allowing different kinds of attitude research (and their associated data) to be synthesised and why it might reasonably be expected to work in a way that offers scientific justification for the resulting models and their conclusions.

Conclusions

This article has shown how ABM works in the context of two models of attitude dynamics. It has presented the emerging methodology of ABM and suggested both what the drawbacks are in not following it (particularly with respect to the use of data for calibration and validation) and what the advantages are in following it. For sociology more generally, a big advantage may be the possibility of incorporating the quantitative and qualitative strands of attitude research into the same models. (The article has also shown how this integration may occur between psychology and research that is often considered sociological.) It has also shown (using BSAS data and psychological experiment data) that if ABM is currently not having much to do with data, this is not because the data is unavailable or unusable even if the quality of match between real and simulated data is only suggestive at this stage. Finally, it has shown that it is possible to produce qualitative similarity with real data using relatively modest amounts of readily available calibration data and suggested how this kind of modelling can help us think in novel ways about the understanding of social processes more generally. (For example, how can we create a system with the variable attitude discrepancy needed to mirror the non-equilibrium observed in the BSAS data rather than declining attitude discrepancy that seems to drive equilibrium in Deffuant *et al.* and its successors?) It seems unlikely that such questions could be formulated clearly or answered without the contribution of the ABM approach.

Notes

[i] It might be argued that the extensive literature on mixed methods (Tashakkori and Teddlie 2010) undermines this claim. In practice, however, it seems that methods are mixed *within a research project* (and could equally well be described as multiple methods) but otherwise continue to be conducted in the traditional manner. They are not, therefore, mixed in the way that ABM mixes them or in the sense that mixed sex schools are. (It would be misleading to describe a school as mixed sex if it maintained separate playgrounds, lunch sittings and classes.)

[ii] The model is presented as being about opinion dynamics but nothing is said about how opinions might differ from attitudes or why it should be a model of one and not the other.

[iii] As with mixed methods, interdisciplinary research often seems to mean the disciplines carrying on doing what they always do but in the context of a single research project.

[iv] An example would be a response to the question “I think abortion should be illegal under any circumstances” with responses coded from 5 (strongly agree) through 3 (neither agree nor disagree) to 1 (strongly disagree). An example of a quantitative theory might be that societies display common patterns of relative social mobility despite significant differences in welfare systems, forms of governance and so on (Grusky and Hauser 1984).

[v] For a typical (though more sophisticated) example, see Ohlander *et al.* (2005).

[vi] In order to make this argument in a reasonable space it is necessary to stylise these large areas of research somewhat. There are exceptions and borderline cases like the quantitative analysis of standardised observational data but these do not detract from the general claim that a large proportion of published social research is straightforwardly qualitative or quantitative in the way that I describe.

[vii] For example, we might believe that more industrial production gives rise to more consumption and more pollution, more pollution gives rise to less health and less health gives rise to less production and more consumption. What will be the dynamic path of the health, consumption, production and pollution variables and will it depend on the starting conditions?

[viii] By contrast, it is widely recognised that building an ABM really does make us think differently about the detail of social processes and how they operate. For example, the coverage measure in harm reduction treats half the population of intravenous drug users having all the hypodermic needles they need and the whole population having half the needles they need as equivalent. Are they in fact equivalent in their effects on the transmission of the blood borne conditions that needle exchange schemes are supposed to reduce? An ABM could set about answering this question.

[ix] What is relatively unproblematic cannot be established *a priori* but only by development and critique of ABM as part of the research process. This process will be illustrated using the ABM discussed later in the article.

[x] In linear systems these two aspects again become degenerate. Collective saving is simply the sum of individual savings. However, an important theoretical finding of ABM is that we should not assume that even very simple social systems are linear (Chattoe-Brown 2013).

[xi] Famously, Gode and Sunder (1993) show that in the very socially attenuated world of the auction, rationality has no added value and so called “zero intelligence” traders can perform very well. In the language of sociology, structure here outweighs agency.

[xii] In turn, this is perhaps because agents and their interactions in an ABM resemble real social actors in important ways (like having different knowledge and capabilities) rather than being mediated by narratives or equations (which often make simplifying assumptions like perfect information). In particular, the agency in the system resides where we think it does in the social world. Inflation is actually *caused* by the interactions of people and institutions and not by the structure of equations (though under certain circumstances those equations may identify regularities which *reflect* those underlying causes).

[xiii] This model was published as Deffuant *et al*. (2002) and developed in Deffuant (2006). It has also formed the basis for extensions by other researchers, for example (Jacobmeier 2006, Malarz *et al*. 2011).

[xiv] Unfortunately, it is rare for ABM researchers to claim explicitly that their models are purely formal and follow the logic of that claim. Instead they tend to hint that their models might be useful anecdotally while neglecting to provide substantive evidence. I will support this claim shortly.

[xv] It might be argued that mathematics has a reasonable chance of real world applicability purely through its relative simplicity but if ABM were truly data free wouldn't it be an extraordinary coincidence if such complicated arbitrary structures did resemble real social processes?

[xvi] This concern can also be expressed in a different way. If this part of the programme is really supposed to be a pile of wheat and that part is really a tiger, there is no point in modelling the situation where the simulated tiger eats the simulated wheat because tigers don't eat wheat. What makes this thing in the computer programme a simulated tiger is that it follows tiger like processes with respect to other objects (such as wheat or antelope). Thinking that we can merely declare something to be a tiger (or an attitude) in an ABM seems to be a mistaken extrapolation from statistical approaches where we really can say that the number 5 represents the attitude of a particular person to abortion. This may work with variables connected by equations but it doesn't work with variables that are part of processes.

[xvii] In statistics, this is over fitting (Skiena 2001) but to my knowledge there is no equivalent procedure for evaluating the complexity of an ABM relative to available data.

[xviii] It is possible to make definite arguments for using ABM rather than statistical models but these also turn out to depend on the examination of real data. See, for example, Chattoe-Brown (2010).

[xix] This figure far from represents a majority and four years is not necessarily a particularly short time period given the volatility of German politics in these decades. For reference, there were 4 Reichstag elections during this period and votes for the Non-Catholic middle parties experienced a change of similar magnitude (from 27.5% to 3.5%) while some other parties maintained relatively stable support throughout. Nazi support also showed a noticeable decline (4.2%) between July and November 1932 while there is no sign of aggregate reversals in the RAI model.

[xx] This point has been made in other contexts, for example in Coleman's (1990, pp. 1-23) claim that meaningful explanation requires causal claims linking different levels of description.

[xxi] This article follows Deffuant *et al*. (2002) in treating agent attitudes as signed variables (between 100 and -100 for example). Such scales are often produced in collecting data about attitudes and that might lead one to conclude that they correspond to the mental content of social actors. However, it is perfectly possible that, cognitively, attitudes are neither stored nor modified as scales. For example, agents might respond emotionally to various mental schema like "Would you let your daughter marry one?" and then report the strength of that emotional response. In this view, attitudes could change not only by modification of emotional response to a particular schema but also by changes in the set of schemata held by an individual so tolerance would involve the process by which "We're all the same under the skin" came to be adopted as a new schema and associated with a positive emotional response. In this case, it is unlikely that an ABM based on signed variables would accurately reproduce aggregate attitude dynamics.

[xxii] See, for example, Goldsmith (1997).

[xxiii] Further, any ABM is subject to later challenge by new data, a more realistic specification and/or a better match between real and simulated data. This is a component of the progressive research strategy advocated in this article.

[xxiv] This also suggests institutional opportunities for ABM. Social science may be much more willing to engage with a research method that, in turn, makes use (perhaps even better use) of what social scientists have already discovered.

[xxv] Attitude change is a borderline case for psychology and sociology because it is not clear how meaningfully social actors can articulate the process by which they are influenced. It may be that a combination of interview and experimental research is therefore more valuable than either approach alone.

[xxvi] Although I have not been able to discover any definite objections to this methodology, the implications of not following it seem to be widely ignored in published ABM research.

[xxvii] For example, see Bertaux and Thompson (1997) on the quantitative domination of social mobility research.

[xxviii] This would be like talking about the correct significance level.

[xxix] The Schelling model (a very popular ABM) provides an instructive example (Schelling 1971). As research has explored this model, it has become clear that its basic result for two agent types (clusters form) occurs more or less regardless of the micro assumptions made. In this situation, more data about the actual behaviour of social actors and/or about real patterns of segregation (like that provided by Hatna and Benenson 2012) are needed to establish which model might actually explain reality.

[xxx] The data is discontinuous because not all questions are asked every year. The challenge for ABM would be, after all, to produce a resemblance to data in just this form.

[xxxi] We have to make claims about turning points with caution. We don't know how much of the change year on year is attributable to reporting error, biases created by exactly who responded in a particular year and so on. We can only work on the assumption that year on year shifts that are much bigger than the rest of the year on year variation might be turning points.

[xxxii] These might involve the discrepancy between attitudes or such things as whether the influencer and the influenced liked or trusted each other. A number of such relevant factors have been identified (Wood 2000).

[xxxiii] Even if the RAI model is intended to be data free (despite its anecdotal professions of relevance) the Hovland *et al*. data suggests it would be more sensible to devote our time to a data free model where attitudes do not always converge when influence occurs. Even one piece of real data trumps any amount of plausible anecdote when models are so far from being calibrated and validated. Of course, if we ever reached the point where (through progressive research) we had two ABM that quite well matched the BSAS data then the actual quality and typicality of the Hovland *et al*. finding might become crucial. But we are very far from that point at present so only the basic result (not all influence converges attitudes) is necessary to distinguish different ABM and their empirical potential.

[xxxiv] To restate the issue concisely it is not necessary that we have exactly the correct model of social influence to question the usefulness of the RAI model. All we have to do is show that the assumptions of the RAI model directly contradict experimental evidence about the more general issue of whether attitudes can *sometimes* diverge rather than converging under social influence. Of course, we would rather know exactly when each effect occurs and how big it is but even the fact that divergence is empirically *possible* seems likely to undermine the value of RAI model which assumes that it does not occur. Even if the RAI model is supposed to be data free it isn't clear why we would study it rather than a data free model compatible with the Hovland *et al*. results.

[xxxv] In the initial stages of ABM development more like need mean no more than has turning points rather than doesn't have turning points. It may only be when ABM shows real signs of performance improvement in the resemblance between real and simulated aggregate data that sophisticated assessments of similarity become necessary. It is for this reason that I have not discussed the extensive literature on validation, for example Moss (2008) and Windrum *et al*. (2007). I shall return to this point.

[xxxvi] Somebody at the pub may temporarily sway you to a more negative view of immigration but the following morning you are likely to realise that the view makes little sense in the light of your other attitudes to equality, tolerance, economic efficiency and so on.

[xxxvii] Although this feature was implemented in the model, it is not reported here, partly for reasons of space and partly because the outcomes appear extremely counter-intuitive.

[xxxviii] This list is not intended to be complete and other mechanisms are clearly possible. For example, extreme attitudes entailing extreme actions (like celibacy) may be tiring or inclined to lapses. Cognitive dissonance may reduction thus result in the erosion of extreme attitudes from within. There are also cases where processes may interact. For example, strong beliefs (whether they are true or not) will both determine and protect some attitudes. For example, as Siraj (2009) shows, Muslim attitudes to homosexuality are justified on religious grounds and such justifications seem less likely to be modified than mere attitudes as long as the underlying religious faith persists. We might then come to a view of attitudes as a cognitive network where changes to one have more or less affect on changes in others.

[xxxix] Other researchers (for example Mckeown and Sheehy 2006) have added media processes to the Deffant *et al.* model but none of their outcomes look anything like BSAS data either.

[xl] There is an equal chance that a campaign will be positive or negative and an equal chance that the media outlet will campaign for any of the attitudes in the corresponding range (-100 to -26, -25 to 25 and 26 to 100).

[xli] Other models have extended the Deffuant *et al.* model to consider networks (for example Stauffer *et al.* 2004) but without reference to either validation or calibration data, as already discussed, it is not clear in what scientific sense these models are better rather than just different. Stauffer *et al.* (2004) also cite no empirical research directly.

[xlii] The ABM is written in NetLogo. This can be downloaded free from <http://ccl.northwestern.edu/netlogo/>. A documented version of the code can be downloaded from the openABM web site <http://www.openabm.org>.

[xliii] The news process is clearly important to the qualitative similarity with BSAS data as one might expect for something making a closed system into an open one. However, when operating alone, far too much attitude variation occurs (30-40%) and no combination of parameter values can be found to reduce this variation significantly. This suggests that it is a *combination* of social processes which gives rise to the BSAS data as one might expect.

[xliv] This period was chosen because it was the only BSAS attitude discussed in this article where continuous data was available over a ten-year period. The comparison was started arbitrarily at the earliest point where continuous death penalty data existed.

[xlv] This raises an interesting technical point. Although a certain amount of variation can be washed out in annualised data, there should be occasions when sharp changes coincide with annual data collection and if anything *overplay* the amount of change that has occurred. Can we determine, by looking at the structure of annualised data, how likely it is that the underlying daily data had a particular degree of smoothness? (We can imagine very smooth simulated data going through the observed data points but also very non-smooth data that could still do so.)

[xlvi] Another example would be the modelling of extremism. In the RAI model, an extremist has an extreme attitude and is not influenced by social interaction and that is all that is possible within the context of the model assumptions. In the model presented here, it is reasonable to ask further empirically accessible questions. For example, is an extremist likely to be extreme in all attitudes or only some? Is someone who is immune to social influence also immune to media influence? Are only extreme attitudes immune to influence or is it possible to be a fanatical moderate? Models that are oversimplified relative to existing social science knowledge render such obvious questions invisible.

[xlvii] Note that this still isn't fitting a model. I am trying to identify the social processes involved in behaviour that *might* lead to the observed patterns. Actual calibration and validation remains to take place. At this stage all I am showing, for example, is that the RAI model isn't even roughly right for real attitude data or that social influence based on experiments like that of Hovland *et al*. will produce different dynamics in the system so it matters if we use empirical data or not.

[xlviii] There are other possibilities. Perhaps memory and the wish to seem consistent limit the rate at which individuals can change their attitudes. Perhaps intergenerational change and imperfect socialisation acts as a source of continuing divergence. But the key insight is to think about to the extent to which social processes are likely to bring everyone to the same attitude (whatever it is) and the extent to which they might maintain non-trivially divergent attitudes (not in isolated populations, cycling arbitrarily or simply initialised amongst a population which isn't susceptible to social influence).

[xlix] There are some interesting issues here worthy of further investigation. As a possible genuine innovation in methodology (i. e. neither an offshoot of qualitative or quantitative research), ABM faces the challenge of re-

evaluating large quantities of pre-existing data. This raises new technical challenges: What should be concluded for the design of ABM from past qualitative research on any given topic and how should those conclusions be justified as systematic? At the same time, are there areas of research that only justify themselves to other converts (social networks matter) and have never had to demonstrate that adding a network element to a theory achieves something that cannot plausibly be achieved in other ways? It may be that only once one has a technique that allows different kinds of data to be synthesised that questions like "how much detail matters in social systems?" (a key bone of contention between qualitative and quantitative approaches) can even be clearly asked, let alone answered. We may find that in some social processes, networks have not been shown to matter because, in fact, they don't. An example of this would be finding that only news matters and that there is no point in building an ABM, instead just adding a variable of news emphasis to more traditional variables in regression.

[1] This has to be something of an article of faith in ABM. Critics of the approach often say: "How can you be sure that the fit you got could not have been produced by another completely different kind of model?" This possibility is called equifinality and, even though the criticism is coherent, I have never seen any critic bother to prove that a system of socially plausible complexity actually does display this property. (As I already said, we are well aware that simple Schelling models are equifinal but equally that they are not serious contenders for socially plausible complexity.) I am not sure it is fair to put the burden of proof to show such a thing cannot happen onto ABM. One presumes, after all, that the same point could be made about statistical models.

References

BERTAUX, D. and Thompson, P. (Eds.) (1997) *Pathways to Social Class: A Qualitative Approach to Social Mobility*. Oxford: Clarendon Press.

CHATTOE-BROWN, E. (2010) 'Is Simulation Forgetting Its History? Two Case Studies', Draft Paper, Department of Sociology, University of Leicester, <https://leicester.academia.edu/EdmundChattoeBrown>

CHATTOE-BROWN, E. (2013) ‘Why sociology should use agent-based modelling’, *Sociological Research Online*, Vol. 18, No. 3, < https://www.socresonline.org.uk/18/3/3.html>.

COLEMAN, J. (1990) *Foundations of Social Theory*. Cambridge, MA: Belknap Press/Harvard University Press.

DEFFUANT, G., Amblard, F., Weisbuch, G. and Faure, T. (2002) ‘How can extremism prevail? A study based on the relative agreement interaction model’, *Journal of Artificial Societies and Social Simulation*, Vol. 5, No. 4, <http://jasss.soc.surrey.ac.uk/5/4/1.html>.

DEFFUANT, G. (2006) ‘Comparing extremism propagation patterns in continuous opinion models’, *Journal of Artificial Societies and Social Simulation*, Vol. 9, No. 3, <http://jasss.soc.surrey.ac.uk/9/3/8.html>.

EGAN, P. and Mullin, M. (2012) Turning personal experience into political attitudes: The effect of local weather on Americans’ perceptions about global warming’, *The Journal of Politics*, Vol. 74, No. 3, pp. 796-809.

EPSTEIN, J. (2007) *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press.

EVANS, G. and Durant, J. (1995) ‘The relationship between knowledge and attitudes in the public understanding of science in Britain’, *Public Understanding of Science*, Vol. 4, No. 1, pp. 57-74.

FLEISHMAN, J. (1986) ‘Types of political attitude structure: Results of a cluster analysis’, *Public Opinion Quarterly*, Vol. 50, No. 3, pp. 371-386.

FORRESTER, J. (1971) *World Dynamics*. Cambridge, MA: Wright-Allen Press.

GILBERT, N. and Troitzsch, K. (2005) *Simulation for the Social Scientist*, second edition. Buckingham: Open University Press.

GODE, D. and Sunder, S. (1993) 'Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality', *Journal of Political Economy*, Vol. 101, No. 1, pp. 119-137.

GOLDSMITH, D. (1997) *Einstein's greatest blunder? The cosmological constant and other fudge factors in the physics of the universe*. Harvard, MA: Harvard University Press.

GRUSKY, D. and Hauser, R. (1984) 'Comparative social mobility revisited: Models of convergence and divergence in 16 countries', *American Sociological Review*, Vol. 49, No. 1, pp 19-38.

HATNA, E. and Benenson, I. (2012) 'The Schelling model of ethnic residential dynamics: Beyond the integrated - segregated dichotomy of patterns', *Journal of Artificial Societies and Social Simulation*, Vol. 15, No. 1, <http://jasss.soc.surrey.ac.uk/15/1/6.html>.

HOVLAND, C., Harvey, O. and Sherif, M. (1957) 'Assimilation and contrast effects in reactions to communication and attitude change', *Journal of Abnormal and Social Psychology*, Vol. 55, No. 2, pp. 244-252.

HUCKFELDT, R., Johnson, P. and Sprague, J. (2004) *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge: Cambridge University Press.

JACOBMEIER, D. (2006) 'Focusing of opinions in the Deffuant Model: First Impression Counts', *International Journal of Modern Physics C: Computational Physics and Physical Computation*, Vol. 17, No. 12, pp. 1801-1808.

MALARZ, K., Gronek, P. and Kułakowski, K. (2011) 'Zaller-Deffuant model of mass opinion', *Journal of Artificial Societies and Social Simulation*, Vol. 14, No. 1, <http://jasss.soc.surrey.ac.uk/14/1/2.html>.

MOSS, S. (2008) ‘Alternative approaches to the empirical validation of agent-based models’, *Journal of Artificial Societies and Social Simulation*, Vol. 11, No. 1, <http://jasss.soc.surrey.ac.uk/11/1/5.html>.

MULLAINATHAN, S. and Washington, E. (2009) ‘Sticking with your vote: Cognitive dissonance and political attitudes’, *American Economic Journal: Applied Economics*, Vol. 1, No. 1, pp. 86-111.

OHLANDER, J., Batalova, J. and Treas, J. (2005) ‘Explaining educational influences on attitudes toward homosexual relations’, *Social Science Research*, Vol. 34, No. 4, pp. 781–799.

O’LESSKER, K. (1968) ‘Who voted for Hitler? A new look at the class basis of Naziism [sic]’, *American Journal of Sociology*, Vol. 74, No. 1, pp. 63-69.

SCHELLING, T. (1971) ‘Dynamic models of segregation’, *The Journal of Mathematical Sociology*, Vol. 1, No. 2, pp. 143-186.

MCKEOWN, G. and Sheehy, N. (2006) ‘Mass media and polarisation processes in the bounded confidence model of opinion dynamics’, *Journal of Artificial Societies and Social Simulation*, Vol. 9, No. 1, <http://jasss.soc.surrey.ac.uk/9/1/11.html>.

SIRAJ, A. (2009) ‘The construction of the homosexual ‘other’ by British Muslim heterosexuals’, *Contemporary Islam*, Vol. 3, No. 1, pp. 41-57.

SKIENA, S. (2001) *Calculated Bets: Computers, Gambling, and Mathematical Modeling to Win!* New York, NY: Cambridge University Press/Mathematical Association of America.

Stauffer, D., Sousa, A. and Schulze, C. (2004) ‘Discretized opinion dynamics of the Deffuant model on scale-free networks’, *Journal of Artificial Societies and Social Simulation*, Vol. 7, No. 3, <http://jasss.soc.surrey.ac.uk/7/3/7.html>.

ST. GEORGE, A. and Robinson-Weber, S. (1983) ‘The mass media, political attitudes, and behavior’, *Communication Research*, Vol. 10, No. 4, pp. 487-508.

TASHAKKORI, A. and Teddlie, C. (Eds.) (2010) *Sage Handbook of Mixed Methods in Social and Behavioral Research*, second edition. Thousand Oaks, CA: Sage.

WHITE, G. (1975) ‘Contextual developments of opinion judgments: Field experimental probes of judgmental relativity boundary conditions’, *Journal of Personality and Social Psychology*, Vol. 32, No. 6, pp. 1047-1054.

WINDRUM, P., Fagiolo, G. and Moneta, A. (2007) ‘Empirical Validation of agent-based models: Alternatives and prospects’, *Journal of Artificial Societies and Social Simulation*, Vol. 10, No. 2, <http://jasss.soc.surrey.ac.uk/10/2/8.html>.

WOOD, W. (2000) ‘Attitude change: Persuasion and social influence’, *Annual Review of Psychology*, Vol. 51, pp. 539–570.