Measuring the deliberative quality of an online experimental mini-public: methodology and early results

Measuring the 'deliberative quality' of an online experimental mini-public:

methodology and early results[1]

Graham Smith*, Corinne Wales*, Peter John+, Sarah Cotterill+, Pat Sturgis*,

Gerry Stoker* and Hisako Namura+

* University of Southampton

+ University of Manchester

gsmith@soton.ac.uk

Paper prepared for ECPR General Conference, Potsdam, September 2009

'Experiments in Deliberative Mini-publics' panel

Abstract

This paper reports the evolving analysis strategy and some early results from a large-scale randomised controlled trial (RCT) that aims to explore the deliberative quality of internet engagement. The RCT divided 6009 participants into deliberation, information-only and control groups allowing us to investigate significant aspects of online democratic engagement. First it allows us to explore the extent to which moderated asynchronous discussion of policy issues (in this case youth anti-social behaviour and social cohesion) leads to reasoned shifts in policy preferences. Second, the design of the experiment provides an occasion to judge the relative impact of informed interaction between citizens as compared to individual reflection on information. Finally, a combination of quantitative analysis of the changes in policy preferences and use of the site and qualitative analysis of contributions to threads allows us to assess the extent to which interactions between citizens in this online environment can be understood as a form of democratic deliberation. More broadly, the experiment provides empirical insights that can inform contemporary debates on the desirability and effectiveness of internet-based participation in political decision making.

Introduction

Interest in mini-publics has gathered pace in recent years. From early experiments in the 1970s with citizens' juries in the US and planning cells in Germany, we now find a variety of designs utilised in the policy process: citizens' juries and planning cells have been joined by consensus conferences and deliberative polls, and in the last few years arguably the most impressive step change in practice, the citizens' assembly model instituted in British Columbia and Ontario. The arguments for engaging citizens through mini-publics are familiar, with particular emphasis placed on creating an inclusive environment within which a representative sample of citizens are enabled to form considered judgements on the issue at hand (Gastil and Levine 2005; Warren and Pearse 2008; Smith 2009). However, critics – and this often includes policy makers – question the legitimacy of designs that involve relatively small numbers; even the designs that engage in excess of 100 participants (planning cells projects, deliberative polling, citizens' assemblies). This is where developments in information and communication technology (ICT) could in principle challenge traditional barriers to citizen participation associated with time and space: in a mini-public that utilises an asynchronous internet-discussion forum citizens can engage at their own pace and participation can be 'scaled-up'. Large numbers can be involved in political deliberations without incurring the costs of physically bringing people together.

We have limited knowledge, however, of what happens when political engagement moves online, particularly in the form of a mini-public. James Fishkin and his colleagues have experimented with an online version of their deliberative poll model (Iyengar et al. 2005; Luskin et al. 2006), but they have made use of voice-operated software in an attempt to ensure that proficiency in ICT does not affect deliberations. Their results are interesting: for example they find that participants' policy knowledge and preferences tend to move in the same direction as for participants in face-to-face equivalents, but that 'changes from online deliberation were less pronounced than in the face-to-face version' (Ackerman and Fishkin 2004: 117; for more detail, see Luskin et al. 2006: 17-23). However, it is our contention that the online deliberative poll (ODP) is unlikely to be the design of choice for public authorities interested in running online mini-publics since it is relatively time-consuming and requires particular software. It is more realistic to believe that if public authorities are to make widespread use of online mini-publics they are more likely to use platforms with which they are already familiar and which can engage large numbers: asynchronous discussion forums being the most likely. ODPs tell us nothing about behaviour and outcomes in such an environment. The use of asynchronous technology to 'scale-up' engagement has (until now) not been explored systematically. It is such an experimental design that tests the efficacy of an asynchronous online mini-public – and in particular our approach to data analysis – that this paper reports.

In analysing asynchronous engagement, we need to be aware of differences between online and face-to-face (offline) worlds that are likely to have an effect. The dynamics of online deliberation differ from face-to-face interactions in at least two significant ways. First, we lose almost all the paralinguistic phenomena associated with interaction. Non-verbal communication plays a significant role in conveying meaning and emotion and its loss is likely to affect the degree of mutual understanding between participants. Second, in the face-to-face world, participants are typically brought together in mini-publics for a number of days. Asynchronous online forums may be open for long periods, but they allow participants to contribute to discussion threads whenever they wish: the participation demands placed on online participants in deliberations are not so high. Arguably it is easier to construct face-to-face mini-publics to promote equality of voice amongst participants (e.g. through active facilitation). If participants are able to choose when (and indeed if) to log-on and to post, it would appear more difficult to ensure the same degree of inclusiveness online.

This paper is primarily methodological in that it discusses the approach we are taking to the analysis of the online experiment. Having offered a description of the experiment, we explain how we are focusing on both outcome and process to evaluate the extent to which online engagement in an asynchronous internet discussion forum can be classed as 'deliberative'.

Where data is available (we are only in the early stages of analysis), we offer some initial findings.

The experimental design

The research team worked with Ipsos-MORI to develop a novel internet-based experiment – a randomised-controlled trial (RCT) – to assess the potential of large-scale asynchronous online mini-publics. The experiment involved 6,009 participants, the sample drawn from the Ipsos-MORI's survey panel that regularly engage in market research (although none would have been invited to take part in research of this type before).

When accepting the invitation, the 6,009 participants were informed that they would be required to complete three relatively short online surveys over three weeks and then may be asked to undertake other tasks. All participants were offered an incentive to complete the survey (entry into a prize draw). When they accepted, they were invited to complete the first survey (T1) which included some basic socio-demographic details and a first wave of questions on youth anti-social behaviour (ASB) and community cohesion. Quota sampling was employed using early questions on the survey instrument to ensure that the 6,009 selected participants were broadly representative in terms of age, gender, education, geographical location and political interest. Given the nature of the panel, we had to accept that the number of non-white participants would be below the national average and

that all participants had some ICT competence; although experience varied

widely with many having little or no experience of internet forums. Having

selected the sample, it was only at this point that the participants were

randomly assigned to one of six groups: four treatment and two control

groups.

The first two treatment groups are termed the <u>deliberation groups</u>. At

this point we use the term 'deliberation' in a relatively broad sense: the

conditions were in place for the free and fair exchange of views between all

participants, with a moderator ensuring civility in interactions. Participants in

these groups were invited to take part in an online forum: the first group on

youth ASB ('The truth about youth'); the second on community cohesion

('Getting on together'). The rationale of running two groups was to

investigate the extent to which participants behaved differently online when

initially confronted with a controversial topic where flaming was more likely

(i.e. community cohesion). The forums were hosted on specially

commissioned phpBB 3.0.x boards, with engagement incentivised with

additional entries into a prize draw each time participants spent 10 minutes or

more online.[2] When logging in for the first time, participants were greeted

with a video from the then Secretary of State for Communities and Local

Government, Hazel Blears, who gave her support for the project and

[2] It was important to design an incentive that could be used easily by public authorities and did not involve extravagant expense. Mini-publics typically offer an honorarium of around £100 to participants. This would be too much of a cost impediment for such large numbers.

committed herself to consider the issues raised by participants. Every two days new threads (topics) were launched on different issues with accompanying information (a combination of written and video materials), starting with discussions of participants' own experience, through to discussions of government policy and proposals. After 10 days, a second survey was administered (T2) that focused on the topic participants had been discussing. At this point, the threads were closed and the groups switched topic: the youth ASB group moved on to discuss community cohesion and vice versa. After another 10 days a third survey (T3) was administered based on this second topic. At this point the experiment ended and participants were thanked for their participation.

The third and fourth treatment groups are termed the <u>information-only groups</u>. They were treated in the same manner (including surveys and incentives) with one exception. The phpBB 3.0.x boards were the same, <u>except</u> that while participants were able to post, they were <u>unable</u> to see the postings of other participants. This design feature enables us to investigate the 'added value' of engaging with other participants: the value of online deliberation as compared to only offering information for reflection.

Finally, the two <u>control groups</u> (groups five and six) were simply surveyed at the same three points as the deliberation and information-only groups. See Figure 1 for the design of the experiment.

Figure 1: the design of the experiment

Delib 1	Delib 2	Info 1	Info2	Control1	Control2

T1 survey --

Youth ASB	Cohesion	Youth ASB	Cohesion	Youth ASB	Cohesion

T2 survey (youth or cohesion issues) --

Cohesion	Youth ASB	Cohesion	Youth ASB	Cohesion	Youth ASB

T3 survey (youth or cohesion issues) ---
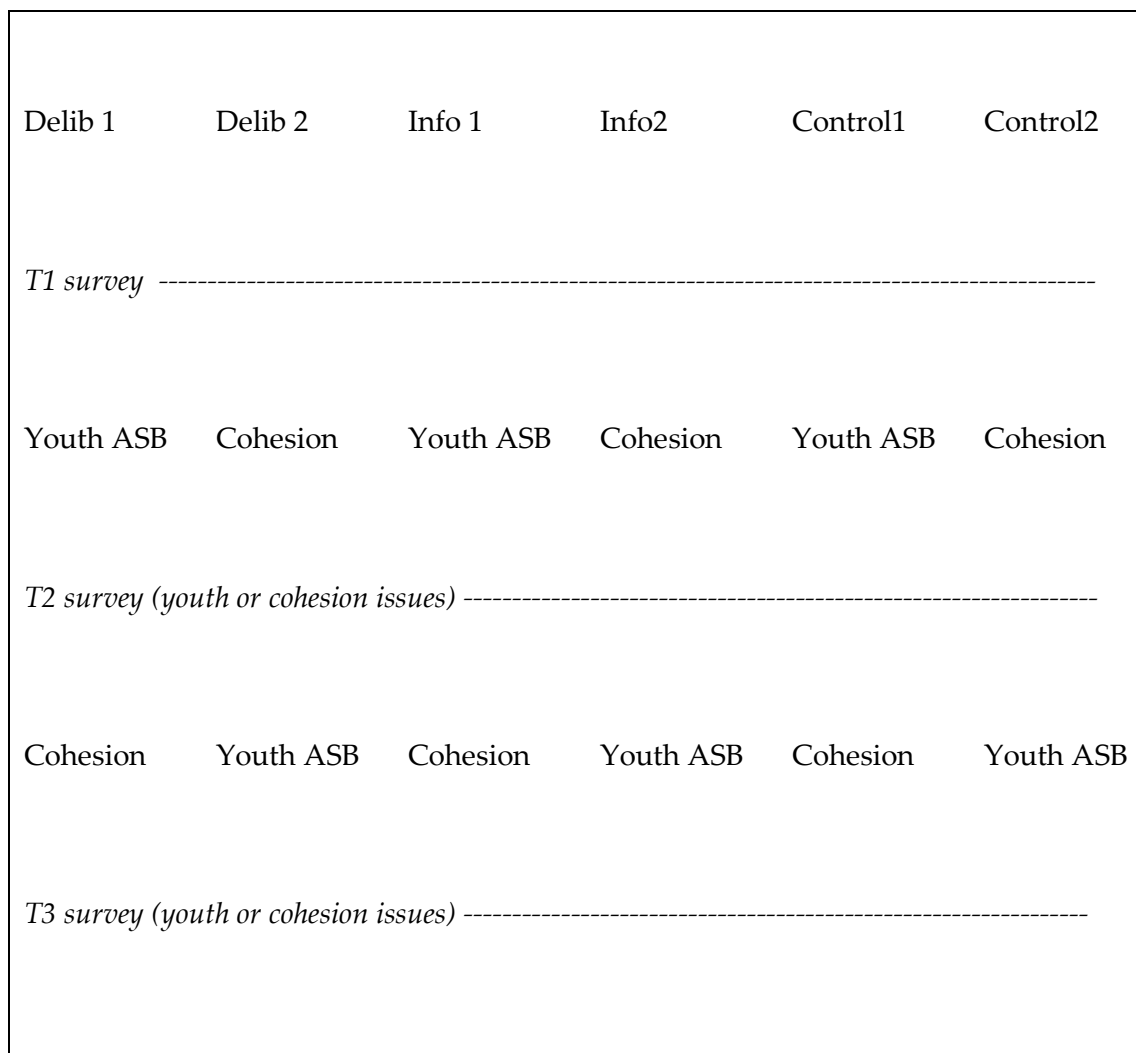
An *outcome* approach to judging deliberative quality: the experiment as a randomised controlled trial (RCT)

Deliberative democrats share the belief that the process of public deliberation between free and equal citizens can have a transformative effect on individuals' preferences. Under conditions of mutual respect, democratic deliberation provides the motivation to articulate preferences that are 'public spirited' in nature. David Miller, for example, stresses the 'moralising effect of

public discussion': the reciprocal requirement to put forward reasons and to respond to challenges that eliminate irrational preferences based on false empirical beliefs, morally repugnant preferences that no one is willing to advance in the public arena, and narrowly self-regarding preferences (Miller 1992: 61). He continues: 'we have good reason to expect the deliberative process to transform initial policy preferences (which may be based on private interests, sectional interest, prejudice and so on) into ethical judgements on the matter in hand' (ibid: 62). Evidence from mini-publics suggests that more considered judgements often do emerge amongst participants in practice (Smith 2009: 94-101) and it is the outcome measure of changes in policy opinion and knowledge – and the extent to which there is some consistency in any shifts – that Fishkin and others have used to judge deliberative quality of mini-publics.

Given the loss of paralanguage and the expectation that participants will spend less time online than in face-to-face mini-publics, it is reasonable to hypothesise that any effect size of shifts in opinion and knowledge will be relatively small when compared to face-to-face forums. However, our experiment is designed to involve large numbers; to 'scale-up' the opportunity to engage. Given the large sample size, we will be in a position to witness any significant shifts in opinion and knowledge and to judge their consistency.

There is another difference between the online and face-to-face worlds that creates a challenge for analysis: what counts as participating in online deliberation? This is significant when it comes to comparing results across groups. The simplest approach to the data is an intention-to-treat (ITT) analysis: comparing results from the completed questionnaires across the deliberation, information-only and control groups. But we can reasonably expect (based on experience from other RCTs) that we will be faced with the problem of non-compliance: although members of the treatment groups were invited to participate in online activities, this does not mean that they accepted the invitation. We need to carry out comparisons with the control groups where such self-selection has not taken place.

But where compliance with interventions is generally easy to define, it is not so simple in this type of internet engagement. Who has engaged? The most obvious answer is posting: deliberation here is defined as actively contributing to one or more threads by posting at least once during the experiment. But this may not be an accurate reflection of online deliberation as understood by many deliberative democrats. Deliberation is not just about talking, but also listening and reflecting on contributions. Playing a spectator role, while others make contributions (which may well reflect your own position), can also be understood as participating in deliberation. After all, where large numbers are involved, not everyone can speak. It would thus seem strange to define only the speakers as those who have participated in

the deliberation. In the online world, such spectators are referred to as 'lurkers' (Jansen and Kies 2005: 331). If we wish to include such lurkers within the definition of compliance, then a more suitable definition would include all participants who <u>login</u> to the boards, irrespective of whether they then post. Thus in analysing the data we will investigate the effects relating to three compliance groups: ITT, LOGIN and POST

The non-compliance problem not only creates this definitional question, but also a technical problem of how to compare the compliance groups to the control. While the randomisation ensures that the intention-to-treat and control groups have broadly similar characteristics, we cannot assume that the characteristics of the compliers in the deliberation and information-only group will be the same. Arguably the most popular approach to dealing with the non-compliance problem in RCTs is propensity score matching (PSM) which in principle allows us to match compliers with an identical population in the control group. But there are two (interrelated) problems with applying PSM to our data. First, we find considerable differences in results depending upon the variables selected for the matching process. Second, although large for a deliberation experiment, the population size is relatively small for effective PSM (Zhao 2004).

Given the limitations of PSM, we instead utilise the Complier Average Causal Effect (CACE) model, a latent variable approach to the estimation of experimental treatment effects in the presence of non-compliance (Jo and

Muthen 2001). A latent class (or finite mixture) framework is used to estimate the compliance status of those in both the treatment and the control conditions, allowing 'fair' comparisons to be made between compliers in the treatment condition and 'potential compliers' in the control group. The basic problem in this regard is, of course, that in the control condition compliance status is unobserved and, without additional assumptions, the model is unidentified. The CACE model achieves identification of the latent compliance classes through application of the 'exclusion restriction' (Angrist et al 1996). The exclusion restriction in the context of non-compliance relates to the actual and potential behaviour of experimental subjects and how this is associated with the outcome. Experimental subjects can be categorized as falling into one of the following 4 categories: (1) compliers – those who comply with experimental procedures; (2) never-takers – those who do not take the treatment when assigned to the treatment condition; (3) always-takers – those who take the treatment even if allocated to the control condition; (4) defiers – those who do the opposite of what they are assigned to do. The exclusion restriction assumes that there are no defiers and no always-takers. In addition to the exclusion restriction, identification of the latent non-complier class is aided by the incorporation of covariates which are predictive of non-compliance.

While this approach to the analysis of responses to the survey instrument allows us to investigate the extent of opinion and knowledge shift,

the design of the experiment also provides a rare occasion to use this method to compare the effect of deliberation as compared to the information-only group. Research designs with information and discussion sandwiched between pre- and post-tests do not usually yield the necessary data to determine whether shifts occurred because of interaction between participants or because of information exposure, or some combination of the two (Muhlberger and Weber, 2006: 4). The two different types of treatment group in our experiment allow us to investigate the value of deliberation: interaction as compared to access to information.

A *process* approach to judging deliberative quality: mixing quantitative and qualitative methods

We have used the term 'deliberation' to differentiate two of the treatment groups. As such we can be accused of pre-judging our findings. These groups may have engaged in asynchronous internet discussion forums, but to what extent can that engagement be classed as 'deliberative' as that term is understood in deliberative democratic theory? To what extent is interaction inclusive and based on reason-giving in the context of reciprocity and mutual respect? This is where we move from the analysis of the experiment as RCT, to a combination of quantitative data relating to the use of the boards and more qualitative content analysis of threads and postings.

A rare candidate for evaluating deliberative quality in a systematic sense is the Discourse Quality Index (DQI) developed by Marco Steenbergen and colleagues (2003). Normatively grounded in Habermasian discourse ethics, the DQI was created to evaluate the quality of parliamentary debates. However, based on our experience of attempting to apply the DQI on a pilot for our online experiment, we found that in a number of important respects its transferability across different empirical conditions (in our case internet engagement) is bounded by this initial research focus. Nonetheless, the DQI has provided a useful benchmark against which our own thinking about coding and analysis has developed.

In evaluating the deliberative quality of our asynchronous internet mini-public we will primarily be taking individual posts as the unit of analysis. Having previously piloted a more complex coding framework, we propose to restrict ourselves to the following categories in our analysis:

- Inclusiveness of participation

- Form and content of justification

- Respect towards groups discussed

- Respect / reciprocity amongst participants

Inclusiveness of participation

In the hands of deliberative democrats, inclusiveness turns our attention to both presence and voice. The sample from the Ipsos-MORI panel will by

definition not include those who suffer from one aspect of the 'digital divide' (Norris 2001): lack of access to the internet. However, care was taken to ensure that participants were selected with different levels of internet use as a proxy of internet competence. While participants had an equal opportunity to participate in terms of posting, these differences in experience within the sample of using the internet – and in particular online forums – as well as more traditional (non-digital) differences in, for example, education and political interest, may well affect participation rates.

A second indicator of inclusiveness is the presence of competing discourses in threads, particularly since the subjects of the discussion forums are controversial (youth ASB and, in particular, community cohesion). After all, '[s]ome basic disagreement is necessary to create the problem that deliberative democracy is intended to solve' (Thompson: 2008: 502). The presence of different perspectives in the discussion threads indicates that those participating (those who have chosen to post) are not a homogeneous group, a condition that can lead to polarisation and less open-minded examination of viewpoints (Price et al 2002; Sunstein, 2000).

To understand the extent to which inclusiveness is realised, we intend to code for:

- The number, frequency and patterns of logins and posts against socio-economic characteristics and internet usage (both overall and on selected threads)

- The presence or absence of competing discourses (on selected threads)

Form and content of justification

In participating in discussions, a number of different types of claims can be made to justify statements. Competing accounts of deliberative democracy offer alternative readings of the status of justifications: for example, difference theorists (e.g. Young 1990; Sanders 1997) are much more sympathetic to appeals to personal experience. The Habermasian roots of Steenbergen and colleagues' DQI and (arguably) their focus on parliamentary debates led them to code for politically sophisticated modes of justification, where (for example) personal experience does not count for very much. This may be appropriate to arenas where finely honed deliberative skills and capacities are expected (Dutwin 2003), but is arguably inappropriate for an arena in which many participants may not have deliberated before online or indeed offline. Our aim is to understand the extent and manner in which online participants offer evidence or reasons to support the claims they are offering. For this reason, we intend to investigate the diversity of sources of justification, coding for the following sources:

- None offered

- Personal experience

- Briefing materials

- External authorities

- Other participants' posts

- Moderator posts

In analysing the content of justification, Steenbergen and his colleagues distinguish between appeals to self- or group-interest and appeals to the common good, dividing the latter between utilitarian and Rawlsian understandings (2003; 28-29). These distinctions may be possible to utilise in analysing parliamentary debates where speeches are often carefully crafted, but proved difficult to implement in our pilot. We remain interested in the extent to which appeals to the common good are made, but will take a more inductive approach, coding those posts where an appeal is made to a particular group or society at large and then drawing out any relevant classifications.

Respect towards groups discussed

In selecting the two broad subjects for the forums – youth ASB and, in particular, community cohesion – we are interested in the extent to which there is empathy towards the groups under discussion. Our experiment is a strong test of such respect since in both cases, members of the groups discussed – young people and minority ethnic groups – are likely to be under-represented in the forums. In particular, the subject of community cohesion is more likely to generate what internet users term 'flaming' – offensive

comments. Our pilot indicated the difficulties involved in coding 'respect' towards groups. As such we intend to code for:

- Explicit disrespect towards groups discussed (e.g. youth, minority ethnic groups)

Respect / reciprocity between participants

Mutual respect and reciprocity between participants is fundamental to deliberative democracy. There is widespread concern that in the online world, participants are often less constrained and political discussions – particularly on controversial issues – can degenerate with excessive 'flaming' and other forms of incivility (e.g. Docter and Dutton 1998). However concern tends to relate to synchronous chat rooms and Usenet groups. As Stephen Coleman notes: 'In Usenet political discussions, people talk past one another, when they are not verbally attacking each other. The emphasis is not problem solving, but discussion dominance' (Coleman 2004: 6; see also Sack 2005: 268). The design of our experimental mini-public differs markedly from such forums – in particular the use of asynchronous interaction and moderation – and is thus more likely to promote more respectful exchange (Janssen and Kies 2004: 4-5). As Scott Wright and John Street suggest: 'the democratic possibilities opened up (or closed off) by websites are not a product of the technology as such, but of the ways in which it is constructed, by the way it is designed' (Wright and Street 2007: 850).

Again, we have to be aware of the difference between the highly stylised and often well-prepared parliamentary debates between practiced rhetoricians that were the subject of Steenbergen and his colleagues' analysis and the form of interaction that takes place on discussion boards. In evaluating the extent of respect towards other participants, we will focus again on the counterfactual:

- Explicit disrespect towards other participants

Reciprocity is included in our analysis because there is a danger in the online world that participants will simply talk past each other; in other words, post without even reading the contributions of others. We are thus interested in the extent to which participants who post are engaging with one another and considering the claims of others, rather than simply making statements in response to the questions asked by the facilitators. Additionally we will investigate whether reciprocity tends to be shown primarily to those with whom the poster agrees. As such we intend to code each post for:

- No response to other posts

- Explicit agreement with other post(s)

- Explicit disagreement with other post(s)

- Non-explicit agreement with other post(s)

- Non-explicit disagreement with other post(s)

Initial findings

As the narrative of the paper suggests, we have yet to undertake the content analysis of threads, so initial findings are limited to some early quantitative analysis of shifts in policy preferences and the inclusiveness of engagement. At present this analysis has only been carried out on a couple of sets of questions related to perception of youth ASB and potential policy solutions (see Appendix 1). We do not consider the ordering of the issues, thus generating three groups of 2,000 (deliberation; information-only; control).

Shifts in policy preference

An analysis of intention-to-treat across the three groups (deliberation, information-only and control) indicates differences on only a very small number of questions. Shifts appear random (there is no obvious consistency between questions where there is some change) and the degree of change means that it is probably down to chance. In other words, we find <u>no systematic differences</u> in opinion and knowledge across the three groups for the ASB questions.

However, when we take into account the problem of non-compliance and apply the CACE model to compliers defined as LOGIN and POST, the findings are different. Table 1 offers a comparison of the results of the CACE model (with compliance defined as LOGIN and POST) compared to the ITT for the two sets of youth ASB questions. We provide the mean effect score and

standard deviation for each question for the deliberation group (with compliance defined as LOGIN and POST) in comparison to the control along with the average effect for the two sets of questions.

Table 1: Effect of deliberation

variable name	control-delib ITT	control-delib (CACE-login)	control-delib (CACE-post)
asb20	-0.009 (0.062)	-0.016 (0.106)	-0.024 (0.224)
asb21	0.093 (0.061)	0.189 (0.114)	0.489 (0.225)*
asb22	0.063 (0.063)	0.134 (0.113)	0.260 (0.277)
asb23	0.097 (0.063)	-0.022 (0.033)	0.357 (0.237)
Average effect for perception of ASB questions	0.066	0.090	0.443
asb27	0.101 (0.055)	0.341 (0.112)*	0.917 (0.234)*
asb28	0.069 (0.053)	0.177 (0.108)	0.351 (0.431)
asb29	0.005 (0.052)	-0.066 (0.517)	-0.321 (0.140)*
asb30	-0.001 (0.051)	0.148 (0.108)	0.622 (0.172)*
asb31	0.009 (0.051)	0.175 (0.109)	-0.417 (0.102)*
asb32	0.096 (0.070)	0.227 (0.119) (p=.058)	0.450 (0.233)(p=.053)
asb33	0.017 (0.070)	0.028 (0.117)	0.105 (0.227)
asb34	-0.107 (0.069)	-0.148 (0.116)	-0.200 (0.208)
asb35	-0.026 (0.070)	0.035 (0.126)	0.245 (0.250)
Average effect for	0.049	0.149	0.403

policy questions			

When comparing the three definitions of compliance – intention-to-treat (ITT), those where compliance is taken to be logging into the site (LOGIN) and those where compliance is taken to be posting at least once (POST) – there is a discernable pattern of opinion change. In all but two of the questions there is a consistent increase in effect size as we move from ITT to LOGIN to POST (ASB23 and ASB31 are the exceptions where LOGIN moves in a different direction). This finding is reinforced when we consider the magnitude of the effect across the two sets of questions. The average effect is very small for ITT, slightly larger for LOGIN and considerably larger for POST. Our first finding, then, is that actively contributing to deliberation in the form of posting has the most significant impact on opinion change. Lurking has little effect.

While effect size is important, on its own it tells us nothing about the consistency of opinions. If we focus on the nine policy questions and take into account effect size (POST minus ITT), then there does indeed seem to be consistency in opinion change with relevant questions grouping together (see Table 2): movement away from legal punishment (ASB27, ASB35) and heavier policing (ASB28) and towards better role models (ASB31) and monitoring by adults (ASB28); and movement away from providing activities (ASB30) or rewards (ASB32) to young people. The two questions which relate to policy options that are arguably least familiar to the public – voluntary parenting

classes (ASB34) and written rules of behaviour (ASB33) – show the least opinion change.

Table 2: Policy questions ranked according to effect size

Number	Question	Direction of change	Effect size
asb27	Apply strong forms of legal punishment for young people who misbehave.	Disagree	0.816
asb30	Provide more activities for young people	Disagree	0.623
asb31	Promote better role models for young people	Agree	0.426
asb32	Give rewards to young people who stop committing anti-social behaviour	Disagree	0.354
asb29	Better monitoring of young people by adults in the neighbourhood	Agree	0.316
asb28	Put more police on the streets	Disagree	0.282
asb35	Take legal action against parents of poorly-behaved young people	Disagree	0.271
asb34	Provide voluntary parenting classes	Agree	0.093
asb33	Establish written rules of behaviour for the neighbourhood	Disagree	0.088

Information or deliberation?

This finding that online deliberation (understood as posting) can generate consistent opinion shifts can be further interrogated by focusing on the results

from the information-only group. In this case we find little or no movement.

These results are presented in Table 3.

Table 3: Effect of information-only

variable name	control-info ITT	control-info (CACE-login)	control-info (CACE-post)
asb20	0.111(0.063)	0.066 (0.046)	0.101 (0.055) (p=.067)
asb21	-0.106(0.064)	-0.078 (0.047)	-0.088 (0.056)
asb22	0.178(0.065)*	0.137 (0.048)*	0.146 (0.059)*
asb23	-0.083(0.065)	-0.065 (0.048)	-0.075 (0.057)
Average effect for perception of ASB questions	0.120	0.087	0.103
asb27	0.046(0.057)	0.068 (0.056)	0.039 (0.169)
asb28	0.071(0.053)	0.077 (0.048)	0.052 (0.108)
asb29	0.013(0.053)	-0.089 (0.047) (p=.056)	0.065 (0.082)
asb30	0.012(0.053)	0.051 (0.049)	-0.113 (0.054)*
asb31	-0.026(0.053)	0.037 (0.046)	0.038 (0.065)
asb32	0.006(0.071)	0.009 (0.052)	-0.009 (0.062)
asb33	0.117(0.070)	0.080 (0.052)	0.086 (0.064)
asb34	-0.063(0.070)	-0.048 (0.053)	-0.070 (0.064)
asb35	0.054 (0.072)	0.034 (0.056)	0.037 (0.072)
Average effect for policy questions	0.045	0.054	0.057

For the information-only groups, average effect sizes are small and differences between ITT, LOGIN and POST are fairly negligible. Certainly the differences we witness for the deliberation group (Table 1) are not present: the average effect size for the POST compliance group is not significantly different from the LOGIN or ITT groups.

This comparison between deliberation and information-only groups suggests that interaction between participants has an effect on their preferences. However, it could be that the deliberation group used the support materials more than the information-only group and as such it was their greater use of information that led to the shifts we have uncovered. A simple way of testing the impact of interaction as compared to information use is to run a correlation based on whether participants who posted in the deliberation and information-only group used the support materials at least once during the life of the forums. The results are in presented in Table 4.

Table 4: Comparing the use of supporting materials for POST

Used supporting materials at least once	Deliberation POST	Information-only POST	Total
No	317 (60.27)	238 (37.01)	555 (47.48)
Yes	209	405	614

	(39.73)	(62.99)	(52.52)
Total	526	643	1,169
	(100)	(100)	(100)

Pearson chi2(1) = 62.7302   Pr = 0.000

This basic correlation shows that for participants who posted there is a significant differential in information use between the two treatment groups. Those in the information-only group (63%) are more likely to access the supporting materials compared to those in the deliberation group (40%). In some ways this should not be such a surprise since the information-only group would not be distracted by the postings of other participants. What these results do suggest then is that it is the interaction between participants in the deliberation group that is shifting opinions rather than the use of information: <u>active engagement</u> with other participants is having the effect.

Who deliberates? A digital divide?

While we have shown that participants who took the opportunity to deliberate (in the sense that they posted at least once during the experiment) were more likely to alter their opinions in relation to youth ASB, we do not know the characteristics of those who complied. To what extent do they replicate the characteristics of the more politically-engaged in the face-to-face

world? In other words, does online engagement reinforce political inequalities?

Table 5: Compliance across groups

	ITT	Completed survey	LOGIN	POST
Deliberation	2,004	1,409	1,073	526
Information	2,003	1,305	1,216	643
Control	2,002	1,622		

As Table 5 indicates, compliance drops-off as involvement in the experiment becomes more demanding on participants: hence the need to apply a CACE model to compare the different compliance groups with the control. Participation takes time and is not appealing to the majority of citizens. But who are the participants who actively engaged in the discussion forum? Table 6 provides the results of a cross-section logit regression for different forms of compliance, with the deliberation POST group highlighted.

Table 5 Factors that affecting compliance according the level of compliance

VARIABLES	deliberation Questionnaire completed	deliberation login	**deliberation post**	Information Questionnaire completed	Information login	Information post
Female	0.341***	0.196*	**0.480*****	0.077	-0.099	0.001
	(0.102)	(0.094)	**(0.108)**	(0.098)	(0.096)	(0.100)
age 25-34	0.018	0.333	**0.064**	0.250	0.318	0.166
	(0.179)	(0.176)	**(0.215)**	(0.184)	(0.185)	(0.211)
age_35-44	0.305	0.481**	**0.195**	0.456*	0.396*	0.436*

	(0.179)	(0.173)	**(0.208)**	(0.180)	(0.180)	(0.203)
age_45-54	0.470*	0.518**	**0.337**	0.833***	0.517**	0.478*
	(0.185)	(0.176)	**(0.210)**	(0.192)	(0.188)	(0.210)
age_55+	0.840***	0.723***	**0.579****	1.037***	0.679***	0.632**
	(0.174)	(0.164)	**(0.195)**	(0.177)	(0.174)	(0.195)
Non-white	-0.084	0.264	**0.288**	-0.306*	0.260	0.364*
	(0.167)	(0.160)	**(0.172)**	(0.151)	(0.155)	(0.155)
level 1 qualification	0.228	0.517*	**0.500**	-0.069	0.136	0.335
	(0.240)	(0.218)	**(0.263)**	(0.236)	(0.217)	(0.235)
level 2 qualification	0.258	0.487*	**0.576***	-0.219	0.234	0.107
	(0.216)	(0.197)	**(0.240)**	(0.211)	(0.195)	(0.214)
level 3 qualification	0.252	0.340	**0.363**	0.051	0.406	0.397
	(0.230)	(0.210)	**(0.258)**	(0.225)	(0.210)	(0.227)
level 4-5 qualification	0.333	0.524**	**0.575***	0.082	0.503*	0.301
	(0.210)	(0.191)	**(0.233)**	(0.212)	(0.196)	(0.212)
other qualifications	0.019	0.380	**0.496***	-0.218	0.161	0.130
	(0.223)	(0.205)	**(0.249)**	(0.217)	(0.200)	(0.219)
Frequent internet user	0.299	0.222	**0.503**	0.360	0.531**	0.309
	(0.214)	(0.207)	**(0.276)**	(0.190)	(0.185)	(0.212)
High political interest	-0.079	0.356***	**0.302***	-0.161	0.221*	0.378***
	(0.113)	(0.103)	**(0.114)**	(0.108)	(0.106)	(0.107)
Constant	-0.162	-1.181***	**-2.579****	-0.207	-0.985**	-2.037***
	(0.334)	(0.319)	**(0.404)**	(0.314)	(0.304)	(0.343)
Observations	2000	2000	**2000**	1999	1999	1999

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Two findings jump out. First, posters are significantly more likely to be female and have an already high political interest. They are also likely to be older and more highly qualified (although there is some discrepancy here given that level 2 qualification is also significant). The significance of gender and age suggest that available time (unsurprisingly) affects levels of engagement. In many ways the online world mirrors the face-to-face, with well-known variables such as age, qualifications and political interest explaining much of the variation in participation (Verba et al 1978). What is more unexpected, however, is that the frequency of internet use does not appear to affect levels of compliance. In this sense the online world does not

appear to create new disadvantage, but rather replicates and reinforces existing political inequalities.

Conclusions

Our experiment is designed to test the extent to which online engagement in an asynchronous discussion forum is 'deliberative' as understood within deliberative democratic theory. This requires a mixed methods approach to the analysis of the data. Our initial findings from the statistical analysis of the response to two sets of youth ASB questions suggest, first, that participation (defined as posting) does lead to consistent changes in policy preference. Second, it appears that these shifts are primarily driven by interaction between participants, rather than reflection on the background materials provided on the forum. Third, participation in the forum was a minority pursuit and, as such, reinforced traditional differentials in participation across social groups. While these initial findings suggest that online engagement can have an effect on outcomes, further statistical analysis of other ASB and community cohesion questions and content analysis of threads is required to understand whether such online engagement can be defined in any way as 'deliberative'. The inequalities in participation found thus far do not bode well.

References

Ackerman, B. and J. S. Fishkin (2004). <u>Deliberation Day</u>. New Haven, Yale University Press.

Angrist, J, Imbens, G. and Rubin, D. (1996) Identification of Causal Effects Using Instrumental Variables. <u>Journal of the American Statistical Association</u>, 91 444-445.

Barber, B. (1998). "Three Scenarios for the Future of Technology and Strong Democracy." <u>Political Quarterly</u> **113**(4): 573-89.

Beetham, D. (1999). <u>Democracy and Human Rights</u>. Cambridge, Polity.

Cederman, L.-E. and P. A. Kraus (2005). Transnational Communication and the European Demos. <u>Digital Formations: IT and New Architectures in the Global Realm</u>. R. Latham and S. Sassen. Princeton, Princeton University Press.

Coleman, S. (2004). "Connecting Parliament to the Public via the Internet: Two case studies of online consultations." <u>Information, Communication and Society</u> **7**(1): 1-22.

Coleman, S. (2005). Direct Representation: Towards a Conversational Democracy. London, IPPR.

Dahl, R. (1998). On Democracy. New Haven, Yale University Press.

Docter, S. and W. H. Dutton (1998). The First Ammendment Online: Santa Monica's Public Electronic Network Cyberdemocracy: Technologies, Cities and Civic Networks. R. Tsagarousianou, D. Tambini and C. Bryan. London, Routlege.

Dutwin, D. (2003) The Character of Deliberation: Equality, Argument and the Formation of Public Opinion, International Journal of Public Opinion Research Vol. 15 No. 3: 239-264

EOS Gallop Europe (2002). Flash Eurobarometer 135: Internet and the Public at Large. Brussels, EOS Gallop Europe.

Ferguson, R. (2006). Digital Dialogues: Interim Report, December 2005 - August 2006. London, Hansard Society / Department for Constitutional Affairs.

Ferguson, R., Ed. (2006). <u>TellParliament.net Interim Evaluation Report 2003-5</u>. London, Hansard Society.

Fung, Archon. "Survey Article: Recipes for Public Spheres: Eight Institutional Design Choices and Their Consequences." <u>Journal of Political Philosophy</u> 11, 3 (2003): 338-67.

Gastil, J. and P. Levine, Eds. (2005). <u>The Deliberative Democracy Handbook: Strategies for Effective Civic Engagement in the 21st Century</u>. San Francisco, Jossey-Bass.

Iyengar, S., R. C. Luskin, et al. (2005). "Deliberative Preferences in the Presidential Nomination Campaign: Evidence from an Online Deliberative Poll." <u>Paper available from The Center for Deliberative Democracy, Stanford University http://cdd.stanford.edu/research/index.html</u>.

Janssen, D. and R. Kies (2004). "Online Forums and Deliberative Democracy: Hypotheses, Variables and Methodologies." <u>Paper prepared for the Conference on 'Empirical Approaches to Deliberative Politics', European University Institute, Florence, 22-23 May 2004.</u> .

Janssen, D. and R. Kies (2005) "Online Forums and Deliberative Democracy", Acta Politica, **40**: 317–335

Jo, B. & Muthén, B. (2001). "Modeling of intervention effects with noncompliance: A latent variable approach for randomized trials". In G. Marcoulides & R.E. Schumacker (eds.) New Developments and Techniques in Structural Equation Modeling (pp. 57-87). Mahwah, New Jersey: Lawrence Erlbaum

Luskin, R. C., J. S. Fishkin, et al. (2006). "Considered Opinions on U.S. Foreign Policy: Face-to-Face versus Online Deliberative Polling, ." Paper available from The Center for Deliberative Democracy, Stanford University http://cdd.stanford.edu/research/index.html.

Norris, P. (2001). Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide. Cambridge, Cambridge University Press.

Pratchett, L. (2006). Understanding E-democracy Developments in Europe. Strasbourg, Council of Europe.

Price, V., J.N. Cappella and Nir, L. (2002). Does Disagreement Contribute to More Deliberative Opinion?". Political Communication **19**: 95-122

Sack, W. (2005). Discourse Architecture and Very Large-scale Conversation. Digital Formations: IT and New Architectures in the Global Realm. R. Latham and S. Sassen. Princeton, Princeton University Press.

Schlosberg, D., S. W. Shulman, et al. (2006). Virtual Environmental Citizenship: Web-Based Public Participation in Rulemaking in the United States. Environmental Citizenship. A. Dobson and D. Bell. Cambridge, MA, MIT Press**:** 209-236.

Smith, G. (2009). Democratic Innovations: Designing Institutions for Citizen Participation. Cambridge, Cambridge University Press.

Sunstein, C. R. (2000). "Deliberative Trouble? Why Groups Go To Extremes." Yale Law Journal **110**: 71-119.

Sunstein, C. R. (2001). republic.com. Princeton, Princeton University Press.

Thompson, D. F. (2008). "Deliberative Democratic Theory and Empirical Political Science." Annual Review of Political Science **11**: 497-520.

Warren, M. E. and H. Pearse, Eds. (2008). <u>Designing Deliberative Democracy: The British Columbia Citizens' Assembly</u>. Cambridge, Cambridge University Press.

Wright, S. and J. Street (2007). "Democracy, deliberation and design: the case of online discussion forums." <u>New Media and Society</u> **9**(5): 849-69.

Verba, S., N. Nie and J. Kim (1978). <u>Participation and Political Equality</u>. Cambridge, Cambridge University Press.

Zhao, Z. (2004). "Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence." <u>The Review of Economics and Statistics</u>, 86(1), 91-107