How Can You Think That?: Deliberation and

the Learning of Opposing Arguments

December, 2013

Abstract

Deliberation is commonly thought to increase people's awareness of both supporting and opposing arguments. But who learns, and who doesn't learn, them? We examine these questions with data from a Deliberative Poll (DP) in Northern Ireland on the future of the school system, whose questionnaire included open-ended questions about the reasons that people supporting and opposing various policies might offer. Data show that people in even deeply divided societies can successfully deliberate about topics that broach sectional lines, and come to learn opposing arguments.

In their everyday lives, many people have little idea of how others holding policy preferences very different from theirs could do so (Mutz, 2002b). The least attentive may not even recognize the existence of opposing opinions. Others may recognize their existence but never trouble to wonder how they could be held. Yet others may recognize their existence but be at a loss to explain them. And yet others may recognize them and attribute them to ignorance, stupidity, prejudice, or naked self-interest.

For that matter, some people—again, the least attentive—may have no conscious reason for their *own* policy preference. Some indeed have no real preference, and report "non-attitudes" in response to survey questions (Converse 1970). Others may have a preference and one or more conscious reasons for it. They may even recognize (without accepting) a few more. But many recognize only a few of the plausible reasons for their own views (and of course hold still fewer).

Deliberation may to some degree remedy all this. Deliberation, as the term is most commonly employed, is discussion—but not just any discussion. It is balanced, exposing its participants to competing arguments—but not just balanced. It is also informative, involving an exchange of information, broadly construed, and earnest, in the participants' considering and weighing each other's views, leaving open the possibility of changing their own. And it is civil.

The process should inject its participants' everyday political reality with a good dose of additional cognition. It should increase their recognition and internalization of the possible reasons for their own existing preferences, enriching and possibly nuancing them, if not necessarily altering their broad thrust. That in itself can be normatively valuable insofar as it moves them closer, at least as a matter of degree or nuance, toward their authentic (full-information) views.

More distinctively, however, deliberation should increase the participants' awareness of the possible reasons for *opposing* preferences. That may be still more valuable, enriching and possibly nuancing their existing preferences—and sometimes altering their broad thrust, when a change would make them more authentic. That is one of two possible benefits. The other is that participants who learn more of the other side's reasons may regard them with greater respect, understanding better how they could legitimately—if mistakenly—disagree.

But who learns how many supporting and opposing arguments? How is that learning conditioned by other factors, including the knowledge with which they enter the process? And to what extent and how does such learning of reasons affect their policy preferences and attitudes toward those holding opposing views? We examine these questions using data from a Deliberative Poll (DP) on the future of the local school system in the Omagh District Council area of Northern Ireland.

Arguments

Let us begin by underscoring some conceptual distinctions. First, we are primarily concerned here with the learning of reasons, not the persuasion or attitude change that may sometimes ensue. To cognize arguments is not necessarily to accept them, much less accept enough of them to change one's views. The participants in a deliberation may reject many, most, or all of the arguments they learn. But they do typically learn some, often many, often on the other side of the issue, or so is our claim. What we are examining here is the movement from "How could you think that?" to "I can see how you could think that" (and from "why do I think that?" to "that's why I think that"). In most cases "I can see how you could think that" would in fuller form be followed by "although I still disagree."

It is important to keep in mind that the "sides" here are sides of a policy debate, not sectarian groups. The question is the learning of arguments, especially those favoring the other side of the debate. This isn't to say that identification with sectarian group cannot impact learning. Where relations between sectarian groups are hostile, inter-group animus may lead some people to ignore arguments made by the opposing group, regardless of its content. Furthermore, if the issue being discussed touches upon sectarian cleavages, people may be especially unlikely to learn arguments that have to do with sharing resources across communities and working together.

Beyond that, some explicit terminology and notation may make *our* argument clearer. For simplicity's sake, assume a binary choice between policy options A and B, although B could be not-A rather than any specific alternative. An argument may be either pro- or anti- A (treating pro-B as anti-A and anti-B as pro-A). Denote the numbers of pro- and anti-A arguments in long-term memory by n_p and n_a, making the combined number of stored arguments $n_c = n_p + n_a$, and the *imbalance* of these arguments—the extent to which they are pro- versus anti-A—be $b = (n_p - n_a)/(n_p + n_a)$. If $n_a = 0$ (they are entirely pro-), $b = 1$; if $n_p = 0$ (they are entirely anti-), $b = -1$; and if $n_p = n_a$ (they are equally pro- and anti-), $b = 0$. The *imbalancedness*—the degree of imbalance in either direction—may then defined as abs(b). These quantities are all implicitly triple-subscripted: for the individual, the topic, and the measurement-wave.

An argument's being "supporting" or "opposing" is a matter of the agreement or disagreement of its thrust (pro- or anti-A) and the individual's own position (pro- or anti-A). In many contexts, it may make most sense to define individual's own position as his or her *current* position, but in the present context, where we are examining change over a period, rather than

instantaneously, it makes more sense to take it as his or her *initial* position. Here *supporting* arguments are those tending to buttress his or her initial position (pro-A if he or she started as pro-A, and anti-A if he or she started as anti-A), while *opposing* ones are those tending to undermine his or her initial position (anti-A if he or she started as pro-A, and pro-A if he or she started as anti-A).

Parallel variables can characterize the arguments aired in the deliberation. Some arguments, imported from outside sources or freshly excogitated by participants, may not reside in the head of any participant at the outset. Most, at the outset, are presumably cognized by some but by no means all. What happens thereafter is our story. The point is that these are outside-the head entities, residing in communications, not necessarily represented in belief systems. Here we are concerned with the numbers of distinct arguments *offered*, N_p, N_a, and N_c; their imbalance $B = (N_p - N_a)/(N_p + N_a)$, and their imbalancedness, abs(B).

Deliberation and the Learning of Arguments

At any given moment, the number of pro-A arguments a person knows (n_p) is a function of his or her historical exposure to pro-A arguments (N_p) and historical incentives (call them I_p) to absorb them (viewing incentives to ignore as the other side of the same variable as incentives to absorb). Illustratively (there is no reason to commit ourselves to a specific functional form):

$$n_p = f(\alpha_0 + \alpha_1 N_p + \alpha_2 I_p),$$

where *f* is some unspecified non-negative function. Similarly,

$$n_a = f(\alpha_0 + \alpha_1 N_a + \alpha_2 I_a),$$

since there is no reason to expect the parameters to be different for anti- versus pro-A arguments.[1]

H1: Deliberation increases the total number of cognized arguments ($n_c = n_p + n_a$)

Deliberation should promote the learning of "new" arguments (previously encountered, perhaps, but not previously cognized), both pro and con. It adds both exposure, ΔN_p and ΔN_a (both definitionally > 0) and incentives, ΔI_p and ΔI_a (both expectedly > 0). Indeed, thinking about the arguments is part of the task, a helpful "demand characteristic." Thus deliberation should increase both n_p and n_a (just as it does factual knowledge, as demonstrated in numerous Deliberative Polls (Luskin et al. 2009; Hansen 2004)).

H2: In everyday life, cognized supporting arguments (consistent with existing attitudes) should outnumber cognized opposing ones ($n_p > n_a$, assuming that "pro" is supporting)

Denote pro-A arguments by individuals who are pro-A and anti-A arguments by individuals who are anti-A as "supporting," and anti-A arguments by pro-A and pro-A arguments by those anti-A as "opposing." In their everyday lives, people know many more supporting (pro-A) than opposing (anti-A) arguments (Mutz 2002b, Baron 2007), for probably much the same reasons they tend to know more attitudinally congenial than uncongenial facts (Delli Carpini and Keeter 1997, Bartels 2002, Shapiro and Bloch-Elkon 2008, Jerit and Barabas 2011; cf. Gaines et al. 2007):   they are exposed mainly to pro-A arguments and face greater incentives to absorb them.

Part of this story is that people increasingly live among others sharing their political orientations (Calhoun, 1988; Bishop and Cushing, 2009; cf. Abrams and Fiorina, 2012).[2] They also voluntarily self-select, choosing, even among their neighbors, to associate mainly with those

sharing their political orientations (Verbrugge, 1977; 1983; Knoke, 1990; Huckfedlt and Sprague, 1995; Mutz and Martin, 2001; see also MacKuen, 1990). And when they nevertheless interact others holding opposing views, an interest in avoiding the possibility of discord, may keep any areas of disagreement off the table.

People generally seek information and opinions confirming their existing views (Siune, 1984; Jecker, 1964). This tendency naturally applies to consumption of policy relevant information (Garrett, 2009; Iyengar and Hahn, 2009; Stroud, 2010). And recent technological advances seem only to facilitate self-selection into attitudinally congenial content (Sunstein, 2006). People process information less as neutral observers and more as biased partisans (Chang 2003; Lord, Ross and Lepper 1979 etc.).

H3: Deliberation Should Narrow This Everyday Imbalance between Cognized Supporting and Opposing Arguments.

Deliberation should narrow this gap. It airs arguments on both sides, more or less equally. At least on average, in the long run, $\Delta N_p = \Delta N_a$. In any given deliberation, that equality may be only "more or less," but so long as the newly encountered arguments are less imbalanced than the previously encountered ones ($\Delta N_p/\Delta N_a < N_p/N_a$), the addition of the former should decrease the overall imbalance (moving B toward 0).[3] Deliberation also increases the incentives to take opposing arguments aboard. Part of the task is to confront them and to interact with those making them. They acquire "faces." It is more difficult to ignore arguments made in discussions with people one (now) knows. Attention and retention may remain somewhat selective, but so long as the new incentives are more balanced than the old ($\Delta I_p/\Delta I_a < I_p/I_a$), the imbalance should be lower, post-deliberation.

Conditioning Factors

Other variables may condition these effects. In particular: (1) Anything that leads to a scanter store of cognized arguments may also depress the number of new arguments learned. (2) Anything that accentuates the initial imbalance between the numbers of supporting and opposing arguments already known has the potential to increase the extent to which it is leveled by deliberation. (3) Anything that decreases the difference in the incentives to absorb supporting versus opposing arguments in deliberation should decrease this leveling effect.

H4: Policy attachments may dampen the learning of opposing arguments and hence prevent the narrowing of the gap between supporting and opposing arguments

Even where $\Delta I_p/\Delta I_a < I_p/I_a$, and $\Delta N_p = \Delta N_a$, $n_p - n_a$ may still increase. In fact, it will do so for all cases where $\Delta I_p/\Delta I_a > 1$. People who are pro-A may be less likely to learn anti-A arguments, while those anti-A may be less likely to learn pro-A arguments. For, as we note above, while deliberation 'forcibly' exposes people to all sides of the argument, selective attention and retention can still influence how many supporting and opposing arguments people learn. We test whether deliberation leads to symmetric gain in arguments, where $n_p - n_a$ is greater among those who deliberate.

H5: Intergroup hostility may dampen both the learning of arguments and the narrowing of the gap between supporting and opposing arguments.

Suppose the population can be divided on some dimension into two groups, G_1 and G_2, like Catholics and Protestants in Northern Ireland. The extent to which member of one of the groups dislikes the other group may limit the attention she gives to arguments made by those belonging to the other group. Thus, intergroup hostility could greatly lessen the argument pool

that an individual is willing to consider as many candidate arguments could be ruled out by their source and hence dampen learning.

If the group membership is noticeably correlated with position on the issue—with, for example, members of G_1 being distinctly more pro-A than members of G_2—members of G_2 may tend to ignore the new pro-A arguments, associating them with the disliked G_1, and members of G_1 to ignore the new anti-A arguments, associating them with the disliked G_2. Note, however, that this sort of tuning-out requires both that groups dislike each other, and that policy positions be correlated.

A third possibility only presents itself when the issue expressly touches upon group rights, resources, and behaviors. For instance, it is very likely that some of the inter-group affect spills over on to attitudes on policies to do with increasing contact, or sharing resources, with another community. I n Northern Ireland, religion is a powerful basis for group identity, with sharp long-standing divisions between Protestants and Catholics. The Protestant-Catholic divide is powerfully supported by residential segregation. For example, "92.5% of public housing in Northern Ireland is divided along religious lines. In Belfast it is even worse – 98%" (O'Hara, 2004; see also Cohen, 2007).[4] Many Catholic and Protestant communities, especially in poorer, working-class areas, are separated by "peace walls" which have (figuratively as well as literally) cemented divisions rather than ameliorated them (see Pogatchnik, 2008).[5] And, Protestants and Catholics tend to marry within their own community, send their children to segregated schools, and, traditionally at least, have tended to work in different industries (Boyle and Hadden, 1994). In sum, a great deal of antipathy pervades Protestant-Catholic relations in Northern Ireland. So in Northern Ireland, those holding negative affect against the other community may be particularly

disinclined to learn about arguments in favor of policies that require children of both Protestants and Catholics to be taught together in the same classroom.

H6: Deliberation may produce a bigger increase in the number of cognized arguments, the more so for those who already have a lot of them.

The overall number of arguments learned ($\Delta n_c = \Delta n_p + \Delta n_a$) may depend on the number of arguments already known ($n_c = n_p + n_a$). Participants who arrive knowing more arguments may tend to learn still more, just as those who arrive with more factual knowledge (largely the same people) tend to acquire still more of it. The rich, as just about always in rich learning environments, get richer (see the discussion and citations in Luskin, Helfer, and Sood 2013).

H7: Deliberation's effect on the overall learning of arguments, and especially of opposing arguments, should be greater, the more attitudinally diverse the deliberators.

There is no reason to expect sociodemographic diversity to have any conditioning effect, except insofar as some of the traits involved are correlated with policy attitudes. But encountering people with opposing policy attitudes (the probability of which is maximized, for the group as a whole, when it is split 50-50 between the sides of the debate) should increase the probability of learning them (higher Δn_a and thus higher Δn_c, even if Δn_p remains the same).

Deliberative Polling and the Omagh DP

We test these hypotheses with data from a Deliberative Poll in Omagh, a district council area of nearly 48,000 inhabitants in County Tyrone, Northern Ireland, on the future of the local school system (Luskin et al., 2012). Omagh has a mixed population of Catholics and Protestants and primary and post-primary schools representing all the major school types (State Controlled, Catholic Maintained, Other Maintained, Voluntary, Integrated and Institutions of Further and

Higher Education). Some have mainly Catholic pupils, some mainly Protestant pupils, and a few

a more even mix. As in Northern Ireland as a whole, the birthrate and thus the school-age

population are declining, which is forcing schools to consider greater coordination or

consolidation (Department of Education 2007).

DPs embody a specific form of deliberation. The Omagh DP adhered to the standard DP

design (see, e.g., Luskin, Fishkin, and Jowell 2002; Fishkin and Luskin 2005; Fishkin 2009). A

random sample of parents of school-aged children was polled (T1) and invited to participate in a

long day of deliberation. Of the 568 initial interviewees, 124 took part in the DP (we discuss this

recruitment rate later). They were sent balanced briefing materials, then during the day engaged

in face-to-face deliberations in small groups, alternating with plenary sessions, during which they

posed questions to panels of policy experts and policy makers, including representatives from all

of the different school types. They were re-interviewed at the end (T2). A separate control group

($n = 150$) was also interviewed, along with a subsample ($n = 93$) of the participants a month later

(T3).

Like other DPs, the Omagh DP thus brought a diverse set of people together to interact in

an environment conducive to a respectful airing of views, involving both similarities and

differences of both opinions and aims. In expectation (in the statistical sense), the participants

were equally exposed to new facts and arguments and experienced an equal increase in their

incentives to learn them. We return to the individual-level variation around these equal

expectations presently.

The setting of this particular DP—a classic "deeply divided society," with a still-recent

history of violent strife—is in many ways distinctive. From the standpoint of those skeptical of

the mass public's ability to deliberate across deep divisions, it was an achievement simply to bring these randomly selected members of the mass public together to talk. As in most other DPs, there was some considerable learning of relevant facts and some significant net change of policy preferences. There also seemed to be some increased in each group's respect for the other (for all this, see Luskin et al. 2013).

Yet the deep divisions between Catholics and Protestants do not necessarily pose a special obstacle for learning the arguments on the other side of these school policy issues. As it turns out, Catholics and Protestants do not hold very different positions on these issues (see Luskin et al. 2013). So anti-A Catholics (where A is policy involving mixing) can expect to hear about as many pro- (and anti-) A arguments from Protestants as from their fellow Catholics.

Granted, enmity toward the other group may be among the factors leading some participants to oppose policies calling for Protestant and Catholic children to be taught together in the same classroom. Yet there is little reason to think that such enmity conditions the learning of new arguments. Again let A be a policy involving religious mixing. Pre-deliberation, Catholics/ Protestants hostile to Protestants/Catholics may have learned many more anti- than pro-A arguments. For them, n_a may be much greater than n_p—as it may also be for other people, for other reasons. During the deliberation, these anti-A participants may be more inclined than other participants to ignore new pro-A arguments, making Δn_p smaller, but may also be more inclined to absorb new pro-A arguments, making Δn_p larger. The net impact, on Δn_c, is unclear.

But what of the effect on the imbalance of the arguments the participants know? The ratio $\Delta n_p / \Delta n_a$ will likely be smaller for anti-A than for pro-A participants. It may even be less than 1, though almost certainly still greater than n_p / n_a, meaning that deliberation is still reducing

the imbalance. Like the ratio $\Delta n_p / \Delta n_a$, this reduction will presumably be smaller for anti-A than for pro-A participants. But that is not the question. The question is whether it is still smaller for the more anti-Protestant Catholics or the more anti-catholic Protestants who are anti-A than it is for other anti-A participants. That seems unlikely, in a process in which anti-Catholic/-Protestant Protestants/Catholics are hearing many supporting arguments from members of the other community and many opposing arguments from members of their own.

Cognized Reasons

The T2 and T3 surveys contained pairs of open-ended questions soliciting the reasons for favoring or opposing four policy proposals relating to how the school system might be reorganized. These followed closed-ended questions about the respondent's preferences on these same policies:

All-Ability Schools: Respondents were asked to rate on a scale of 0 (opposing something as strongly as possible) to 10 (supporting as strongly as possible) the option of "having a system of all-ability schools, all providing the same wide curriculum". We rescaled the responses so that they lay between 0 and 1.

Consolidating schools: On a scale of 1 to 10, where 0 is opposing something as strongly as possible, 10 is supporting it as strongly as possible and 5 is exactly in-between, how strongly would you support or oppose the option of Omagh schools combining primary and post-primary pupils (for example, ages 7-14). Scores were rescaled to lie between 0 and 1.

School collaboration: How strongly would you agree or disagree with the statement that if schools of different religious composition enter partnerships, the children from both schools should at least sometimes be taught in the same classroom – strongly disagree, tend to disagree,

neither agree nor disagree, tend to agree, strongly agree. Responses were coded 0, .25, .50, .75, and 1 respectively.

Balanced enrolment: Respondents were asked to place their views on a 0 to 10 scale on the importance of religious homogeneity of children's schools. Scale point 0 meant it was important children should attend school only with other children of their own religion and 10 meant that it was important for children to attend schools that have a balanced enrolment of Protestant and Catholic pupils.) Scores were rescaled to lie between 0 and 1.

The open-ended follow-ups ran:

All-ability schools: Respondents were asked to list up to five reasons that people who "strongly support" having "a system of all-ability schools, all providing the same wide curriculum" would provide. They were also asked to list up to five reasons that people who "strongly oppose" the proposal would provide.

Consolidating schools: Respondents were asked to list up to five reasons that people "who strongly support schools combining primary and post-primary pupils" would "give for their position". Following this question, they were asked to list up to five reasons that people who "strongly oppose" the proposal would "give for their position".

School collaboration: Respondents were asked to list up to five reasons that people "who strongly agree" that "if schools with different religious compositions enter partnerships, the children from both schools should at least sometimes be taught in the same classroom" would "give for their position". Again, they were asked to list up to five reasons that people who strongly disagree with the proposal would "give for their position".

Balanced enrolment: Respondents were asked to list up to five reasons that people "who

think that children should attend schools that have a balanced enrolment of Protestant and Catholic pupils" would "give for their position". As before, they were also asked to list up to five reasons that people who "think that children should attend schools only with other children of their own religion" would "give for their position".

We coded the responses to the open-ended questions based on the substance of the argument (See Appendix A. for details about how responses were coded).

Other Measures

Inter-community Affect. Respondents were asked to rate how favorably or unfavorably did they feel towards Catholics and Protestants on a scale of 0 to 10, where 0 meant "as unfavourably as possible", 10 "as favourable as possible", and 5 "exactly in-between". Responses from the questions were used to create a difference score, difference between in-group ratings and out-group ratings, tallying inter-community affect. The variable range from 0 (a 10 for out-group, and 0 for in group) to 1 (a 10 for in-group and 0 for out group).

*Trait Rating*s. Respondents were asked to rate to what extent they thought 'most Protestants' and 'most Catholics' were trustworthy and open to reason on a 0-10 semantic scale ("untrustworthy"/"not open to reason" to "trustworthy"/"open to reason"). We averaged the ingroup – outgroup differences (for Catholics, Catholics – Protestants, for Protestants the reverse) to create a variable ranging from 0 (the same score for both groups on both traits) and 1 (10 for in-group and 0 for out-group on both items).

Factual Knowledge. Respondents were asked seven closed-ended knowledge items relating to the school system (listed in Appendix B). Correct responses were scored 1, all other responses 0. The scores were linearly averaged to produce an index.

Group Knowledge: Mean factual knowledge of participants in the group excluding the individual at T1.[6]

Attitudinal Heterogeneity: Square root of average variance across the four attitude items minus the covariance (Anderson 2003; see also Luskin et al. 2009).

Age, recorded in years.

Gender, coded 1 for female 1, 0 for male.

Religious affiliation, coded 1 for Catholic dummy, 0 for Protestant.

Education, coded into four ordinal categories: in descending order, bachelors or higher, BTEC (higher)/GCE "a" level etc., GCSE/CSE etc., no qualifications.

Research Design

The data allow us to make three kinds of comparisons: (1) T2 participants (*n* = 124) versus the Control Group (*n* = 150), (2) T2 participants who also completed the T3 survey (*n* = 93) versus the Control Group, and (3) T3 participants (*n* = 93) versus the Control Group. Comparison (1) estimates the treatment's short term effect. Comparison within participants between T2 and T3 allow us to answer to what degree the treatment effects last. Comparison (3) allows us to see whether participants still know more arguments than those in the control group.

These comparisons are complicated by sampling variation – gods of random assignment may make the treatment and control groups 'unbalanced' on some covariates, self-selection (from the initial random sample into the treatment group), systematic attrition (from T2 to T3), and treatment spillover (the accidental treatment of part of control group). Comparisons are also complicated by the fact that a DP actually involves not one uniform treatment but a collection of parallel treatments. The participants are (randomly) assigned to the small groups, which vary in

ways that may affect learning of arguments (and other consequences of deliberation). Groups in which one's fellow participants know more arguments may expose their members to more arguments (see Luskin et al. 2012 for a similar hypothesis about the learning of factual information). A more attitudinally heterogeneous group may foster richer discussion, in which the arguments aired are both more numerous and more diverse. It may also create more new incentives to learn opposing arguments.

Before we move to presentation of results, we discuss some potential problems that can vitiate comparisons between the CG, and the treatment samples (T2 and T3). Estimates of treatment effect gotten through raw comparisons between the CG and treatment groups means can be biased due to confounding differences between the treatment and the control group, and treatment spillover – accidental "treatment" of portion of the control group. We analyze below to what degree these factors are at play, starting with discussion of differences between the CG and T2 and T3 samples.

Differences between Control Group and T2 and T3: Composition of the CG and the treatment samples may differ as a result of sampling variation, self-selection into treatment, and non-random attrition between T2 and T3. When it comes to differences between the CG and the treatment groups, the most severe concerns relate to 'confounding differences' – differences in composition of the groups that can at least partly explain the differences across dependent variables of interest. For instance, if more knowledgeable are also more likely to select into the treatment group, differences between the CG and the treatment groups in the average levels of knowledge will provide positively biased estimates of the treatment effect. With that in mind, let us investigate self-selection into treatment, and attrition between T2 and T3 –

Self-Selection into Treatment: Of the people who took the initial questionnaire and were then invited to participate in the DP event, a majority chose not to come. To test whether there were systematic patterns to who came and who didn't, we estimated a model predicting decision to attend DP using theoretically relevant variables – knowledge, education, gender, and religion. In line with prior findings, we expect the more knowledgeable and more educated to be more likely to attend (Westwood and Sood, 2010). Our predictions for religion are based on the contextual information. Since the Protestant population is falling vis-à-vis Catholics, Protestants have more reason to turn up and fight their corner.

However, results show there aren't many significant differences in the makeup of the participants and the non-participants on the variables we entered. Compared to the non-participants, participants were no more knowledgeable, and held similar mean levels of education. The proportion of females did not differ significantly across participants and non-participants. However, compared to Protestants, and as expected, Catholics were significantly less likely to attend the DP (see Table 3).

Attrition between T2 and T3: Of the 124 DP participants, only 93 filled out the T3 survey. However, the loss of approximately 25% of the sample between T2 and T3 doesn't appear to be systematic. Notably, factual knowledge, education, number of arguments offered at T2, and religious status did not predict whether the participant filled out the survey at T3 or not (see Table 3).

Aside from self-selection, and systematic attrition, sampling variation can cause CG and treatment samples to differ. So we compared the CG and the T2 and T3 samples on the key confounding variables. Both multivariate specifications (see Table 3) and bivariate comparisons

of crucial confounding variables revealed no significant differences across the CG and the participant samples. We acknowledge that absence of significant coefficients doesn't mean that no differences exist – small sample sizes mean some medium sized differences may not reach statistical significance – but it does suggest that there were no large asymmetries that would make comparisons greatly problematic. However, to quell any remaining doubts, we also present results from a model that controls for crucial potentially confounding background variables.

Treatment Spillover: 11.3% of the control sample indicated that they had heard about the Deliberative Poll on education in Omagh. Of those who had heard about the DP, 94.1% said that they had "read or heard about issues of education in Northern Ireland in newspapers or magazines or on television or radio" over the past month, as opposed to 69.2% of those who hadn't heard about the Deliberative Poll. Similarly, 94% of those who had heard about the DP said that they "discussed issues of education in Northern Ireland with family, friends, or co-workers", as opposed to 73.7% of those who hadn't, and 47.1% of those who had heard about the DP said that they had "sought information about issues of education in Northern Ireland in the library or on the internet", opposed to 27.1% of those who hadn't.

At least part of the difference is explained by who heard about the DP – the more knowledgeable and educated were more likely to hear about the DP.[7] However, the treatment (here – hearing about the DP) made people still more likely to read about, and discuss these issues.[8] If we ignore the spillover, our estimates will be biased, albeit the bias will run in a conservative direction – increased attention to media etc. inarguably has a non-negative impact on the number of arguments control group participants report. We opt for this more conservative strategy.

Results

The awareness of arguments in the control group provides a window on the everyday world. We follow it with an analysis of the short-term impact of the treatment on knowledge of arguments, followed by analysis of somewhat longer-term impact of the treatment. We then test the robustness of our estimates by analyzing the data controlling for potential confounders. Next we analyze whether learning was selective. Before exploring consequences of learning of arguments, we analyze whether learning of arguments differed across small groups.

Knowledge of Arguments in the CG

On average CG respondents provide less than one argument in support or opposition of each of the four policies (see Table 1). For instance, on the issue of 'balanced enrolment', the CG respondents gave on average 1.27 arguments in support of the policy, and an average of 1.25 arguments against the policy. If we don't credit people for reasons that don't give too much credit to the other side – people who oppose this policy are prejudiced – the means look somewhat weaker.[9] Removing the tiny fraction of the reasons that bear little relation to the policy issue at hand leaves the results mostly unchanged.

There doesn't appear to be an imbalance between knowledge of attitudinally congenial and uncongenial arguments. For example, people who opposed 'all ability schools' knew essentially similar number of arguments in support as against it (see Table 4). Essentially the same null pattern held for 'balanced enrolment' and 'school collaboration'.

Learning of Arguments

At T2, the participants knew significantly more—more than 35% more—arguments than the CG (see Appendix C). In particular, there were sizable and significant differences on all-

ability schools and balanced enrolment. Participants also gratifyingly knew significantly more

arguments in support of inclusive classrooms than respondents in CG. Differences across other

remaining items did not breach significance though were notably positive and healthy for reasons

in support for consolidating schools. Essentially the same result is obtained when one compares

the CG with the subset of the T2 sample that also responded at T3 (see Table 2).

(Table 2 about here)

Once again, participants knew overall far more arguments than respondents in the CG. And

iterating over each policy, we again see that participants surveyed right after deliberation knew

significantly more arguments both for and against all-ability schools, balanced enrolment, and in

support of an inclusive classroom.

Next, we compare responses from the survey of participants conducted

contemporaneously with the CG survey. The results are an attenuated version of the results we

report above. Participants, when surveyed a month later, listed overall more arguments than the

CG respondents; the difference was significant for a one-tailed test (see Table 1). On a less

positive note, there was a sharp and significant decline in the number of arguments people were

able to list between the time right after deliberation and the survey conducted a month later. This

decline is plausibly explained by either forgetting (which means that some of the gains from

deliberation were somewhat fleeting), or lower motivation to respond effort-fully a month later.

Lastly, for testing robustness of the estimates of treatment effect that we present above, and for

formally testing some moderators, we estimate the treatment effect controlling for some potential

confounders – factual knowledge, education, and gender, and by including some theoretically

motivated interactions. For instance, we add religious affiliation to our specification and formally

test whether there were differences in how much Catholics learned vis-à-vis Protestants. We also account for one additional potential complication: A closer look at Table 2, which has items arranged in order they appear in the questionnaire, reveals a slightly odd pattern – people appear to offer greater number of reasons on questions posed earlier in the questionnaire, than on questions asked later on in the questionnaire. Since we had no a priori (prior to the design of the questionnaire) reasons for suspecting that respondents knew fewer arguments on the questions posed later, we suspect the pattern, if it exists, is a consequence of respondents, perhaps tired of responding to one too many open-ended questions, satisficing on latter questions. A significant order effect would not change our results except lend greater understanding to why we see somewhat fewer arguments being noted on proposals for 'inclusive classrooms' and 'consolidating schools'. So we control for the order in which the question was posed in the questionnaire, and formally check whether the effect of order was different across treatment and control group.

(Table 3 about here)

Results indicate that the effect of the treatment is still positive and highly significant (T2: b = .79, p < .001; T3: b=.62, p < .01, see Table 3). In all, we can safely conclude that DP participants did learn arguments over variety of different policies over the course of the treatment. When comparing T2 participants to CG, we find that question order effects were sharper in treatment group than in CG.[10]

Selective Learning

 As noted before, given lower birth rates among Protestants vis-à-vis Catholics, some Protestants may believe that they have more to lose in supporting policies such as 'balanced

enrolment' and 'inclusive classroom'. Given the contextual reality, one may plausibly conjecture

that Protestants would be less likely to learn arguments in support of 'balanced enrolment' and

'inclusive classroom', or more deviously, more likely to learn arguments against the two

policies. Results indicate that these concerns are ill-founded. Splitting the sample by

respondents' religion, it becomes amply clear that Protestants learn if anything more arguments

than Catholics in support of both 'balanced enrolment' and 'inclusive classroom' (see Appendix

C). On the proposal for balanced enrolment, when we compare the entire T2 sample with the CG

sample, it is Protestants who appear to have learned more arguments in support than Catholics

(Diff. $_{Protestants}$ = .68, p < .05, Diff. $_{Catholics}$ = .33, p < .1). In doing so, they all but bridge the small

gap in knowledge of arguments in support that separates the two communities in the CG sample.

Gratifyingly, Catholics and Protestants learn arguments against balanced enrolment at

about an equal clip (Diff. $_{Protestants}$ = .59, p < .1, Diff. $_{Catholics}$ = .63, p < .01). [11] In the case of

learning of arguments in support of inclusive classroom, again it appears Protestants come out

slightly ahead (Diff. $_{Protestants}$ = .53, p < .1, Diff. $_{Catholics}$ = .3, n.s.). In all, results indicate that

concerns about selective learning based on religious affiliation are ill-founded. A formal test of

whether learning of arguments differed across all policy issues, taken together, also yields null

results.

Continuing with the theme of selective learning, we next explore whether initial knowledge or

education were potential moderators of learning of arguments. When comparing T2 participant

sample with the CG, we find, expectedly, that learning of arguments was healthier among those

with higher education (see Table 3). Unexpectedly, initial levels of factual knowledge had no

impact on rate of learning of arguments. When we compare T3 participant sample to the CG, we

find that neither education nor extent of factual knowledge moderated the treatment effect (see Table 3).

Next, we explored to what extent learning of attitudinally congenial arguments outstripped that of uncongenial arguments, with congeniality tallied by initial policy position. To do so we split the participant sample based on whether the person supported a particular policy in the pre-deliberation survey, and split the CG sample by policy attitudes measured contemporaneously. Results indicate that both, those who supported the policies initially and those who opposed the policies ended up learning more arguments in support of policies (see Table 4). One explanation for the result is asymmetry in the initial participant sample – given a substantial majority of participants started with attitudes that were in support of the policies, it is possible that fewer negative arguments were voiced during deliberations.[12]

(Table 4 about here)

Given DP participants learned more reasons in support of the policies, than in opposition, there are natural worries about whether the treatment was tantamount to one-sided persuasion. For instance, some may worry that this asymmetric learning led to changes in opinion in the direction of supporting the policy. However no such changes were observed. Respondents who opposed respective policies in post-deliberation survey also seem to know more arguments in support.[13]

Until now we have thought of the DP as one single treatment. However, not everyone receives the same treatment in a DP. As we noted before, within a DP, participants are randomly assigned to small groups. Consequently the small groups vary randomly in their composition. We suspect random variation in the number of arguments known by other members of a given group. Attitudinal heterogeneity, among other things, affects how much people learn. To test

these hypotheses, we specified a hierarchical linear model predicting number of arguments offered on any item using mean knowledge of other members in a group (a proxy for number of arguments known), attitudinal heterogeneity, individual background factors thought to predict initial knowledge of arguments, and learning of arguments (education and initial factual knowledge), and to what extent individuals read the briefing materials; we allowed for separate intercepts for each small group, individual, and item.

(Table 5 about here)

In line with our expectations, similar participants in more attitudinally heterogeneous gave more arguments in the post-deliberation survey than those assigned to less diverse groups (b =1.83, p < .05). Again, as predicted, similar participants assigned to groups where other participants held more factual knowledge, gave more arguments in the post-deliberation survey, than those assigned to groups where participants knew less (b =, p <). As for individual level factors, both initial knowledge and education were expectedly positively correlated to number of arguments given by the respondent in the post-deliberation survey (see Table 5). These results have some bearing on how we conceive of the effect of the DP. For discussion and analyses of what effects of group composition mean for how we conceive of the effect of the DP, see Appendix D.

Consequences of Learning

We tested whether the impact of the treatment on attitude change was mediated by learning of arguments (Imai, Keele, and Tingley, 2010).[14] Of the four policy issues, only one showed significant change; we limit our attention to it. We can specify two kinds of models – one where attitude change is mediated by difference in arguments learned in favor and in opposition, and another where attitude change is mediated by overall number of arguments

learned. We test both but expect only the latter to be successful.[15] We find no evidence of

learning of arguments mediating effect of treatment on attitude (see Table 6).

(Table 6 about here)

Next we check whether learning of arguments mediated changes in beliefs about the out

religious groups. Particularly, learning about arguments in favor of school mixing and integration

should be transformative.

Discussion

Participating in a Deliberative Poll, does indeed induce learning of arguments, even in the

context of a deeply divided society and when the topic of discussion is contentious. Among other

things, the deliberators learn reasons about why other people oppose and support policies.

As far as we know, this is the first empirical demonstration of learning of opposing

arguments. Past research hitherto has shown that those who choose to expose themselves to

"cross-cutting" exposure are more informed about arguments on the other side (Mutz 2002). But

it is possible, nay likely, that those who choose to expose themselves to "cross-cutting" exposure

are very different from those who don't in many unmeasured ways. We think it possible for

everyday discussion to upend some of the problems of imbalance, on the rare occasions it is

sufficiently engaging, sustained and balanced. We just await empirical confirmation. We also

contend, though again leave it for future research to investigate, that *deliberation's* effect is

likely stronger than that of everyday discussion.

A variety of issues limit generalizability of the results. This is one Deliberative Poll, on

one topic. Northern Ireland is a deeply divided society, which affects a great many things about

its mass (and elite) politics, including much of what can be expected from deliberation. As we

explain above, our fears are mostly that the estimated effect size is if anything a lower bound of what can be expected in similar such experiments conducted elsewhere with a shallower history of rancor.

One reasonable way to conceive of 'true preferences' is to think of them as preferences a person would hold if she or he were well informed, had heard competing opinions, and understood the consequences of alternative policies (Fishkin, 1991; 2009; see Dahl, 1989, p. 180-81, 307-8; Arendt, 1968).[16] Most people, however, know and have thought very little about most policy issues. Deliberative Polling, by providing people not only factual information but also the opportunity to hear competing reasons, affords a glimpse of what a more informed citizenry that has also had the opportunity to consider competing arguments would look like. Our results on consequences of learning of argument on opinion change are negative but we believe a stronger research design may well show a different result on this particular point.

A research design including open-ended items canvassing for arguments measures, both pre- and post-treatment would be helpful. It might also be helpful to tweak the design of these questions, for instance providing inducements when respondents seem to lose steam in answering them. The recording of the arguments offered in response could also be made easier, for instance by using automated anonymous telephone systems allowing respondents record arguments more easily than on a sheet of paper (as in this study). All that is for the future. For now, we conclude with the important finding that even in deeply divided societies, people can come together and deliberate about issues that touch upon the cleavages, and learn opposing arguments.

References

Abrams, Samuel J., and Morris P. Fiorina. 2012. "The Big Sort" That Wasn't: A Skeptical Reexamination. *PS: Political Science & Politics*, 45 (02), pp. 203-210

Altemeyer, Robert. 1996. *The Authoritarian Specter.* Cambridge, MA: Harvard University Press.

Anderson, Theodore W. 2003. "An Introduction to Multivariate Statistical Analysis." New York: John Wiley and Sons.

Arendt, Hannah. 1968. Truth and politics. In H. Arendt (Ed.) *Between Past and Future: Eight Exercises in Political Thought*. pp. 227-265. New York: Viking Press.

Atkin, C. 1973. Instrumental utilities and information seeking. In P. Clarke (Ed.), *New Models for Mass Communication Research* (pp. 205–242). Beverly Hills, CA: Sage.

Baron, Jonathan. 1998. *Judgment Misguided: Intuition and Error in Public Decision Making*. Oxford University Press.

Bartels, Larry M. 2002. Beyond the Running Tally: Partisan Bias in Political Perceptions. *Political Behavior*, 24: 117–150.

Bennett, Stephen E., Flickinger, R. S., and Rhine, S. L. 2000. Political talk over here, over there, over time. *British Journal of Political Science*, 30, 99-119.

Bishop, Bill, and Robert G. Cushing, 2009. *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart.* Mariner Books.

Boyle, Kevin and Tom Hadden. 1994. Northern Ireland: The Choice. Penguin Books, London, UK.

Chaffee, Steven H. 1972. The interpersonal context of mass communication. In F. G. Kline and P. J. Tichenor (Eds.), *Current perspectives in mass communication research* (pp. 95–

120). Beverly Hills, CA: Sage.

Chaffee, Steven H., and McLeod, J. M. 1973. Individual vs. social predictors of information-seeking. *Journalism Quarterly, 50,* 237–245.

Cho, Wendy Tam, James Gimpel, and Iris Hui. 2012. Voter Migration and the Geographic Sorting of the American Electorate. *Annals of the Association of American Geographers.* Forthcoming.

Cohen. Nick. 2007. Stop this drift into educational apartheid. Available from (last accessed: April 24th, 2011).

http://www.guardian.co.uk/commentisfree/2007/may/13/comment.politics

Converse, Phillip E. 1970. Attitudes and non-attitudes: Continuation of a dialogue. In Edward R. Tufte (Ed.), The quantitative analysis of social problems (pp. 168-189). Reading, MA: Addison-Wesley.

Converse, Phillip. 1964. *The Nature of Belief Systems in Mass Publics* in Ideology and Discontent, ed. Apter, D. New York: Free Press.

Dahl, Robert. 1989. *Democracy and Its Critics*, New Haven: Yale University Press.

Delli Carpini, Michael X., and Scott Keeter. 1996. Stability and Change in the U.S. Public's Knowledge of Politics, *Public Opinion Quarterly*, 55 (1991), 583–612

Delli Carpini, Michael X., and Scott Keeter. 1997. *What Americans Know About Politics and Why It Matters.* New Haven, Conn.: Yale University Press.

de Tocqueville, Alexis [1835, 1840] 1969. *Democracy in America*. Translated by George Lawrence, edited by Mayer, J. P. New York: Anchor Books.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper.

Fishkin, James S. 1991. *Democracy and Deliberation*. New Haven, CT: Yale University Press.

———.2009. *When the People Speak: Deliberative Democracy and Public Consultation.* Oxford: Oxford University Press.

Gaines, Brian J., James H. Kuklinski, and Paul J. Quirk. 2007. "Same Facts, Different Interpretations: Partisan Motivation and Opinion on Iraq," *Journal of Politics*, November 15:1-21

Garrett, Robert K. 2009. Politically Motivated Reinforcement Seeking: Reframing the Selective Exposure Debate. *Journal of Communication*, 59: 676–699.

Gilens, Martin. 2001. Political ignorance and collective policy preferences. *American Political Science Review*, 95, 379–396.

Hansen, Kasper M. 2004. Deliberative Democracy and Opinion Formation. Odense: University of Southern Denmark Press

Huckfeldt, Robert, and Sprague, J. 1995. *Citizens*, *Politics and Social Communication: Information and Influence in an Election Campaign*. New York: Cambridge Univ. Press

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. A General Approach to Causal Mediation Analysis. *Psychological Methods*, 15(4), 309-334.

Iyengar, Shanto. 1990. "Shortcuts to political knowledge: selective attention and the accessibility bias," in Ferejohn, J. and Kuklinski, J. (eds.) *Information and Democratic Processes*. Champaign: University of Illinois Press.

Iyengar, Shanto, Kyu Hahn, Jon Krosnick, and John Walker. 2008. "Selective exposure to campaign communication: The role of anticipated agreement and issue public membership." *Journal of Politics*, 70: 186-200.

Iyengar, Shanto, and Kyu S. Hahn. 2009. Red Media, Blue Media: Evidence of Ideological
Polarization in Media Use. *Journal of Communication*, 59:18-39.

Jerit, Jennifer, and Jason Barabas. 2011. Partisan Perceptual Bias and the Information
Environment. Presented at the Southern Political Science Association conference in New
Orleans, LA

Kim, Nuri. 2011. Applying the contact hypothesis to deliberative environments: The case of
Australian Aboriginals. Paper presented at the 2011 ICA annual conference, Boston, MA.

Krippendorff, Klaus. 1980. Content analysis: An introduction to its methodology. Beverly Hills,
CA. Sage.

Krippendorff, Klaus. 1986. Information theory: Structural models for qualitative data. Beverly
Hills, CA. Sage.

Krippendorff, Klaus. 2004. Content analysis: An introduction to its methodology. Beverly Hills,
CA. Sage.

Chang, Chingching. 2003. "Party Bias in Political-Advertising Processing." *Journal of
Advertising,* 32 (Summer): 55–67.

Lord, Charles G., Lee Ross, and Mark R. Lepper. 1979. "Biased Assimilation and Attitude
Polarization: The Effects of Prior Theories on Subsequently Considered Evidence."
Journal of Personality and Social Psychology, 37(11), 2098–2109.

Luskin, Robert C., and James S. Fishkin. 2002. Deliberation and 'Better Citizens'. Joint Sessions
of Workshops of European Consortium for Political Research, Turin, Italy.

Luskin, Robert C., James S. Fishkin, and Roger Jowell. 2002. Considered opinions: Deliberative
polling in the U.K. *British Journal of Political Science*. 32: 671-686.

Luskin, Robert C., Gaurav Sood, James S. Fishkin, and Nuri Kim. 2009. *Deliberation and Learning: Evidence from the Deliberative Polls*. Paper presented at the American Association of Political Science Meeting at Chicago, US, September 2009.

Luskin, Robert C., Ian O'Flynn, James S. Fishkin, and David Russell. Forthcoming. Deliberating across deep divides. *Political Studies*.

MacKuen, Michael. 1990. Speaking of politics: Individual conversational choice, public opinion, and the prospects for deliberative democracy. In John Ferejohn and James H. Kuklinski (Eds.), *Information and democratic process* (pp. 59–99). Urbana: University of Illinois Press.

Marcus, George, J. L. Sullivan, Theiss-Morse, E., and Wood, S. 1995. *With Malice Toward Some: How People Make Civil Liberties Judgments.* New York: Cambridge University Press.

McDonald, Ian. 2011. "Migration and Sorting in the American Electorate: Evidence from the 2006 Cooperative Congressional Election Study." *American Politics Research* 39(3):512–533.

Mutz, Diana C. 2002a. "The consequences of cross-cutting networks for political participation." *American Journal of Political Science*. 46 (4):838-855.

———.2002b. "Cross-cutting networks: Testing democratic theory in practice." *American Political Science Review*. 96 (2):111-126

Mutz, Diana C., and Martin, P. 2001. Facilitating Communication across Lines of Political Difference: The role of mass media. *American Political Science Review, 95*, 97–114.

O'Hara, M. 2004. Self-imposed Apartheid. *Guardian*. Available from (last accessed April 24,

2011): http://www.guardian.co.uk/society/2004/apr/14/northernireland.societyhousing

Pogatchnick, S. 2008. Despite Peace, Belfast Walls are Growing in Size and Number. *USA Today.* Available from (last accessed April 24th, 2011): http://www.usatoday.com/news/topstories/2008-05-03-1826820552_x.htm

Shapiro, Robert Y. and Bloch-Elkon,Y. 2008. Do the Facts Speak for Themselves: Partisan Disagreement as a Challenge to Democratic Competence. *Critical Review* 20:115-39

Sunstein, Cass R. 2002. *Republic.com*. Princeton University Press, Princeton, NJ.

Stroud, Natalie J. 2008. Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior* 30, no. 3, 341-366.

Verbrugge, L.M. 1977. The structure of adult friendship choices. *Sociological Forces,* 56:576-597

———. 1983. A research note on adult friendship contact: a dyadic perspective. *Sociological Forces*, 62:78–83

Table 1: Coefficients from binomial logistic regressions predicting choice to attend DP, respond at T3, propensity to be in CG vs. T2, and T3 participant samples

	T1 P Vs T1 NP	T3 P Vs. T3 NP	T2 P Vs. CG	T3 P Vs. CG
Knowledge[1]	.20	-.09	-.63	-.68
Female	.21	.07	.20	.24
Education	.00	.46	-.39	-.26
Catholic	-.38+	-.09	-.42	-.46
N	528	110	242	213
AIC	534.92	136.92	339.63	290.54

Note: + p < .10; * p < .05; ** p < .01; *** p < .001; [1]Pre-deliberation measure used for estimating self-selection into treatment (T1P Vs. T1 NP), T2 measures used for estimating who chose to fill the T3 survey (replacing it with T1 doesn't alter our results), CG Vs. T2 P and CG Vs. T3 P comparisons use pre-deliberation measures for participants and contemporaneously obtained measures for CG. Intercept not reported.

Table 2: Mean number of reasons given by Deliberative Poll participants at T2 and T3 compared to average number of reasons given by T3 Control Group respondents

Measure	T2	T3	T3 CG	T2 – T3 CG	T3 – T3 CG	T2 – T3 P
All-ability schools						
Support	1.86	1.30	1.05	.81***	.25+	-.56+
Oppose	1.47	.95	.95	.52**	-.01	-.53
Balanced Enrollment						
Support	1.69	1.51	1.27	.42*	.24+	-.18+
Oppose	2.03	1.69	1.35	.69***	.34*	-.34*
Inclusive classroom						
Support	1.41	1.16	1.02	.39*	.14	-.25
Oppose	.87	.80	.85	.02	-.05	-.08
Consolidating schools						
Support	.90	1.03	.73	.18	.31*	.13*
Oppose	.80	1.00	.83	-.03	.17	.20
Total	11.03	9.43	8.04	2.99**	1.39+	-1.60+

Note: + p < .10; * p < .05; ** p < .01; *** p < .001. $n = 150$ for the CG, $n = 93$ for participants (T2 and T3).

Table 3: Coefficients from generalized hierarchical model estimating effect of treatment controlling for order of questions, treatment, participant knowledge, and education, with random effects for each item, and individual

| | Number of Reasons Offered On an Item | |
| | T3 CG and | T3 CG and |
Predictors	T2 P	T3 P
Treatment	.79***	.62*
Order	-.04	-.04
Treatment*Order	-.09***	.00
Knowledge[1]	1.19***	1.19***
Treatment*Knowledge	-.29	-.55
Education	.54***	.52***
Treatment*Education	.54**	-.20
N	243	214
Log Likelihood	-2733	-2161

Note: + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$; [1]Pre-deliberation measure used for participants and contemporaneously obtained measures for CG; Intercept and random effects not reported.

Table 4: Comparing mean number of reasons given in support and in opposition by people who supported a policy pre-deliberation versus those who opposed it pre-deliberation with supporters and opponents of the respective policies in the Control Group

Measure	Reasons offered at T2		Reasons offered at T3		Reasons offered in Control Group		T2 Minus Control Group		T3 Minus Control Group		T2 – T3 (Participants)	
	T1 Support	T1 Oppose	T1 Support	T1 Oppose	CG Support	CG Oppose	Support	Oppose	Support	Oppose	Support	Oppose
All-ability schools												
Reasons in Support	1.81	1.83	1.28	1.60	1.02	1.45	.78***	.37	.25	.15	.53	.23
Reasons in Opposition	1.34	1.74	.92	1.00	.92	1.36	.42*	.38	.00	-.36	.42	.74
Support Minus Oppose	.47	.09	.36	.60	.11	.09	.36*	.00	.25+	.51	.11	-.51
N												
Balanced Enrollment												
Reasons in Support	1.67	1.64	1.57	1.10	1.21	1.63	.45**	.01	.35*	-.53	.10	.54
Reasons in Opposition	2.00	1.64	1.72	1.50	1.32	1.53	.68***	.12	.40**	-.03	.28	.14
Support Minus Oppose	-.33	.00	-.15	-.40	-.11	.11	-.23	-.11	-.04	-.51	-.18	.40
N												
Inclusive classroom												
Reasons in Support	1.48	1.32	1.14	1.19	1.04	.77	.44*	.55+	.10	.42+	.34	.13
Reasons in Opposition	.94	.83	.82	.79	.88	.54	.06	.29	-.06	.25	.12	.04
Support Minus Oppose	.55	.49	.32	.40	.17	.23	.38*	.26	.15	.17	.23	.09
N												
Consolidating schools												
Reasons in Support	.84	1.06	1.04	1.00	.78	.90	.06	.15	.26	.10	-.20	.06
Reasons in Opposition	.79	.89	1.06	.93	.76	1.10	.03	-.21	.31	-.17	-.27	-.04
Support Minus Oppose	.05	.17	-.02	.07	.02	-.19	.03	.36	-.04	.26	.07	.10
N												

Note: + p < .10; * p < .05; ** p < .01; *** p < .001.

Table 5: Coefficients from generalized hierarchical model predicting number of reasons given on an item using initial knowledge, education, gender, order of the question, attitudinal variance within the group, and mean knowledge of the group, with random effects for each group, individual, and item

Predictors	Number of Reasons Offered On an Item
Order	-.12[+]
Initial Knowledge	1.08**
Education	1.09***
Attitudinal Heterogeneity	1.83*
Group Knowledge	.81
Read Briefing Materials	-.02
N	108
Log Likelihood	-1309

Note: + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$;
Intercept and random effects not reported

Table 6: Testing mediation by change in number of arguments on the impact of participation in Deliberative Poll on attitudes and beliefs

Policy Attitudes	Mediated Effect	Direct Effect
All-ability schools	-	-
Balanced Enrollment	.00	.09*
Inclusive classroom	-	-
Consolidating schools	-	-

Note: + p < .10; * p < .05; ** p < .01; *** p < .001.

Appendix A: Coding of Open-Ended Responses

We classified the arguments offered by the respondents by the substance of the argument. For instance we folded all arguments that dealt with financial impact of the policy under the category 'practical reasons'. The coding was done as follows - a student was asked to come up with initial content categories after being extensively briefed on the schooling issue. This initial scheme was tested by coding of a test sample. Based on the testing, we clarified the use of categories, and added new ones. Once the final coding scheme was decided upon, two undergraduates, blind to the hypotheses being tested, were independently assigned to code the arguments according to the coding scheme described below. Codes chosen by the two coders were largely identical, with both agreeing with the other 9X.5% of the time. Inter-coder reliability was measured using Krippendorff's α, ''the extent to which the proportion of the differences that are in error deviates from perfect agreement, $\alpha = 1$'' (Krippendorff 1980; 1986; 2004). Formally, $\alpha = 1 - \frac{D_o}{D_e}$ ''where Do is a measure of the observed disagreement and De a measure of disagreement that [is] expected due to chance alone.'' The inter-coder agreement was very high, XX. A third coder, also blind to the hypotheses, was asked to adjudicate conflicts and assign responses to a category that s/he thought best suited. The third coder was also asked to check if there were any discernible systematic patterns to disagreement. None were found. Of the reasons that people gave, the relation to the policy proposal was not apparent. For instance, some people mentioned that one of the reasons that people "who strongly support" system of all-ability schools would give is that it would "decrease prejudice". Whether to credit people for such reasons is not clear but since such reasons were given only by a small minority, removing or keeping them doesn't appreciably alter the result.

Off the granular classification, we reclassified some kinds of responses into more coarse bins,

which we use intermittently throughout the paper. We collapsed reasons such as those who held

them are prejudiced under the nominal category of 'giving too little credit to the other side'.

Coding Scheme

All-Ability Schools

Support: Think first about the proposal of having a system of all-ability schools, all providing the same curriculum. Regardless of your own opinion, what reasons would the *people who strongly support this* give for their position?

1. Practical (pecuniary, logistics kind of reasons)
a. Funding
b. Convenience
2. Quality ~ Increases Academic quality
a. Students will be given a better education
b. Students will learn from one another
c. Academic resources will improve
3. Equality ~ Increases equality
a. Students will be given the same opportunities and resources
b. Lessen elitism
4. Decreases Prejudice, Promotes Understanding
a. Increased tolerance
b. Increased integration
c. Ease Sectarian Tensions
5. Other benefits for students (around +11) issues, age and development
a. Less stigma for +11

90. Other
91. Reasons that do not make sense
92. Missing; N/A
93. Vague Responses

Oppose: And what reasons would the *people who strongly oppose this* give for their position?

1. Practical (pecuniary, logistics kind of reasons)
a. Expense
b. Convenience
c. Think that changes will be too difficult to implement
d. Content with Status Quo
2. Academic Quality

a.   Preference for grammar schools
b.   Loss of attention for students, particularly for those in the top and bottom quartiles
3.   Are prejudiced – ignorant, dislike the other religion
4.   Fear prejudice
5.   Fear other problems with mixing
a.   Fear problems (fights, drugs, etc.)
b.   Fear bullying
6.   Self Interest
a.   Grammar schools don't want to give up elite status
b.   Parents of students in Grammar schools don't want to give up status
7.   Prefer single sex schools
8.   Tradition
9.   Other benefits for students
a.   Maintains identity

90.   Other
91.   Reasons that do not make sense
92.   Missing; N/A
93.   Vague responses

<u>Balanced Enrolment</u>

Oppose: Now think about the issue of whether children should attend schools only with other children of their own religion or should attend schools that have a balanced enrolment of Protestant and Catholic pupils. Regardless of your own opinion, what reasons would the *people who think that children should attend schools only with other children of their own religion* give for their position?

1.   Practical (Pecuniary, logistics, etc.)
2.   Academic Quality
3.   Are Prejudiced
a.   Ignorance
b.   Bigotry
4.   Fear prejudice
a.   Fear of repercussions
b.   Loss of standing in community
5.   Fear other problems with mixing
a.   Fear problems (fights, drugs, etc.)
b.   Less bullying
6.   Tradition, Community, Group Identity
a.   Group identity
b.   Sense of community
c.   Worry that mixing students will weaken student's cultural and moral moorings
7.   Want religion in schooling – for example - religion is an integral part of schooling, etc.

90. Other
91. Reasons that do not make sense
92. Missing; N/A
93. Vague responses

Support: And what reasons would *people who think that children should attend schools that have a balanced enrolment of Protestant and Catholic pupils* give for their position?

1. Practical
a. More funding
b. Saves money
c. Better for children of mixed marriages
2. Academic Quality
3. Equality
4. Promotes Understanding in Children and Parents
a. Increased tolerance
b. Increased integration
c. Better social development/ mirrors real world
d. Greater understanding
e. Way to break violence cycles
5. Religion shouldn't be part of public schools ~ as religion is a personal matter, has no place, or doesn't need to there in schools.

90. Other
91. Reasons that do not make sense
92. Missing; N/A
93. Vague responses

School Collaboration

Support: Now think about the issue of whether, if schools with different religious compositions enter partnerships, the children from both schools should at least sometimes be taught in the same classroom. Regardless of your own opinion, what reasons would the *people who strongly agree with this* give for their position? Provide up to five reasons in the spaces below.

1. Practical
a. Makes Economic Sense
2. Academic Quality
3. Equality
4. Promotes Understanding
a. Increased tolerance
b. Increased integration
c. Better social development/ mirrors real world

d.   Less conflict
e.   Better understanding
5.   Small Step Towards Integrated Schools

90.   Other
91.   Reasons that do not make sense
92.   Missing; N/A
93.   Vague responses

Oppose: And what reasons would the people *who strongly disagree with this* give for their position? Provide up to five reasons in the spaces below.

1.   Practical (Pecuniary and logistics)
a.   Disrupt School efficiency
2.   Quality
3.   Are prejudiced
4.   Fear Prejudice
a.   Bigotry
b.   Fear
c.   Ignorance
d.   Could increase tension / segregation
5.   Other problems of mixing
a.   Fears of bullying, being discriminated
6.   Only a half-measure to Integrated Schools
7.   Prefer single sex classrooms, opposed to gender mixing
8.   Tradition, Religious and Moral Focus in School
a.   Sense of community
b.   Favor their own religion
c.   Worry that mixing students will weaken student's cultural and moral moorings
d.   Child may lose a sense of Identity
e.   Indoctrination – children may be indoctrinated

90.   Other
91.   Reasons that do not make sense
92.   Missing; N/A
93.   Vague responses

Consolidating Schools

Support: Now think about the proposal for schools combining primary and post-primary pupils (for example, ages 7-14). Regardless of your own opinion, what reasons would the *people who strongly support schools combining primary and post-primary pupils* give for their position? Provide up to five reasons in the spaces below.

1.   Practical (pecuniary, logistics kind of reasons)

a. Easier transportation
b. Easier for parents
c. Keep rural schools open
d. Earns school more money
e. More Cost-Effective
2. Quality
3. Individual student level reasons: Career, stress, personal
a. Lets children mature before choosing specialization
b. Children too young at 11 – less stress/pressure
c. Improve performance
d. Prevents children from changing schools so young
e. Siblings and friends stay together longer
4. Equality or do away with +11
90. Other
91. Reasons that do not make sense
92. Missing; N/A
93. Vague responses

Oppose: And what reasons would the *people who strongly oppose schools combining primary and post-primary pupils* give for their position? Provide up to five reasons in the spaces below.

1. Practical
a. Don't think that the schools or teachers are equipped to handle the change
b. Need new teachers
c. School closings?
d. Operational Inefficiency
e. Too Many School Changes
2. Academic Quality
a. Would increase class size
3. Individual student level reasons: Career, stress, personal
a. Helps younger students mature
b. Improve performance
c. Prevents children from changing schools so young
d. Older children not stimulated by the company of the younger ones
e. Older Students Ready to Move on
f. Some Children too Young to Change Schools
4. Fear Other problems of mixing
a. Bullying
b. Younger students may feel uncomfortable around 14 year olds
5. Satisfied with status quo
6. Tradition

90. Other
91. Reasons that do not make sense

92.   Missing; N/A
93.   Vague responses

Appendix B: Knowledge Items

Below are the items with which factual knowledge was measured. Correct answers are in parenthesis.

1.   What percentage of majority-Protestant or majority-Catholic schools in Northern Ireland have at least 10% of the other religion in their enrolment? - 40-50%, 20-30%, 5-10%, Less than 1% (5-10%)

2.   By approximately what percentage has the number of children entering Omagh schools increased or decreased over the past five years? Increased by 20%, Increased by 10%, Stayed about the same, Decreased by 10%, Decreased by 20% (Decreased by 10%)

3.   The new entitlement framework requires that… Every school provides all 14-year-olds with a choice of at least 24 subjects, Every child has the right to attend any school his or her parents wish, Every child be provided a free school meal every school day, Every child receives free textbooks (Every school provides all 14-year-olds with a choice of at least 24 subjects)

4.   The new entitlement framework requires that… Every child receives tuition in the language of his or her parents' choice, Every child receives free transportation to and from school, Every denominational group has the right to run its own schools, One-third of all the subjects offered must be applied (One-third of all the subjects offered must be applied)

5.   Which of the following is true of what pupils in Northern Ireland do after they leave school? - About one-quarter go directly into employment, About one-quarter leave school to be unemployed, About three-quarters of grammar school pupils attend university, About three-quarters of secondary school students attend Further Education College (About three-quarters of grammar school pupils attend university)

6.   Which of the following is true of current school funding? - Schools receive more funding for older pupils, Schools receive the same funding for all pupils, regardless of age, Schools receive more funding for younger pupils (Schools receive more funding for older pupils)

7.   Which of the following is true of the employing authority in the schools? –The official employer for all teachers is the Education and Library Board, The official employer for all teachers is the Department of Education, The official employer for all teachers in Catholic schools is the CCMS (Council for Catholic Maintained Schools), The official employer for all teachers in voluntary grammar schools is the school's Board of Governors (The official employer for all teachers in voluntary grammar schools is the school's Board of Governors)

Appendix C: Difference between the entire T2 participant sample, and Control Group

Measure	T2	T3 CG	T2 – T3 CG
Support All-ability schools			
All	1.77	1.05	.72***
Catholic	1.81	1.04	.77***
Protestant	1.63	1.11	.53
Oppose All-ability schools			
All	1.38	.95	.43*
Catholic	1.48	1.00	.48*
Protestant	1.22	.86	.35
Support Balanced Enrollment			
All	1.70	1.27	.43**
Catholic	1.71	1.38	.33+
Protestant	1.68	1.00	.68*
Oppose Balanced Enrollment			
All	1.98	1.35	.63***
Catholic	2.03	1.40	.63**
Protestant	1.78	1.19	.59+
Support Inclusive classroom			
All	1.37	1.02	.35*
Catholic	1.39	1.10	.30
Protestant	1.34	.81	.53+
Oppose Inclusive classroom			
All	.86	.85	.02
Catholic	.82	.94	-.12
Protestant	.98	.54	.44+
Support Consolidating schools			
All	.93	.73	.20
Catholic	1.00	.77	.23
Protestant	.83	.57	.26
Oppose Consolidating schools			
All	.81	.83	-.01
Catholic	.81	.93	-.12
Protestant	.83	.59	.23
Total			
All	10.81	8.04	2.77**
Catholic	11.05	8.56	2.49*
Protestant	10.29	6.68	3.62*

Note: + p < .10; * p < .05; ** p < .01; *** p < .001. $n = 150$ for the CG, $n = 124$ for participants (t2 and t3). Four of the participants, and eight of the CG respondents, did not provide their religious affiliation. These 11 were omitted from sub-group analyses. Among participants, $n = 79$ for Protestants, and $n = 41$ for Catholics. Among CG, $n = 105$ for Protestants, and $n = 37$ for Catholics.

Appendix D: Re-evaluating effect of the DP

If the small group composition matters, as we show above, and as prior research (Luskin et al. 2009) has shown, then the effect of the DP may depend on particular random assignments to small groups. For the same sample, different random assignments to small groups may produce different effects. Traditional estimates of effect of an experiment, constrain potential outcomes to two - when a subject is treated, and when she is not. However here, the potential outcomes of each person are a great many more because treatment here varies by the small group. Hence, to estimate a more general range of 'effects' of a DP – conditional on the same sample, and other fixed characteristics of the treatment (outside of particular random assignment to small groups) – we must simulate all possible random assignments to small groups and evaluate the final mean using a model that captures how small group composition matters. Assuming the model used to generate Table 5 captures how small-group composition matters[1], we estimate a distribution of effects. The effect of the DP (comparing the entire T2 participant sample to the CG) lies between X.XX and X.XX.

---

[1] Small sample size, small number of groups, limited variation across groups, limited data on individuals, among other things, limit our ability to precisely estimate a variety of ways by which small group composition may matter. All these are potential liabilities because the conservative thing to do here may be to reject the null.

---

[1] It is tempting to think that "negativity bias" may make the parameters in the na equation exceed their counterparts in the np equation, but recall that anti-A includes pro-B.

[2] This trend is at least partly driven by preferences among partisans to move to areas with higher concentrations of co-partisans (Cho et al. 2012; McDonald 2011)

[3] If, for example, 80% of the arguments a person already knows are supporting, compared to only 60% of the arguments he or she learns from deliberating (more than half but a smaller proportion), the deliberation will have made his or her argument pool less imbalanced.

[4] Or, more precisely, public housing is divided along nationalist lines. The conflict in Northern Ireland is not a religious conflict. It is a conflict between two national groups, British unionists and Irish nationalists, who are in dispute over whether Northern Ireland should remain part of the United Kingdom or become part of a united Ireland. Yet since most unionists are Protestant and most nationalists Catholic, 'Protestant' and 'Catholic' are used as popular shorthands for British unionist and Irish nationalist, respectively (on religious identity as a 'marker' of national conflict, see McGarry and O'Leary 1995, p. ?).

[5] There are 53 Northern Ireland Office maintained peace lines in four towns and cities in the region -- 42 in Belfast, five in Londonderry, five in Portadown and one in Lurgan. For a useful discussion, see http://news.bbc.co.uk/1/hi/northern_ireland/8121362.stm [Accessed January 2012].

[6] Participants learn a great deal between the time they are first sampled (T1) and when they arrive at the deliberation site (arrival). However, we do not have an arrival questionnaire. Lacking that, T1 knowledge scores are used as a proxy for arrival knowledge.

[7] The percentages for those who had not heard about the DP are impressive but not surprising. For reasons already explained (particularly, the need to consolidate the school system in response to the declining birthrate) education reform was a major news story at around the time DP was conducted.

[8] We predicted activity (average score on the three types of activities) using education and treatment. Treatment was significant. Other more elaborate specifications produced similar results.

[9] Results are available from the authors upon request.

[10] It is plausible that the difference between control group and treatment sample is in not in the rate of the decline of motivation but in the initial intercept – that the 'treatment effect' is an artifact of differential motivation across treatment and control samples. However, we have no power to disambiguate whether intercept differences are 'true' or due to varying motivation.

[11] There are far fewer Protestants in the sample than Catholics. And non-significant results among Protestants, as opposed to significant results among Catholics, are partly a consequence of the fewer Protestants.

[12] Regressing learning of specific kinds of arguments on the mean position on the policy held by the group proved XXXX.

[13] Results are available from the authors upon request.

[14] Given attitude is measured at the same time as arguments offered, there is no way to disambiguate whether arguments were antecedent to attitude change, or whether they are post-hoc justifications.

[15] The argument traces its ancestry to Luskin, Fishkin, and Jowell 2002, that postulates and finds

evidence consistent with factual knowledge as a mediator of attitude change. We differ here in the specification here as, unlike Luskin, Fishkin, and Jowell 2002, we formally test mediation.

[16] For any given individual, this idealized counterfactual needn't always be far from observed opinion – in fact in a dichotomous choice condition, a random toss will half the time net the person her "true preference".