European Journal of Political Research ••: ••–••, 2012
doi: 10.1111/j.1475-6765.2011.02052.x

Which decision-making arrangements generate the strongest legitimacy beliefs? Evidence from a randomised field experiment

PETER ESAIASSON, MIKAEL GILLJAM & MIKAEL PERSSON
University of Gothenburg, Sweden

Abstract. How can democracies satisfy citizens' demands for legitimate decision making? This article reports findings from a randomised field experiment designed to mimic decision making in large-scale democracies. Natural collectives of individuals with a shared history and future (high school classes) were studied. They were asked to make a decision about how to spend a sum of money under arrangements imposed by the researchers and distributed randomly across classes. Within this setting, empirical support for three ideas about legitimacy enhancing decision-making arrangements is tested: participatory constitution-making; personal involvement in the decision-making process; and fairness in the implementation of arrangements. Throughout the analyses it was found that personal involvement is the main factor generating legitimacy beliefs.

Keywords: democratic decision making; legitimacy beliefs; participatory constitution making; participatory democracy; procedural fairness; randomised field experiments

How should democracies go about satisfying citizens' demand for legitimate decision making? Which decision-making arrangements generate the strongest legitimacy beliefs? Questions like these are best addressed through cross-institutional comparisons. Since institutional arrangements are hard to change, most empirical studies on their legitimising capacity rely on natural variations (e.g., Norris 2004; Grofman & Reynolds 2001) or on surveys of citizens' preferences (e.g., Hibbing & Theiss-Morse 2002; Bowler et al. 2007). While such studies have generated important insights, endogeneity problems encourage the use of complementary methodological approaches.

Within the social sciences the prime way to deal with endogeneity problems is to conduct experiments in which, prototypically, subjects are randomly assigned to a treatment group or a control group (Morton & Williams 2010; Druckman et al. 2011a). During the past decade and half, political scientists have increasingly turned to the experimental method to study causality within a wide range of political phenomena (Druckman et al. 2011b). Like all other methods, experiments are associated with certain limitations – for the present research at least, the most pressing issue relates to external validity – but if these are acknowledged, experiments can help to sort out causal relationships

Published by Blackwell Publishing Ltd., 9600 Garsington Road, Oxford, OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

in research fields that are dominated by observational methods (e.g., Kinder 2011).

This article reports findings from an exploratory randomised field experiment designed to replicate, in miniature, decision-making arrangements found in large-scale democracies. Our substantive aim is to evaluate the relative effectiveness of three broad ideas about legitimate decision making: citizen participation in the making of fundamental law as in *participatory constitution making* (e.g., Fehrenbacher 1989); personal involvement in the decision-making process itself as in *direct democracy* (e.g., Pateman 1970; Barber 1984); and fairness in the implementation of institutional arrangements as in *procedural justice theory* (e.g., Lind & Tyler 1988; Tyler et al. 1997). In order to examine the legitimising capacity of these ideas we manipulate a wide range of arrangements such as election-based representation, expert decision making, fairly and unfairly implemented direct voting, and participatory constitution making. Furthermore, we manipulate two theoretically important but rarely implemented forms of decision making: lottery based representation, and consensual decision making.

Our design integrates features of three lines of experimental research. Like field experimental research, we randomise natural collectives of individuals with a shared history and future (Gerber & Green 2009). Like judicial legitimacy research, we manipulate decision-making arrangements that are present in large-scale democracies (Gibson et al. 2005). Like experimental economics, we use money to create an incentivised environment (Palfrey 2009).

A major challenge for field experimental research in this domain is to find a sufficient number of natural collectives to allow for random assignment. In this study we have turned to a high school setting. A total of 21 classes containing 484 individuals – all from a single, well-functioning high school in Sweden – were each provided a sum of money and asked to decide collectively whether to spend it on charity or on an activity for their own pleasure. Decision-making arrangements were determined by us and distributed randomly across classes. Following the decision, participants assessed the fairness of the arrangement. We take these assessments as our measure of legitimacy beliefs associated with the decision-making arrangement used. The field experiment was preceded by a series of studies in which corresponding decision-making arrangements were manipulated by means of vignettes, and which involved both students and adult participants. We use findings from these experiments for validation.

In what follows, we first develop the rationale for the respective ideas about decision-making arrangements. Thereafter, we discuss our experimental manipulations along with other study details. We then present the results, with the findings indicating that personal involvement through direct majoritarian

voting is the most effective way to generate legitimacy beliefs. The concluding section discusses the validity of our suggested approach.

Decision-making arrangements and legitimacy beliefs

Participatory constitution making, personal involvement and fair implementation of arrangements are three broad ideas for legitimacy-enhancing decision-making arrangements. While each idea can be realised in a multitude of ways, their focus can be ordered along a continuum: how to decide how to make specific decisions; how to make specific decisions; and how to apply the rules for the making of specific decisions. Moreover, each idea figures regularly in the literature, and our current knowledge about their legitimising capacity emanates largely from observational data.

Participatory constitution making

Theory concerning constitutionalism considers the higher law of constitutions to be a major source of democratic legitimacy. It assumes that the right to make authoritative decisions derives from a body of fundamental law codified in the constitution (e.g., Fehrenbacher 1989). Traditionally, constitution making has been the privilege of politicians and judicial experts. Lately, however, practitioners and scholars have begun to show interest in reforming this process by acknowledging democratic values such as transparency and citizen participation (e.g., Benomar 2004; Moehler 2008).

Participatory constitution making can generate legitimacy beliefs through two different mechanisms. The first focuses on citizens' ability to choose effective constitutional alternatives; it works to the extent that citizens are better than expert constitution makers at choosing effective arrangements. The other mechanism focuses on the additional value of taking part in the constitution-making process itself; it works to the extent that arrangements generate stronger legitimacy beliefs when chosen endogenously than when exogenously imposed. The latter mechanism is the one most often addressed in the literature.

Participatory constitution-making processes have been used in developing democracies in sub-Saharan Africa (Uganda and Kenya), in post-communist Europe, in established democracies (Canada and the Netherlands) and within the European Union. Empirical evaluations of their success are scarce, but a study of the Ugandan case identifies a number of obstacles regarding, for example, citizen access to neutral information (Moehler 2008). The findings from laboratory experiments on social dilemma situations are more

favourable: using a public goods game, Sutter et al. (2010) find that individuals are more willing to cooperate when they make their own institutional choices than when arrangements are determined exogenously by the experimenter. Hence:

> *H1*: Individuals will bestow more legitimacy upon arrangements of their own choice than upon exogenously imposed arrangements.

Personal involvement

Participatory democrats have long argued the benefits of personal involvement in decision-making processes. The basic idea, advocated by Jean Jacques Rousseau as well as by recent theorists, is that 'being one's own master' increases willingness to play along with collectively binding decisions (Pateman 1970: 26–27; Barber 1984). In general, individuals are expected to appreciate the feelings of control, shared responsibility and mutual respect that come with personal involvement. In a similar vein, referring not only to politics but to life in general, political philosophers value people's right to make their own choices, even when these fail to promote personal well-being (Scanlon 1998; Dworkin 2000; Duus-Otterström 2011).

In large-scale democracies, the decision-making arrangement with the highest degree of personal involvement is direct voting (as in referendums and in people's initiatives) (Smith & Tolbert 2004). Alternative and less participatory forms of authoritative decision making – which of course are used much more frequently – involve elected representatives in legislative assemblies, and experts in the form of judges in courts and administrators in state agencies. Indeed, direct majoritarian voting, representation and expert decision making are the three generic forms of decision making in democracies.

Empirical research that explicitly addresses the legitimising capacity of direct majoritarian voting is scarce. However, literature reviews usually side with the claims of participatory democrats (e.g., Lupia & Matsusaka 2004). A field experiment, in which a large number of Indonesian villages were randomly assigned to make a collective decision through either representative-based meetings or direct voting comes to the same conclusion (Olken 2010). Vignette experiments that compare individuals' reactions towards direct legislation to other forms of decision making provide further confirmation (Gash & Murakami 2009).[1] Hence:

> *H2*: Individuals will bestow most legitimacy upon arrangements that allow personal involvement in the decision-making process.

Because our design enables detailed manipulations we can derive an additional, and we believe original, implication of the idea of personal involvement. According to democratic elitism theory (Gibson & Duch 1991), elected representatives bestow more legitimacy upon their political system than do the citizens they represent (Sullivan et al. 1993). This is ascribed to higher levels of education and other high-status social characteristics, as well as to socialisation effects (e.g., Rohrschneider 1996). However, it can be noted that elected representatives also differ with regard to involvement; for them, participation in decision making is an everyday matter. Our approach enables us to evaluate whether personal involvement affects representatives' legitimacy beliefs while keeping social origin and socialisation under control. Specifically, *H2* will be supported to the extent that we find stronger legitimacy beliefs among participants who are selected as representatives than among those who are not.

With regard to expert decision making versus representation, the latter allows more citizen participation. This notwithstanding, research on judicial legitimacy typically finds that American citizens ascribe more legitimacy to court decisions than to decisions taken by elected representatives (Gibson et al. 2005). In accordance with citizen scepticism towards representation, stealth democracy theory maintains that the average citizen associates political processes with bickering, personal interest, conflicts and inefficiency (Hibbing & Theiss-Morse 2002). Moreover, attitudinal studies from newly established democracies in Eastern Europe find that many citizens prefer to have authoritative decisions taken by experts rather than by elected representatives (Rose et al. 1998). Overall, while empirical evidence does not motivate a clear hypothesis, representative decision making appears to have a relatively low legitimising capacity.

To draw even further on the analytical leverage offered by the experimental method, we test support for two complementary ideas that are theoretically important but rarely implemented in large-scale democracies. First, over the past decades, deliberative democratic theory has emerged as a source of inspiration for democratic renewal (e.g., Chambers 2003). The deliberative ideal overlaps to some degree with participatory ideals, but it puts more emphasis on the quality of public argumentation. Because of its belief in the power of the better argument, for some theorists its preferred mode of decision making is consensual (Cohen 1997). While there are sceptical voices (see Thompson 2008), we identify consensual decision making as a subcategory of direct decision making and test empirical support for the following:

> *H3*: Individuals will bestow more legitimacy upon arrangements for direct consensual decision making than upon direct majoritarian voting.

Second, Manin (1997) points out that elections are a partly elitist way to select representatives. To be elected, an individual must be perceived as more qualified than others with regard to relevant characteristics. The more egalitarian way of selecting representatives, which was first practiced in ancient Athens, is to appoint them by lottery. The idea of selecting representatives through lottery has drawn the interest of political theorists (Stone 2009), but it has rarely been tested empirically. Arguably, since lotteries give every citizen an equal chance of being selected, lottery-based representation allows for a higher level of personal involvement than representation. Hence:

> *H4*: Individuals will bestow more legitimacy upon lottery-based representation than upon election-based representation.

Fairness of the actual decision-making process

Normative theory on authoritative decision making (Christiano 1996; Estlund 2008) and empirical theory on procedural justice (Lind & Tyler 1988; Tyler et al. 1997) agrees that fair applications of decision-making rules are crucial for legitimacy beliefs. In particular, procedural justice research maintains that people care about procedural fairness because it is a moral right fulfilled, it assures them that they are respected by the decision-making authority, it indicates that the decision-making authority is trustworthy and it enables them to reduce uncertainty about the fairness of substantial outcomes.

Numerous empirical studies confirm these claims (see MacCoun (2005) for a review). Specifically, observational studies find a strong positive association between the perceived fairness of procedures and various indicators of legitimacy across policy domains and national contexts. Moreover, experimental studies demonstrate that actual procedures that are non-biased and non-arbitrary, and that allow participants to voice their views, generate stronger legitimacy beliefs than procedures with the opposite qualities. Hence, we expect the following to be true:

> *H5*: Individuals will bestow more legitimacy upon decision-making arrangements that are fairly implemented than upon procedurally flawed arrangements.

The dependent variable: Procedural fairness assessment

However defined, 'legitimacy' belongs to the family of inherently abstract concepts that are hard to measure directly (Thomas 2010). We adopt a psychological (subjective) understanding of the concept (Tyler 2006), and we

address it indirectly by gauging fairness assessments of each decision-making arrangement. Such subjective legitimacy beliefs play a central role in psychological legitimacy theory. According to this theory, assessments of decision-making procedures are of key importance to the perceived legitimacy of authorities and institutions, which in turn affect peoples' willingness to accept their decisions and rules (Tyler 2000, 2001, 2006; Gangl 2003).

Procedural fairness assessments are captured by two items: 'How fair do you think matters were when the decision was taken?', and 'How fairly do you think you were treated when the decision was taken?' Allowing for minor variations of wordings, these are standard indicators in procedural fairness research (e.g., Skitka et al. 2003). For both items, responses were registered on a seven-point scale with designated endpoints 'not fair at all' and 'very fair'. As responses were highly correlated (0.72), the dependent variable is the arithmetic average of the two items.

Design of the study

The main study was sited in a single high school located in Gothenburg, Sweden. The school is fairly large with a student population of about 1,000, and it is generally well functioning (students' graduation marks are slightly above the national average). In the Swedish educational system, high school classes share a common curriculum for the better part of the three years of secondary education (Persson & Oscarsson 2010), and the class is therefore a meaningful unit for students.

The siting of the study allows us to study natural collectives of people with a shared history and future rather than individuals brought together only to participate in a social science inquiry. Nevertheless, the young age of participants is a concern, as is the choice of national context. Developmental psychologists have demonstrated that high school students from different cultural contexts make judgments about democratic government similar to adults (see Helwig et al. (2007) and the literature cited therein). Moreover, university students are often used as subjects in laboratory experiments. In this comparison, high school students are probably more diverse than university students. In addition, we will use our vignette experiments, which involved both students and adult participants, for empirical validation. With regard to national context, Sweden is a fairly typical representative strong-party democracy in which representational decision making is the default arrangement.

Following the lead of experimental economics, we used money to create an incentivised environment (Palfrey 2009). However, while experimental

economics rewards participants individually, participants in our experiment were rewarded collectively, following the outcome of the decision-making process. This incentive structure was set up to ensure that participants cared about the outcome of the decision. Specifically, each class was asked to use a sum of money for either charity (Doctors Without Borders was the designated recipient) or for a festivity of their own choosing. The average sum allocated to each class was the equivalent of US$290 or €250.[2]

The choice between charity (altruism) and a festivity (material well-being) represents a decision over distributive policies. While the type of policy decision is a constant in our field experiment, we can use findings from our vignette experiments to see whether results generalise to regulatory policies as well. Indicating that they do, student participants in these experiments made similar legitimacy assessments whether they were considering regulatory policies on banning religious symbols in schools and allowing teachers to seize students' mobile phones, or distributive policies on spending money from fundraising activities on charity or on a school trip (Gilljam et al. 2010).

We recruited 21 classes from the school, and then we randomly assigned each of seven treatments to three classes with an average of 23 individuals per class. Precisely, we decided in advance the number of classes to be studied, and then randomly decided the order in which each treatment condition should be applied. Acting independently from us, and often on short notice, teachers contacted experimenters with a go-ahead for a trial in their particular class.

In such cluster randomised designs, statistical precision is affected by both the number of clusters per treatment and the number of individuals within clusters (Raudenbush 1997). Because we wanted to contrast several arrangements within a unified framework, and findings from our vignette experiments were clear and consistent, we settled for a small number of clusters per treatment (three), and thus for a design of relatively low statistical power. As results are generally unambiguous, our design strategy worked out satisfactorily.[3]

A series of nested analyses of variance (ANOVA) indicate a successful randomisation. Results show that the observed variables 'gender', 'particular horizontal trust (in classmates)', 'political interest', 'self-reported left-right placement', 'parents' social class' and 'immigrant origin' were about equally distributed across treatment conditions ($p > 0.05$). From previous studies we know that getting one's preferences fulfilled has a strong impact on perceived legitimacy (Esaiasson 2010). Within the respective treatment conditions, the proportion of outcome winners (whose initial preferences were fulfilled) and outcome losers (whose initial preferences were denied) were equally distributed.

Experimental proceedings

The experiment was conducted in the classroom of each class, and lasted on average for 50 minutes. With some deviations between treatments (see the online appendix for details), the experiment proceeded as follows. In the introductory phase, experimenters declared that the study was part of a research project on decision making, and then informed participants about the decision they were about to make. Participants then reported their initial preference regarding the use of money in a written questionnaire. Following this, experimenters primed participants to consider institutional arrangements by reminding them that collective decisions can be taken in different ways. Thereafter, the manipulation was introduced as participants were told which arrangement would be used in their class. This sequencing mimics real-world politics in which typically decision-making arrangements are known to citizens at an early stage of the policy-making process.

Experimenters then initiated a discussion about the pros and cons of the two alternative outcomes – that is, giving money to charity or using the money for a festivity of their own choosing. This was done because we wanted to hold pre-decision deliberation constant at a relatively high level.[4] When the intensity of the discussion began to fade (usually after 20 minutes), participants were again surveyed about their personal preference, and the experimenters organised the decisive vote. The clear majority of classes (16 of 21) decided to give the money to charity. As measured by post-discussion attitudes, 132 participants (28 per cent) were denied their preferred outcome (according to pre-discussion attitudes, the number of losers thus defined was slightly higher at 159). We find no evidence that reactions towards the decision differ between those who preferred one or the other outcome.

After learning about the outcome, participants were asked further survey questions about their judgments regarding the outcome, procedural fairness assessments and background information. There was no formal de-briefing, but participants were promised (and later given) feedback on study results.[5]

Treatment conditions

To evaluate support for our hypotheses, we constructed five main and two complementary treatment conditions. Table 1 presents an overview of the seven conditions, and how they relate to the hypotheses.

Participatory constitution making (*H1*) is represented by a treatment (T1) in which participants were initially asked to decide how they wanted the substantive decision to be made. They were offered a menu of decision-making arrangements: direct voting (secret ballot vote); consensual; representation;

Table 1. Experimental treatments

Participatory constitution-making (*H1*)
T1. Participants choose their preferred decision-making arrangement.
Personal involvement (*H2*)
T2. Direct majoritarian voting (secret ballot vote).
T3. Election-based representation (three representatives were elected by participants).
T4. Expert decision making (decision taken by the class teacher in his or her role as expert administrator).
Consensus and lottery-based representation (complementary hypotheses *H3* and *H4*)
T5. Consensual decision making (unanimous agreement among participants).
T6. Lottery-based representation (three representatives were randomly selected).
Fairness in the implementation of arrangements (*H5*)
T7. Experimenters corrupted an arrangement for direct majoritarian voting.

and expert decision by their teacher. Following this, each class used the preferred arrangement to make the decision. All classes assigned to this treatment opted for direct decision making, either through majoritarian voting (two classes) or through a consensual process (one class). A detailed description of this and other treatments is provided in the online appendix.

To evaluate the idea of personal involvement in the decision-making process (*H2*), we constructed three treatments, one for each generic arrangement for decision making. In classes assigned to these treatments, the decision was made by direct voting (T2), by three representatives who were elected by participants (T3) or by the teacher in his or her role of expert administrator (T4). These arrangements (as well as T5 and T6) were imposed by the experimenters and implemented in a fair way.

The complementary idea about consensual decision making (*H3*) is represented by a treatment (T5) in which participants were instructed to make a unanimous decision. In this condition, the discussion continued until all participants agreed to support one of the alternatives. The complementary idea about lottery-based representation (*H4*) is represented by a treatment (T6) in which the decision was made by three representatives selected by lottery.

Finally, the idea of (un)fair implementation of decision-making procedures (*H5*) is represented by a treatment (T7) in which experimenters corrupted the arrangement. For practical reasons we opted to corrupt only direct majoritarian voting, which was hypothesised to generate the strongest legitimacy beliefs. Following the example of Van Prooijen et al. (2006), we manipulated the neutrality of the mediating third party (the experimenters). After the

announcement of the outcome of a fairly implemented majoritarian vote, experimenters started to question the epistemic quality of the decision and provided additional information that favoured the losing alternative. After a renewed discussion, participants were forced to vote a second time.[6]

Although not full factorial, the design enables us to contrast each experimental treatment with theoretically relevant comparison groups. The logic of our analysis – the exact comparison of groups – is presented in the results section.

As a robustness check we present complementary findings from a series of experiments in which decision-making arrangements were manipulated by means of vignettes. These studies contribute variation on three contextual factors: type of policy decision (regulatory policies and distributive policies); age and type of participants (high school students, university students, teachers, and adult citizens); and type of decision-making setting (high schools and large-scale democracies). Most importantly for the purpose of generalisation, for one set of studies, which all deal with a regulatory policy proposal (to ban the use of religious symbols in schools), experiments cover both high schools and large-scale democracies, and both student and adult participants. In the school setting, the decision was taken by students in a referendum, by elected representatives in a student council or by the collective of teachers. Correspondingly, in the large-scale democracy setting the decision was taken by citizens in a local referendum, by elected representatives in the city council or by expert administrators in the local school board. In the school setting participants were high school students (n = 387) and adult teachers (n = 201). In the large-scale democracy setting, participants were political science students at two Swedish universities (n = 132), and a convenience sample of the general public recruited at the central station in the city of Gothenburg (n = 451).

In these and other vignette experiments referred to in the results section, participants were given an identically worded text introducing them to the policy proposal in question. Thereafter they received randomly distributed information about the procedure for decision making, and about the outcome of the procedure. Following this, they assessed the fairness of the procedure using identical items to those in the field experiment.[7]

Results

Figure 1 presents the observed level of legitimacy belief associated with each treatment condition. Importantly, our manipulations generate different levels of legitimacy beliefs, with a high of 6.5 and a low of 5.0 (T2 and T3, respectively). In absolute terms, and given that the dependent variable ranges from 1 to 7, this corresponds to a substantial 25 per cent of the maximum effect size.

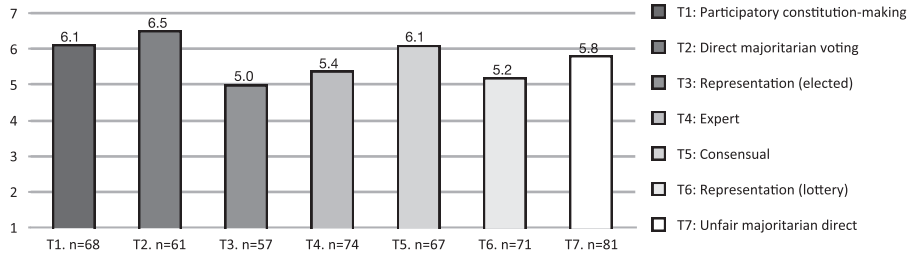

Figure 1. Observed mean levels of legitimacy beliefs.

By looking at the observed means we learn moreover that treatments that allow for some form of personal involvement (T1, T2, T5 and T7) generate higher levels of legitimacy beliefs than representation (T3 and T6) and expert decision making (T4).

Observed means are indicative, but systematic hypothesis testing requires us to consider the hierarchical character of the data (Snijders & Bosker 1999). To account for this, we estimated an intercept-only model with individuals (i), in classes (j):

$$\text{Procedural fairness assessment}_{ij} = \alpha + e_{ij} + u_j \quad (1)$$

The intra-class correlation for this model shows that 16 per cent of the total variance is located at the level of classes (see Table 1 in the online appendix).[8] Indicating a reliable data structure, this is within the range of between cluster variations that are found in American studies on student achievement (Hedges & Hedberg 2007, cited in Konstantopoulos 2009).

To facilitate interpretation of the different treatment contrasts, the main text will report the predicted level of procedural fairness assessment associated with respective treatment and comparison group as yielded by our multilevel models. To increase statistical efficiency, and to correct for remaining imbalances on observed measures post-randomisation, we include covariates that might explain participants' initial attitudes towards decision-making arrangements (Duflo et al. 2008; Gerber et al. 2010). Specifically, we include gender, political interest, outcome winner/loser status, generalised horizontal trust, vertical trust (in teachers and principal) and particular horizontal trust (in classmates) as individual-level covariates, and grade level as a cluster-level covariate.[9] Readers should refer to the online appendix for detailed statistics.

H1*: Participatory constitution making*

Our test of *H1* compares the level of legitimacy beliefs associated with participatory constitution making (T1) to two different comparison groups. The

first consists of exogenously imposed direct voting, election-based representation and expert decision making (Treatments 2, 3 and 4). While this mechanism is rarely discussed in the literature, the comparison will tell us whether participants choose decision-making arrangements that are more effective than the average for the three generic alternatives.

Figure 2 displays the predicted levels of legitimacy beliefs and error bars illustrating 95 per cent confidence intervals for T1 and the comparison group thus defined. *H1* gains some, albeit weak, support from the results. In accordance with predictions, procedural fairness assessments are higher for the endogenously chosen decision-making alternative (6.2) than for the average of the generic alternatives (5.7) that are exogenously imposed, but the difference barely reaches a standard level of statistical significance ($p = 0.09$, one-tailed).

The second comparison group, which reflects the core mechanism of the idea, consists of the same form of decision making as was chosen by participants in T1, but imposed exogenously by experimenters. This comparison will tell us whether participation in the constitution-making process adds legitimacy over and above exogenously imposed arrangements. As all three classes assigned to T1 opted for direct decision making – either a majoritarian secret ballot vote (two classes), or consensual (one class) – the relevant comparison groups are T2 and T5, respectively.

Results run against the hypothesis. In fact, the predicted level of procedural fairness assessments is slightly higher when the arrangement is exogenously imposed (6.4 versus 6.3 for majoritarian secret ballot vote, and 6.2 versus 6.0 for consensual decision making) albeit insignificant.[10] Thus, contrary to the claims of those advocating reform of constitution-making processes, and contrary to findings from experimental economics (Sutter et al. 2010), there is no support for the idea that participation in constitution making in and of itself generates stronger legitimacy beliefs.

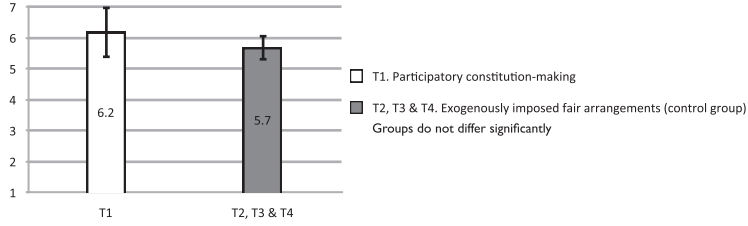

Figure 2. Legitimacy beliefs associated with participatory constitution-making: predicted values and 95 per cent confidence intervals.
Note: Estimates are derived from model 2 presented in Table 2 in the online appendix.

H2*: Personal involvement*

From the idea of personal involvement (*H2*) follows the prediction that direct decision making will generate stronger legitimacy beliefs than representation and expert decision making. To determine whether this is supported by our data, Figure 3 displays the predicted level of procedural fairness assessments associated with T2 (majoritarian direct voting), T3 (election-based representation) and T4 (expert decision), respectively. To test a further implication of the idea of personal involvement, Figure 3 also includes corresponding information for participants (18 out of 128) who were selected as representatives in T3 and T6 (election-based and lottery-based representation).

Our findings provide strong support for *H2*. Majoritarian secret ballot vote is associated with significantly higher fairness assessments than the two main alternatives (6.4 compared to 5.4 (expert) and 4.7 (representation), respectively, $p < 0.05$, one-tailed). Our randomised field experiment thus generates data that are consistent with the claims of participatory democrats and recent field experiments (Olken 2010).

Adding further support to the hypothesis, the small group of participants who were elevated to the status of representative report equally high levels of procedural fairness assessments as those who made the decision by majoritarian secret ballot vote (6.3 versus 6.4). While much previous research attributes a high level of system support among elected representatives to their high-status social origin and to socialisation factors (Sullivan et al. 1993; Rohrschneider 1996), the findings here suggests that representatives' personal involvement in the decision-making process plays a causal role as well.

With regard to the relative effectiveness of representational and expert decision making, our findings indicate a relatively weak position for representational decision making. A comparison of T3 and T4 shows that procedural fairness assessments are lower for representational decision making than for expert decision making ($p = 0.06$, two-tailed). This finding confirms previous

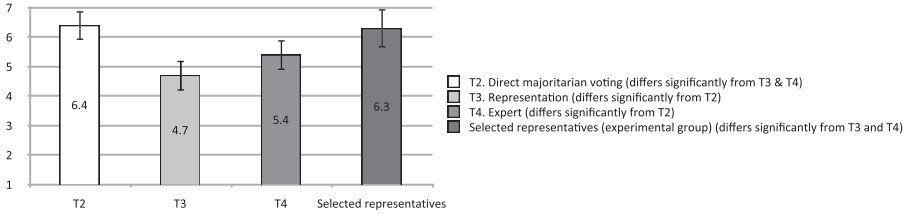

Figure 3. Legitimacy beliefs associated with exogenously imposed fair arrangements: predicted values and 95 per cent confidence intervals.
Note: Estimates are derived from model 5 presented in Table 4 in the online appendix.

research comparing representational decision making to high-status judicial decision making (Gibson et al. 2005; Gash & Murakami 2009). Since the decision in our study is taken by an administrator – the teacher – with less formal status than judges and court institutions, it suggest a relative advantage of experts over representatives in terms of perceived neutrality and, perhaps, knowledge and judgment.

To check the robustness of our findings, we turn to the series of complementary vignette experiments. Figure 4 present results from four experiments that deal with a regulatory policy proposal (to ban the use of religious symbols in schools) in both a high school setting and a large-scale democracy setting, and that involve samples of high school students, university students, teachers, and adult citizens. In all these experiments, participants gave significantly higher procedural fairness assessments when the decision was taken by direct majoritarian voting rather than by representation or expert decision making.

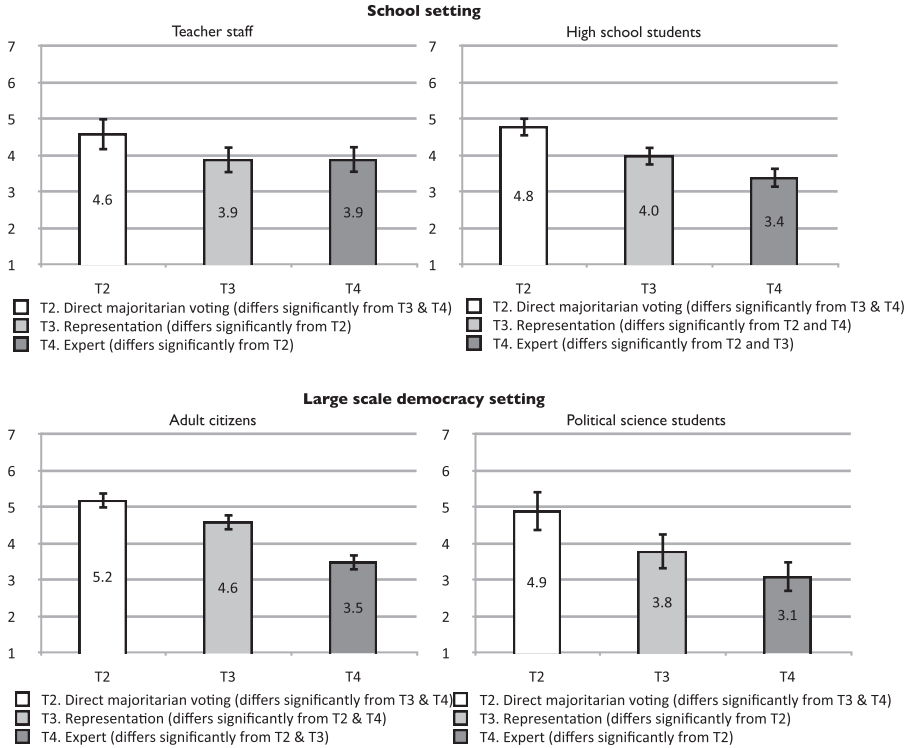

Figure 4. Effects of forms of decision making on legitimacy beliefs: results from vignette experiments.

Note: Estimates are derived from models 6–9 presented in Table 5 in the online appendix.

With regard to representation versus expert decision making, results are mixed. In the school setting, and among political science students, the difference is statistically insignificant, whereas representation generates somewhat stronger legitimacy beliefs among adult participants in the setting of a large-scale democracy. We will return to the issue of external validity in the concluding section, but overall findings from our studies strengthen beliefs that the legitimising capacity of representational decision making is relatively weak.

H3*: Consensual decision making and* H4*: Lottery-based representation*

Regarding the appeal of consensual decision making over majoritarian decision making, we turn back to the main experiment and compare legitimacy beliefs associated with T2 (direct voting) and T5 (consensual decision making). Looking at the observed means reported in Figure 1, it can be noted that both secret ballot voting and consensual decision making are successful relative to other forms of decision making, but that contrary to the prediction of *H3* the former is the most powerful generator of legitimacy beliefs (6.5 versus 6.1). A systematic comparison within a multilevel analytical framework confirms this impression: direct majoritarian voting generates somewhat higher levels of procedural fairness assessments (predicted values 6.4 versus 6.1, $p = 0.38$, two-tailed). Thus, results do not indicate that deciding in a consensual way adds more legitimacy to the process than taking a direct vote by secret ballots (for detailed information, see Table 4 in the online appendix).

Addressing, then, the value of lottery-based representation, we compare legitimacy beliefs associated with T3 (election-based representation) and T6 (lottery-based representation). Both treatments show relatively weak legitimising capacity, but lottery-based representation is associated with a somewhat higher level of procedural fairness assessments than election-based representation (5.1 versus 4.7; $p = 0.11$, one-tailed).

Notably, on this point our vignette experiment tells a partly different story. In one further study on the ban of religious symbols in schools we contrasted lottery-based representation with election-based representation resulting from both a high turnout and a low turnout election. Results showed that both types of election-based representation generated clearly stronger legitimacy beliefs than lottery-based representation (Gilljam et al. 2010). While the need for further investigation is obvious, existing empirical evidence does not suggest that lottery-based representation is particularly appealing to citizens.

H5*: (Un)Fair implementation of decision-making arrangements*

Figure 5 compares legitimacy beliefs associated with fair and unfair implementation of direct voting (T2 and T7). To assess the relative effectiveness of implementation fairness (*H5*) and personal involvement (*H2*), the figure also presents corresponding information for the fair implementation of representation (T3) and expert decision making (T4).

In accordance with expectations, a fair implementation generates higher legitimacy beliefs than an unfair implementation of the same form of decision making (6.4 versus 5.8; p = 0.05, one-tailed). However, while *H5* is supported, it should be noted that the most important factor seems to be personal involvement. This is evident when we compare unfair direct voting with fair representation and expert decision. Contrary to the expected outcome if procedural fairness was a major concern for participants in this context, unfair direct voting generates a significantly higher level of procedural fairness assessments than decisions made in a fair way by elected representatives (5.8 versus 4.7; p = 0.00, one-tailed). The corresponding difference between majoritarian voting and fair expert decision making is also positive but statistically insignificant (5.8 versus 5.4; p = 0.14, one-tailed). To reiterate: procedurally flawed majoritarian voting generates clearly stronger legitimacy beliefs than fair representation.

As always in experimental research, the fairness manipulation might fail to capture the true effect of the explanatory factor. For instance, although the third party was biased, our manipulation allowed participants to voice their opinions prior to the decision. Because the opportunity to voice opinions is a crucial aspect of decision-making procedures, this might override some of the negatives associated with a biased third-party mediator. However, there is also a more substantive interpretation to be made. In collective decisions like the present one, group-serving tendencies may

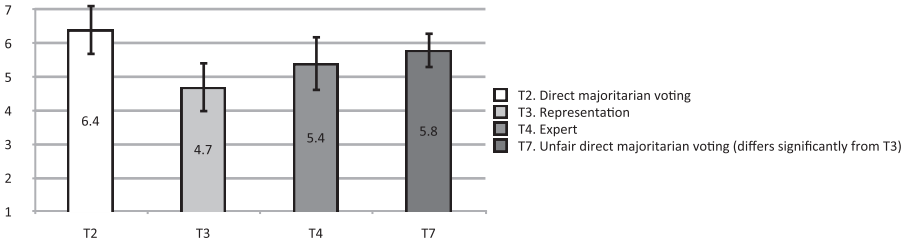

Figure 5. Legitimacy beliefs associated with fair and unfair arrangements: predicted values and 95 per cent confidence intervals.
Note: Estimates are derived from model 10 presented in Table 6 in the online appendix.

undermine procedural fairness effects (Leung et al. 2007). Moreover, most experimental studies on actual procedural fairness effects are laboratory-based and consequently much more artificial than the setting of our study. It might be that procedural violations appear less alarming in complex decision-making situations.

While implications for procedural fairness theory are unclear, the findings do provide further support for the idea of personal involvement. Regardless of whether the arrangement was fairly or unfairly implemented, participants appreciated the opportunity to be personally involved in the making of the decision.

Conclusion

This article reports findings from an exploratory randomised field experiment designed to mimic decision making in large-scale democracies. We have tested empirical support for three broad ideas for legitimate collective decision making: participatory constitution making, personal involvement in the decision-making process, and fairness in the implementation of arrangements. Perhaps surprisingly, the results clearly favour one idea over the others: throughout the analyses we find that personal involvement through direct voting increases legitimacy beliefs substantially, whereas support for the other ideas is mixed.

As we believe it is a novel observation, we draw particular attention to the finding that individuals who were selected as representatives ascribe more legitimacy to the process than their peers. Thus, the legitimising capacity of personal involvement not only applies to the many that occasionally make decisions in direct arrangements, but also to the few who are appointed as representatives.

The article is one of the first to bring experimental evidence to the discussion about legitimate decision making in large-scale democracies. To ensure external validity we manipulate institutional arrangements that resemble the ones used in real-world democracies, we study natural collectives of individuals with a shared history and future, we use money to create an incentivised environment, and we replicate our findings by means of vignette experiments. These efforts notwithstanding, there are many and obvious differences between our field experiment and nation-state democracies: we deal with students of young age, none of whom have yet had the opportunity to vote in a general election; our natural collectives are small (20–25 individuals); and we focus on a single decision on a fairly simple

question rather than on a continuous flow of decisions on complicated matters. Clearly, there is need for replication in other populations, with other issues and with repeated decisions.

Reflecting further on external validity, our manipulation of representational decision making differs the most from large-scale conditions. Despite our efforts to make a case for representational decision making (we refer to the appendix for details), real-world representatives are clearly more experienced, socially exclusive, physically distant and engaged than the ones who were elected in our study. It might well be that the benefits of representative democracy are more obvious to affected individuals in nation-state settings and when really important questions are dealt with. However, while the need for replication in other settings is obvious, we emphasise that very little research addresses head on the relative effectiveness of direct voting, election-based representation and expert decision making, and that our finding about the limits of representational decision making falls in line with research on judicial legitimacy (Gibson et al. 2005; Gash & Murakami 2009).

Returning finally to our main substantive finding, it deserves repeating that it favours the claims of participatory democrats (Pateman 1970; Barber 1984). The results suggest that direct democratic voting has a special legitimising appeal among citizens. This finding has also been confirmed in comparative survey research by Bowler et al. (2007), who show that individuals in a large number of countries are positive toward direct democracy and decision making in referendums.

It should be noted, moreover, that the finding contradicts one of the fundamentals of stealth democracy theory (Hibbing & Theiss-Morse 2002). In accordance with the theory we register widespread scepticism towards political representation, but in contrast to the theory we find that individuals, when provided the opportunity to take an active part in decision making, value personal involvement. Our study thus supports recent efforts by Neblo et al. (2010) to portray citizens as more easily engaged than the rather passive bystanders described by stealth democracy theory.

National-level democracies are complex systems of government. Well-functioning institutional arrangements must strike a balance between different democratic and instrumental values (Bowler & Donovan 1998; Smith & Tolbert 2004; for a survey of the arguments against direct democracy, see Gilljam et al. 1998). However, while acknowledging the complexity of national-level democratic government, the findings of this article add to the growing literature that ascribes to direct majoritarian voting a special quality in generating legitimacy. One policy implication that follows is that democratic governments in search of legitimacy for difficult decisions are well advised to involve citizens in the process.

Acknowledgements

A previous version of this article was presented at the annual meeting of the Midwest Political Science Association, Chicago, 2010. We would like to thank Jenny de Fine Licht for excellent research assistance and Rune Romhed for making the study possible. We have benefited from the constructive advice of Carl Dahlström, Donald P. Green, Andrej Kokkonen, Victor Lapuente, Elizabeth J. Miller, G. Bingham Powell, Anders Sundell and four anonymous reviewers. Research for this article was supported by a grant from the Swedish Research Council (SRC-721-2004-3445).

Notes

1. For a rare dissenting voice, see Morrell (1999).
2. Since each participant was assigned the equivalent of US$13 or €11, the precise sum of money varied slightly across classes. No participant objected to the designated recipient.
3. To compute a formal level of statistical power, we followed the procedure for two-level balanced cluster design recommended by Konstantopoulos (2009). We used findings from our vignette experiments to set an expected value of the effect size parameter δ to 0.80; a survey of intraclass correlations for achievement data in American schools reported in Konstantopoulos (2009) to set expected intraclass correlation $\tilde{\rho}$ to 0.20; and our knowledge about conditions in the selected school setting to set the expected number of individuals in each cluster to 25. Using these assumptions, the statistical power of the experiment is 0.58, two-tailed, at the 0.10 level.
4. In another experiment, we manipulate the presence of pre-decision deliberation (Persson et al. forthcoming).
5. Upon departure, the experimenters informed participants that other classes would be taking part and asked for their discretion about details of their experience. As all classes appeared genuinely curious during the introduction phase, we have no indication that participants did not adhere to this call for discretion.
6. This treatment is inspired by the way some European governments allegedly have handled unwelcome outcomes of referenda regarding further European integration.
7. Except for the study with high school students that only includes one of the indicators of procedural fairness assessments: 'How fair do you think matters were when the decision was taken?'
8. This and subsequent models were estimated with the xtmixed command in STATA11. We treat the dependent variable linearly, although it could also be treated as an ordinal variable. However, results from ordered logit show that both estimation methods yield substantially the same results.
9. Political interest, generalised horizontal trust, vertical trust (in teachers and principal) and particular horizontal trust (in classmates) are all measured on seven-point scales. When used in the regression models, all control variables are rescaled to theoretically vary between '0' and '1' and they were subsequently mean-centred to facilitate interpretation of the dummy variables for the treatments. Although we include a fairly large

number of controls, models do not suffer from multicollinearity. None of the independent variables have a VIF above the critical value 10.

10. For detailed statistics, see Table 3 in the online appendix.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table 1. Empty variance components models, effects of legitimacy beliefs. Maximum Likelihood Estimation

Table 2. Effects of Participatory Constitution-making on Legitimacy. Maximum Likelihood Estimation

Table 3. Effects of Participatory Constitution-Making on Legitimacy Beliefs. A Comparison Between Endogenously (experimental group) and Exogenously (control group) Arrangements for Decision-Making. Maximum Likelihood Estimation

Table 4. Effects of Fair Forms of Decision-Making on Legitimacy Beliefs. Maximum Likelihood Estimation

Table 5. Effects of Forms of Decision-Making on Legitimacy Beliefs. Results from vignette experiments. Maximum Likelihood Estimation

Table 6. Effects of (Un)fair Forms of Decision-Making on Legitimacy Beliefs. Maximum Likelihood Estimation

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Barber, B.R. (1984). *Strong democracy*. Berkeley, CA: University of California Press.

Benomar, J. (2004). Constitution-making after conflict: Lessons for Iraq. *Journal of Democracy* 15: 81–95.

Bowler, S. & Donovan, T. (1998). *Demanding choices: Opinion, voting and direct democracy*. Ann Arbor, MI: University of Michigan Press.

Bowler, S., Donovan, T. & Karp, J. (2007). Enraged or engaged? Preferences for direct citizen participation in affluent democracies. *Political Research Quarterly* 60: 351–377.

Chambers, S. (2003). Deliberative democratic theory. *Annual Review of Political Science* 6: 307–326.

Christiano, D. (1996). *The rule of the many*. Boulder, CO: Westview Press.

Cohen, J. (1997). Deliberation and democratic legitimacy. In J. Bohman & W. Rehg (eds), *Deliberative democracy*. Cambridge, MA: MIT Press.

Druckman, J.N. et al. (eds) (2011a). *Cambridge handbook of experimental political science*. New York: Cambridge University Press.

Druckman, J.N. et al. (eds) (2011b). Experimentation in political science. In *Cambridge handbook of experimental political science*. New York: Cambridge University Press.

Duflo, E., Glennerster, R. & Kremer, M. (2008). Using randomization in development economics research: A toolkit. In T.P. Schultz & J. Strauss (eds), *Handbook of development economics*. Amsterdam: North-Holland.

Duus-Otterström, G. (2011). Freedom of will and the value of choice. *Social Theory & Practice* 37: 256–284.

Dworkin, R. (2000). *Sovereign virtue*. Cambridge, MA: Harvard University Press.

Esaiasson, P. (2010). Will citizens take no for an answer? What government officials can do to enhance decision acceptance. *European Political Science Review* 2: 351–371.

Estlund, D. (2008). *Democratic authority*. Princeton, NJ: Princeton University Press.

Fehrenbacher, D.E. (1989). *Constitutions and constitutionalism in the slaveholding South*. Athens, GA: University of Georgia Press.

Gangl, A. (2003). Procedural justice theory and evaluations of the lawmaking process. *Political Behavior* 25: 119–149.

Gash, A. & Murakami, M.H. (2009). Courts, Legislatures and Ballot Initiatives: How Policy Venue Affects Public Acceptance. Paper presented at the APSA meeting, Toronto.

Gerber, A.S. & Green, D.P. (2009). Field experiments and natural experiments. In R. Goodin (ed.) *Oxford handbook of political science*. New York: Oxford University Press.

Gerber, A.S., Huber, G.A. & Washington, E. (2010). Party affiliation, partisanship and political beliefs: A field experiment. *American Political Science Review* 104 (4): 720–744.

Gibson, J. & Duch, R. (1991). Elitist theory and political tolerance in Western Europe. *Political Behavior* 13: 191–212.

Gibson, J.L., Caldeira, G.A. & Spence, L.K. (2005). Why do people accept public policies they oppose? Testing legitimacy theory with a survey-based experiment. *Political Research Quarterly* 58: 187–201.

Gilljam, M., Esaiasson, P. & Lindholm, T. (2010). The voice of the pupils: An experimental comparison of decisions made by elected pupil councils, pupils in referenda and teaching staff. *Educational Assessment, Evaluation & Accountability* 22: 73–88.

Gilljam, M., Pesonen, P. & Listhaug, O. (1998). The referendum in representatives democracies. In A. Todal Jenssen Pesonen, P. Pesonen & M. Gilljam (eds), *To join or not to join: Three Nordic referendums on membership in the European Union*. Oslo: Scandinavian University Press.

Grofman, B.N. & Reynolds, A. (2001). Electoral systems and the art of constitutional engineering: An inventory of main findings. In R. Mudambi, P. Navarra & G. Sobbrio (eds), *Rules and reason: Perspectives on constitutional political economy*. Cambridge: Cambridge University Press.

Hedges, I.V. & Hedberg, E. (2007). Intraclass correlation values for planning group randomized trials in education. *Educational Evaluation and Policy Analysis* 29: 60–87.

Helwig, C.C. et al. (2007). Mainland Chinese and Canadian adolescents' judgments and reasoning about the fairness of democratic and other forms of government. *Cognitive Development* 22: 96–109.

Hibbing, J.R. & Theiss-Morse, E. (2002). *Stealth democracy*. Cambridge: Cambridge University Press.

Kinder, D.R. (2011). Campbell's ghost. In J.N. Druckman et al. (eds), *Cambridge handbook of experimental political science*. New York: Cambridge University Press.

Konstantopoulos, S. (2009). Using power tables to compute statistical power in multilevel experimental designs. *Practical Assessment Research & Evaluation* 14(10). Available online at: http://pareonline.net/getvn.asp?v=14&n=10

Leung, L., Tong, L. & Lind, E.A. (2007). Realpolitik versus fair process: Moderating effects of group identification on acceptance of political decisions. *Journal of Personality & Social Psychology* 92: 476–489.

Lind, E.A. & Tyler, T.R. (1988). *The social psychology of procedural justice*. New York: Plenum.

Lupia, A. & Matsusaka, J. (2004). Direct democracy: New approaches to old questions. *Annual Review of Political Science* 7: 463–482.

MacCoun, R. (2005). Voice, control and belonging: The double-edged sword of procedural fairness. *Annual Review of Law and Social Science* 1: 171–201.

Manin, B. (1997). *The principles of representative government*. Cambridge: Cambridge University Press.

Moehler, D.C. (2008). *Distrusting democrats: Outcomes of participatory constitution-making*. Ann Arbor, MI: University of Michigan Press.

Morrell, M.E. (1999). Citizen's evaluations of participatory democratic procedures: Normative theory meets empirical science. *Political Research Quarterly* 52: 293–322.

Morton, R.B. & Williams, K.C. (2010). *Experimental political science and the study of causality: From nature to the lab*. New York: Cambridge University Press.

Neblo, M.A. et al. (2010). Who wants to deliberate – and why? *American Political Science Review* 104: 566–583.

Norris, P. (2004). *Electoral engineering: Voting rules and political behavior*. Cambridge: Cambridge University Press.

Olken, B.A. (2010). Direct democracy and local public goods: Evidence from a field experiment in Indonesia. *American Political Science Review* 104: 243–267.

Palfrey, T.R. (2009). Laboratory experiment in political economy. *Annual Review of Political Science* 12: 379–388.

Pateman, C. (1970). *Participation and democratic theory*. Cambridge: Cambridge University Press.

Persson, M. & Oscarsson, H. (2010). Did the egalitarian reforms of the Swedish educational system equalise levels of democratic citizenship? *Scandinavian Political Studies* 33(2): 135–163.

Persson, M., Esaiasson, P. & Gilljam, M. (Forthcoming). *The effects of direct voting and deliberation on legitimacy beliefs*. Manuscript under review.

Raudenbush, S. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods* 2: 173–185.

Rohrschneider, R. (1996). Institutional learning versus value diffusion: The evolution of democratic values among parliamentarians in Eastern and Western Germany. *Journal of Politics* 58: 422–446.

Rose, R., Mishler, W. & Haerpfer, C. (1998). *Democracy and its alternatives: Understanding post-communist societies*. Baltimore, MD: Johns Hopkins University Press.

Scanlon, T.M. (1998). *What we owe to each other*. Cambridge, MA: Belknap Press.

Skitka, L.J., Winquist, J. & Hutchinson, S. (2003). Are outcome fairness and outcome favorability distinguishable psychological constructs? A meta-analytic review. *Social Justice Research* 16: 309–341.

Smith, D. & Tolbert, C. (2004). *Educated by initiative: The effects of direct democracy on citizens and political organizations in the American states*. Ann Arbor, MI: University of Michigan Press.

Snijders, T.A.B. & Bosker, R.J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Stone, P. (2009). Sortition and Democratic Legitimacy. Paper presented at the APSA meeting, Toronto.

Sullivan, J. et al. (1993). Why politicians are more tolerant: Selective recruitment and socialization among political elites in Britain, Israel, New Zealand and the United States. *British Journal of Political Science* 23: 51–76.

Sutter, M., Haigner, S. & Kocher, M. (2010). Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Review of Economic Studies* 77: 1540–1566.

Thomas, M.A. (2010). What do the worldwide governance indicators measure? *European Journal of Development Research* 22: 31–54.

Thompson, D.F. (2008). Deliberative democratic theory and empirical political science. *Annual Review of Political Science* 11: 497–520.

Tyler T.R. (2000). Social justice: Outcome and procedure. *International Journal of Psychology* 35: 117–125.

Tyler, T.R. (2001). A psychological perspective on the legitimacy of authorities and institutions. In J.T. Jost & B. Major (eds), *The psychology of legitimacy: Emerging perspectives on ideology, justice and intergroup relations*. Cambridge: Cambridge University Press

Tyler, T.R. (2006). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology* 57: 375–400.

Tyler, T.R. et al. (1997). *Social justice in a diverse society*. Boulder, CO: Westview Press.

Van Prooijen, J.W. et al. (2006). How do people react to negative procedures? On the moderating role of authority's biased attitudes. *Journal of Experimental Social Psychology* 42: 632–645.

Address for correspondence: Peter Esaiasson, Department of Political Science, University of Gothenburg, PO Box 711, SE-405 30 Gothenburg, Sweden. E-mail: peter.esaiasson@pol.gu.se