

Democratic Auditing of the European Union, Past and Future.

Christopher Lord, ARENA, The Centre for European Studies, The University of Oslo.

1. Introduction.

Much of the controversy about democracy and the EU in recent years has revolved around an exchange in which Andreas Føllesdal and Simon Hix (2006) disputed attempts by Andrew Moravcsik (2002) and Giandomenico Majone (2005) to question whether the EU really suffers from a democratic deficit. Where, however, the protagonists in this debate end up depends a good deal on the particular definitions of democracy from which they start out. Moravcsik tells us that ‘the classic justification for democracy is to check and channel arbitrary power’ (2002: 6). Small wonder, then, he feels that multiple checks on Union power are reason to question how far it is in democratic deficit. In contrast, Føllesdal and Hix affirm Schattsneider’s view that democracy ‘is a competitive political system in which competing leaders and organisations define the alternatives of public policy in such a way the the public can participate in the decision-process’ (2006: 547). Small wonder then that they infer a democratic deficit from the absence of elections contested on EU issues.

Before, however, I sound unduly critical, it is important to note that the problem of how to specify definitions and standards that have much chance of satisfying everyone is common enough to attempts to assess the democratic quality of political systems. Indeed, it is still, arguably, the core problem forty years on from what is widely acknowledged to be the first attempt – by Robert Dahl – to propose indicators of democracy (1971). I say this even though the indicators developed by David Beetham and his collaborators (Democratic Audit) are to my mind remarkably full, precise and sensitive to different concepts of democracy.

Yet, defining indicators can only be the first step towards resolving the problem of standards. At least two further questions also need to be answered. First, how should different indicators relate to another? Is it really satisfactory just to aggregate scores given to each indicator without attention to how different indicators or combination of indicators should be prioritised, treated as necessary or sufficient, regarded as pass/fail conditions, traded off against one another and so on? Second, what should count as satisfactory levels of performance? Should the ideal, the feasible or the comparative be used as benchmarks? And who, in any case, should answer questions such as these? Experts? Those who live under a political system? Or must we include some element of universality, so that we can put some limits on which political systems can count as democracies, and make some comparisons across systems?

It is on the problem of setting and applying standards in democracy assessment that I want to focus in this paper, since it is that problem that makes the European Union so interesting to any one concerned with democratic auditing. No one really knows what kind of a polity the EU may become, or even what kind of a polity it is right now, as every undergraduate essay which repeats Jacques Delors’ joke that the Union is an unidentified political object will tell you. Given that the object of assessment is so mysterious, it is perhaps unsurprising that there is also little agreement on what standards of democracy should apply to the European Union, and, therefore, what standards should in an academic assessment. Indeed, the notion that the EU should be democratic at all is deeply contested. Perhaps the nearest to a consensus view is that the Union should be democratically controlled. Yet many, of course, continue to believe that could be done through the democracies of its member states and that it does not, therefore, follow that the Union should be a democracy itself. Even assuming that the Union should itself be a democracy there are many models on offer, and behind each of those models is wide disagreement about about the value, justification and nature of democracy, as well, of course, what is feasible at the Union level. Moreover, this discussion never really frees itself from uncertainty about the nature of the beast, and the core question of whether an as yet imperfectly classified non-state political system that exercises political power beyond the state be assessed by democratic standards and indicators that have largely been developed through experience with assessing democracy within states.

in turn, be used in any academic assessment only is the. However, I want to argue here that there is a role for experimentation, as well as deliberation, in resolving 'problems of standards'. I make this argument in three stages. First, I identify general problems and possibilities of democracy assessment. Second, I illustrate some of those problems and possibilities from my own experience with democratic auditing of the European Union. Third, I sketch – and I emphasise that I do no more than that – my own suggestion for a more 'experimental-deliberative' approach to democracy assessment.

2. The challenges of democracy assessment.

Numerous attempts have been made to design methods of assessing democracy (Dahl 1971; Bollen 1980; Hadenius 1992; Beetham et al 2002; Economist Intelligence Unit 2007; Vanhanen 2008; Bühlmann *et al* 2011). To the extent, however, that just two of those measures – those developed by Freedom House (2012) and POLITY IV – are used to generate further conclusions in several other fields of study from right across the social sciences, their weaknesses deserve special attention. At least the following problems have been identified.

One problem is one of *consistency*. Freedom House assesses democracy using indicators of civil liberties and of political rights. In contrast POLITY IV uses indicators of participation, executive recruitment and constraints on the executive. Defenders of Freedom House and POLITY IV argue that it does not matter that they differ in their indicators and their methods, since they correlate in their scores. Yet, Casper and Tufis argue that it does not follow from this that Freedom House and POLITY IV measures can be used interchangeably (2003: 197). To the contrary, different explanations for levels of democratisation are statistically significant, or have positive or negative impacts, depending on which of the surveys are used. Other commentators question whether the Freedom House and POLITY IV scores correlate at all once the easy cases – those political systems with near perfect scores – are removed (Coppedge, Gerring *et al* 2011: 151). Thus far from being mutually confirming, it is questionable whether the two measures of democracy that are most used in social science research can both be right.

A second difficulty is one of *arbitrariness*. The use of expert judgements in Freedom House and POLITY IV surveys to score democratic standards has been criticised as reinforcing biases in the indicators themselves. There is no requirement that the judgements be deliberated or justified, and it is not even clear whether they are consistently and transparently checked by inter-coder reliability tests (Coppedge and Gerring 2011: 151). The arbitrariness of the Freedom House and POLITY IV measures is, arguably, then compounded by their fondness for aggregating the scores they give to individual aspects of democratic performance into overall country scores. But does it make sense to blend indicators into overall country scores or are some indicators so basic that a country that fails to satisfy them can hardly count as democratic at all? Can any system of aggregation really identify complex interactions between the different components of democracy that mean shortcomings are sometimes compensated, sometimes compounded, by other attributes of political systems? (Lord 2004; Coppedge and Gerring 2011: 250).

A third criticism is that Freedom House and POLITY IV surveys are *blind to reasonable disagreement* about the nature of democracy itself (Abromeit 2004). Even those who are fully reasonable in the sense they are committed to agreeing understandings of democracy with others, soon discover deep differences in human conceptions of freedom and equality on which democracy depends. Thus people with equally sincere and well-thought out understandings of democracy would be likely to score Freedom House and POLITY IV indicators differently, assuming that they are even prepared to acknowledge that those indicators test what needs to be tested in any assessment of democratic standards (Coppedge, Gerring *et al* 2011; Bühlmann *et al* 2012).

However, a satisfactory method of assessing how democratic are political systems cannot just be identified negatively, that is, by reviewing existing approaches in the hope of identifying where they go wrong and what might be needed to repair them. Ensuring that methods of assessing democracy are

'fit for purpose' plainly also requires a more positive effort to identify what those purposes might be. One obvious possibility is that democracy assessment may be needed to meet what the political scientist Robert Putnam once described as 'the ancient obligation of our craft' to study conditions of 'good government' through 'rigorous appraisals of institutional performance' (1993: 63). Assuming that democracy is a principal component of 'good government' a key question is whether there is anything about democracy that requires democracy to be able to assess democracy.

On the one hand, democracies may need to be able to assess themselves. Of crucial importance here is the argument of many democratic theorists that self-rule can only be self-rule where publics can make informed choices between forms of democracy. Thus James Bohman has argued that 'democracy is that set of institutions by which individuals are empowered as free and equal citizens to form and change the terms of their collective life together, *including democracy itself*' (2007:2). Richard Bellamy likewise argues that 'for democracy to mean "people rule", the *demos* should be free to redefine the nature of their democracy whenever they want and not be tied to any fixed definition' (2007: 90). Of course, there are some who might question just how far publics ever get to choose whole social and political systems. However, I take it that even those who believe that *bricolage* or incremental adjustment is as much as can be expected, would at least agree that publics should, as Philip Pettit puts it, be able to 'invigilate' democratic systems (2008: 53). If, however, publics are to be able to invigilate their political systems it would help if academic research could do everything that might reasonably be expected from a rigorous method of assessing democracy, such as a) justify standards and make them explicit; b) correctly attribute causes, effects and constraints so that problems can be diagnosed; and c) distinguish problems that arise in the delivery of other attributes of democracy from those that are 'pure loss'.

On the other hand, democracies may need to be able to assess the level and quality of democracies in other countries. Democratic systems have done just that in the belief that decisions on development aid, membership of international organisations and international stability should be linked to attributes of democratic rule. However, strictly speaking, it is not necessary to believe in 'democratic conditionality' or 'democratic peace' to hold that democracies need a reliable means of distinguishing democracies from non-democracies. Democratic governments cannot claim that political power is justified where it is controlled by individual citizens as equals whilst seeming to question the right of other states and societies to decide in just that way (Rawls 1999). Thus they have certain duties – of respect and of constraint - to other democracies that follow from a need to avoid contradicting the very principles that legitimate their own political systems. Since, moreover, the aim here is presumably to be seen to act towards other democracies in ways that maintain the overall credibility and legitimacy of democratic rule, democratic governments are presumably constrained in how they can categorise other systems as democratic or non-democratic on no other basis than their own caprice and convenience, and without regard for the credibility of their assessments. Access to reliable and defensible methods of democracy assessment would be one way of ensuring that credibility.

Now that we have said something about its problems and purposes, it is useful to distinguish two kinds of challenge that arise during the course democracy assessment: those of theoretical relevance and of empirical observability. Perhaps the main challenge of theoretical relevance is one of ensuring that standards of assessment match reasons for valuing democracy in the first place. This seemingly banal – even self-evident requirement - is actually fraught with difficulty. Let me give a huge example. It is often said that freedom and equality jointly form the core values of democracy. Yet freedom and equality are valued by different people for different reasons. Republicans famously differ from liberals in how far they believe that freedom is merely a 'negative' liberty from non-interference as opposed to a 'positive' liberty to act to ensure freedom from all kinds of arbitrariness. Social democrats famously differ from liberals in how far they believe that political equality also requires social and economic equality. Those who belong to these different schools of thought would be unlikely to be satisfied with the same indicators of democracy, nor would they be likely to score the same indicators in the same way.

As it happens, a number of recent contributions have pleaded for a more pluralistic approach to democracy assessment which acknowledges that there may be multiple ways of valuing democracy and multiple ways of realising democratic values. Bühlmann et al (2012) have developed indicators of minimalist, medium and maximalist democracy, designed to test whether political systems meet bare requirements for electoral democracy (minimum), achieve 'high quality participation and representation' (medium) and 'social justice' (maximum). Whilst, however, Bühlmann et al associate their levels with different values, their rather Millian attempt to organise those values into 'higher' and 'lower' pleasures seems to cling on to the assumption that democratic values are ultimately uni-scalar, rather than sometimes non-commensurable. Michael Coppedge and John Gerring go further towards suggesting that we should simply develop indicators of different models of democracy. However, unless I am mistaken, their plea for an historical approach implies that any one political system should be assessed against just the one model of democracy that is closest to that it has adopted historically. What, though, of political systems that cannot yet be said to have adopted any one model of democracy definitively? Might there even be circumstances in which it might make sense to assess them against more than one model? I will come back to this both in the discussion of the EU and in that of the benefits of introducing more experimental methods to democracy assessment.

Yet, greater pluralism in standards cannot be the only solution either. Whilst it is important that standards should accommodate different reasons for valuing democracy, it plainly cannot be the case that democracy is just what anyone says it is. In my book on the Democratic Audit of the EU, I suggested that one solution might be to construct indicators on the assumption that democracy is a 'boundedly contested concept' as opposed to an 'essentially contested concept' (2004: 12-3). That is to say, varieties of democracy can only vary within the bounds of what is needed to provide *some* means of meeting core requirements, such as political equality and public control. I take it that Bohman has something similar in mind so far as he argues that each people should be free to choose its own form of democracy, and, yet, there is, none the less, a 'democratic minimum' that everyone has to observe (2007). Some way of framing democratic indicators so that they differentiate between 'compulsory cores' and 'optional' ways of meeting those core requirements may also be the best way to allow scope to evaluate each democracy against its own freely chosen standards whilst retaining some criteria that allow judgements and comparisons are to be made across systems

What, though, of the challenge of framing tests that are empirically verifiable? In part the challenge is a basic one common to all forms of enquiry. Methods of assessing democracy can only be as meaningful as the conditions that validate them, that is, they need to be framed so that those using them can present reasons and evidence that show that they are entitled to reach their assessments. However, once again a seemingly obvious requirement turns out to be anything but straightforward. The core problem here is that democracy assessment is not – and cannot possibly be – a purely empirical undertaking. It is necessarily also both evaluative and prescriptive. Up to a point, this actually reinforces the need for democracy assessments to get things right empirically. If, as seen, one purpose is to diagnose shortcomings and identify how things might be put right, democracy assessments will need to be able to analyse exactly which institutions or practices are causally responsible for any deficiencies. Thus democracy assessment can no more do without the state-of-the-art methods of identifying causation that have been developed by analytical political science than it can manage without state-of-the-art political philosophy to justify its standards. Indeed, this interdependence becomes acute with the core evaluative task of democracy assessment. At some point data collection and analysis has to draw to a close and a judgement has to be made on 'how well?' standards are satisfied. No amount of empirical data can on its own answer the question what justifies a decision to count one threshold rather than another as 'good enough'? At this point, democracy assessment is unavoidably a judgement rather than a measurement; or, to be precise, it is a judgement of measurements made.

3. Challenge of conducting a Democracy Assessment of the EU

What do the possibilities and problems discussed in the last section imply for the challenge of designing a democratic audit of the European Union? Here are a few selected thoughts.

a): Defining the purpose and scope of a Democratic Audit of the European Union. A clear difficulty for any democracy assessment of the EU is one of how to appraise a political system that is still, arguably, at a stage of deciding what kind of democracy it should be. It is tempting to conclude that academic appraisal should wait until citizens and their representatives have resolved this question. Yet, it is precisely at the formative stage of institution building that democratic assessments may be most needed, if one of their purposes is to help people make choices of institutional design. However, this already begins to beg the question of what should be included in any democratic audit of the European Union. As Moravcsik has remarked the Union is not a ‘stand-alone’ political system. Rather, it is constructed in a way that presupposes that its democratic qualities, if any, are jointly supplied (Lord and Harris 2006), part by its own institutions, part by those of its member states. Thus, the EU depends on the electoral administration of its member states to organise free and fair elections, on national party systems to structure choice and competition in European elections, on a critical, free and informed media in each member state to provide public debate and so on. Indeed, the view of representation to which the Union has subscribed in the Lisbon Treaty presupposes that accountability and control run through national parliaments as well as Union institutions. Thus a comprehensive democratic assessment of the European Union would probably have to include a fairly thorough-going assessment of democracy in each member state.

Indeed, I argued earlier that democratic assessments are not just needed to help with questions of institutional design. They are also needed to deal with ‘inside-outside’ problems of democratic rule. On the one hand, democracies are systems of self-rule. On the other hand, they have external effects and, arguably, external duties that, to some extent, imply that their internal democratic qualities are shared concerns. In the case of the European Union these difficulties are especially acute. Each member state takes part in what is to some extent a common and collectively binding form of rule that can significantly affect the distribution of values and individual life chances. Exit options are constrained, and existing policies hard to amend (Scharpf 2009). Yet, as just seen, the democratic quality of those collectively constraining decision-making process is, in part, the product of the democratic quality of each member state. Once again, it looks as though a comprehensive democratic assessment of the Union would have to be a 27+1 solution that would ideally link democratic audits in member states with an audit of Union institutions themselves.

b): Setting the standards for a Democratic Audit of the EU. If, however, a democracy assessment of the Union is to help everyone equally in deliberating and choosing what kind of a democracy the Union ought to be it will need to avoid presupposing the superiority of any one solution and settle instead for assessing the Union simultaneously against different models that can be defended as a fair representation of the feasible possibilities and normative preferences that are in play at any one time. Of course, this is a tall order, given the huge range of possibilities available. However, table 1 shows how core democratic standards might in, principle, be adapted so that they can test for different models. In keeping with earlier observations that the best way to reconcile the universal and the plural in democracy assessment may be to specify core tests that can be filled out in different ways, the first column of the table suggests some generic tests. The second then specifies tests for a model in which the Union is controlled through the democratic institutions of its member states, and the third tests for a model in which the Union develop a democratic political system of its own.

c): Deciding what is good enough. It is plainly possible to prefer either of the models in table 1 for different reasons. Some people might have different value preferences for different degrees of political community and, therefore, democracy at the national or European levels. Others, however, may feel that the choice between the two models depends at least somewhat on empirical questions of feasibility or necessity. It is worth labouring this point, since, I think, it confirms the question of what is a satisfactory level of democratic performance is not always self-evident.

Consider the second model in table 1. Some might dismiss it outright, perhaps on the grounds that democracy beyond the state is a category mistake, or, in other words, an absurdity. As it happens, I think this is an argument that needs to be taken very seriously indeed. It might run as follows. If a political system is to be democratic, it will have to satisfy a long and demanding set of conditions,

such as: a) freedoms of speech and association; b) free and fair elections; c) appointment of the legislature and leading executive positions by popular vote; d) a form of political competition that offers voters choices relevant to the control of the political system; e) a civil society in which all groups have equal opportunity to organise to influence the polity; f) a public sphere in which all opinions have equal access to public debate and g) a defined *demos*; or, in other words, agreement on who should have votes and voice in the making of decisions binding on all. Yet, achieving all these conditions simultaneously may be hard in any institution that operates beyond the state and is not, therefore, itself a state. The capacity of the state to concentrate power, resources and legal enforcement has been useful in all kinds of ways to democracy: in ensuring that the decisions of democratic majorities are carried out; in guaranteeing rights needed for democracy; in drawing the boundaries of defined political communities; in motivating voters and elites to participate in democratic competition for the control of an entity which manifestly affects their needs and values.

Whilst, though, some people believe that democracy beyond the state is a category mistake, others believe it may even be a categorical imperative (Bohman 2007). Even though this is argument conflicts with the last one, maybe it too needs to be taken seriously. It might run as follows. Many of the important systems on which modern life depends - not least ecological and financial systems - require international bodies that are powerful enough to control negative externalities, provide public goods, and avoid moral hazard. If however power needs to be exercised internationally, and if democratic control is a necessary condition for the rightful exercise of political power, then democracy must at least be able to reach beyond the state. It may also need to be partly organised beyond the state if that is what is needed to solve critical collective action problems whilst meeting basic obligations of justice, political equality or freedom. We, arguably, have obligations to develop any institutions that are, in turn, needed to deliver our obligations (Buchanan 2002).

Now all this, it seems to me, is compatible with a wide range of different beliefs about what is reasonable level of democratic performance to expect of the European Union. We could, for example, take a kind of Scharpf-like view that it is most unlikely that the Union will be able to reproduce the conditions needed for democracy and yet it is also likely to be needed none the less to solve collective action problems. That kind of view might justify conducting a democracy assessment of the EU with low expectations, as befits Scharpf's view that the democratic deficit is something to be managed, not solved. On the other hand, we could take the view that ideals are ideals and they remain valid measures, however, much we might want to acknowledge reasons why it is difficult to achieve them. As a third position we might even want to insist on some comparison with standards within member states as a way of assessing how much is gained or lost – and in relation to which indicators and attributes of democracy – when competence is moved between the levels. We might also want to resist making macro-judgements about the Union's overall democratic quality at the expense of more fine grained comparisons a) of its own democratic performance over time, b) of the relative democratic quality of its different institutions and practices and, as seen, as c) of how the democratic qualities of each of its member states interacts with those of the Union to produce variation in how far individuals living in different countries can exercise public control of EU policies as equals.

4) Where next? A Deliberative-Experimental Future?

Over the course of the paper I have filled out my initial claim that the 'problem of standards' includes much more than the bare question of which indicators should be used in a democracy assessment. Methods of democracy assessment also have to indicate what is a good, bad or indifferent level of performance against each indicator. Here, the main question is whether benchmarks should be ideal, feasible or comparative. On top of all that, methods of democracy assessment have to decide how (if at all) individual indicators should be combined to reach an overall evaluation of a political system. Here the main question is whether there are relationships of necessity or sufficiency that preclude a simple aggregation without regard for how low scores on some indicators could 'fail' a system outright, whilst others can be safely traded off against one another.

Coppedge and Gerring have argued that the best response to questions such as these is to ‘let the people decide’. In other words, the principal objection to the most widely used methods of democracy assessment (Freedom House and POLITYIV) is that they assume *a priori* and arbitrary answers to the questions of how standards should be defined, assessed and combined, one obvious solution might be to research what those who live under a system believe about standards. Several Democratic Audits have attempted to do just that by introducing a deliberative element to assessment processes. As David Beetham put it in his Parliamentary Affairs article:

‘There are a number of significant differences between the idea of a democratic audit and other methods of democracy assessment. Where the others involve outsiders sitting in judgement on another country or countries, a democratic audit is undertaken as a domestic project by citizens of the country being assessed, as part of an internal debate about the character of political institutions and public life. This does not mean that there is no external or comparative reference point for the criteria and standards employed – indeed, such a reference point is essential – but that the assessments being made involve local expertise and critical judgement, and form part of an internal political debate, rather than an externally defined agenda... In a democratic audit, the final responsibility for the judgements made is a domestic one’ (1999: 568).

Using a variety of assessors who represent a range of critical voices, and getting them to justify their assessments, all within a process of public debate, are all, arguably, improvements on assessment methods used by Freedom House and POLITY IV. A further, as yet untried, possibility would be to use a citizens’ jury, or even a process resembling a deliberative opinion poll (Fishkin 2009) to deliberate standards.

However, it seems to me that there is a role for experimentation, as well as deliberation, in resolving complex questions about the use of standards in democracy assessments. Before explaining exactly what I mean by this, I first set out some theoretical reasons for expecting experimentation to be central to any process standard setting. Almost a century ago, John Dewey, asked ‘Shall our political philosophy be experimental, or shall it be *a priori* and absolutist (Boydston 1979: 415)? Since then the ‘pragmatic turn’ in philosophy has investigated just how far concepts and standards are formed through the commitments that actors enter in their everyday practice and use of language (Searle 2010: 87). This is hardly the place to review the state-of-the-art in the philosophy of language and of meaning. It is enough for our purposes to note how that literature implies that any process of standard setting is doubly experimental.

First, a shared standard is self-evidently intersubjective. In Rawlsian terms, it can only be established through arguments capable of convincing others (Rawls 1993 following Sibley 1953: 553). In Habermasian terms it can only be established by through action aimed at mutual understanding (Habermas 1984: 285. See also Eriksen and Weingard 2003: 24) . Assuming no *a priori* standards, we actors agree standards experimentally by putting forward suggestions that are then accepted, rejected or ignored by others.

Yet, even where we succeed in convincing one another of a set of standards, we also need to be able to ensure that they bring about the values we want them to achieve. This is the second way in which standard setting is experimental. Often our standards themselves assume that ‘ought implies can’, and that they hold only as long as we can make them work in daily experimentation with practical problem solving, otherwise not (Putnam 2004).

None of this means that our standards are only ever justified by their consequences as opposed to the inherent value we put on them. But it certainly does mean that the notion that we can ever make claims of fact without making claims of value or *vice versa* is deeply and utterly incoherent. If we are to convince others of any kind of claim – empirical ones included- we need, as Robert Brandom puts it, already be able to do everything we need to do to make normative statements, since we need to be able to

justify and commit ourselves (Brandom 2008: 115). If on the other hand we are to make any claim of value we must at the least be able to indicate the ‘states of the world’ under which it holds.

All this probably explains why assessments of fact and value are so deeply intertwined in democracy assessments. However, rather than follow up that thought just now, I want to focus on the point that standard setting is an experimental exercise. If we can expect to infer what actors believe about democratic standards from the ways in which they appear to use them in their practical problem solving, and from the positions to which they commit themselves in making, accepting, rejecting or ignoring claims (See Saward (2010) for the use of a similar approach in the study of representation), then we might expect to be able to do much of that by just observing the normal operation of the political process, on the assumption that it is itself a great big live continuous experiment that unfolds in front of us. Yet a very different kind of experiment, namely, controlled experiments – in which beliefs about standards are analysed under laboratory conditions – can be expected to yield further understanding that is unlikely to be gained by just observing the political process. Although this might sound somewhat fanciful, two examples illustrate the potential of such a strategy.

First, philosophers have recently used laboratory experiments to test what real people think about a number of questions that are at least as abstract as beliefs about democratic standards, including questions of justice (Frohlich & Oppenheimer 1992), ethics (Appiah 2008), consciousness (Knobe and Prinz 2008) and free will (Nichols 2011). In the spirit of pragmatic philosophy, these experiments infer norms and meanings, and how they change under different conditions, by analysing how participants respond to variations in real world problems.

My second example comes from the work of Eleanor Ostrom, who devoted a large part of her research to testing the conditions under which people are prepared to accept norms and institutions in order to solve collective action problems. As she pointed out, a great limitation of just observing the political process is that the latter may only get to consider a narrow range of the different ways in which institutions could be designed (Ostrom 1998). In an ideal world, we would be able to vary each of the many different ways in which institutions could be designed in a controlled manner – changing just one attribute at a time – to gain a complete picture of how, for example, people’s beliefs about democratic standards vary over the range of institutional choices available to them. Controlled experiments can do that. Mere observation of the political process cannot (See also Druckman et al 2006).

Some recent examples provide some fascinating examples of how controlled experiments can be used in political science. Of course, there is only space here to present an *amuse-gueule*. Imagine, for example, an experiment designed to test different factors that affect the likelihood that participants would:

- a): vote in European elections on European issues or national issues, where the one course excludes the other;
- b): accept collectively binding forms of joint decision-making at the Union level, even where there is a risk that this will over-ride national democratic institutions;
- c): support national vetoes, even where there is a risk that collective action will fail at the European level; and so on.

Each of these is a trade-off question. It presupposes a choice has to be made. Moreover, each choice is also relevant to that between the two models in table 1. By deciding one way or the other between the three questions participants would ‘reveal’ their preferences between the standards represented by the models. Moreover, the experiments could be refined in further insightful ways. Following Chong and Druckman (2010), content analysis of official documents and media debates could be used to identify real world problems and framings of problems to present to participants in the experiments to discover

how responses to questions a)-c) vary. That would be the kind of thing that might help us test the pragmatic expectation that beliefs about standards vary over states of the world.

Conclusion.

Freedom House and POLITY IV remain the most widely used methods of democracy assessment in political science. Yet they have been criticised for making arbitrary assumptions about democratic standards. Thus it perhaps fortunate that Freedom House and POLITY IV do not attempt to assess EU institutions.

The Union is not the only polity in the world in which different people value democracy for different reasons. However, it is one in which those differences are compounded by differences in the value that different people put on democratic political community at the national and European levels, as well as differences in beliefs about the feasibility and necessity of attempting to democracy at the Union level. There is therefore a strong case for testing the Union simultaneously against indicators that reflect these different assumptions about the kind of democracy that the Union can or should be, especially if democratic assessment is to have a role in informing institutional choice.

Providing the Union with an assessment method that avoids criticisms levelled at Freedom House and POLITY IV would, however, require more than assessing it against plural standards. It would also require some means of ensuring that the whole range of decisions that need to be taken about standards – what should counts as a satisfactory level of performance and how should standards be combined together – should not just be based on researchers’ assumptions. As far as possible they should reflect public beliefs. Here it is useful that Democratic Auditing anticipates a need to deliberate and justify standards. However I have also suggested here that experiments could be used to infer public beliefs about standards under alternative states of the world.

Last but not least, a comprehensive democratic audit of the European Union would include assessments of democracy in its member states. All of this is an agenda for a research community, not for an individual

Table 1. Adapting indicators to different models of European Union Democracy

Based on the first and second Recon Models. (Eriksen and Fossum 2012: 22-9)

Generic Indicator	Model 1 Delegated Democracy (The	Model 2 Supranational Democracy (The
-------------------	-------------------------------------	---

	Union is democratically controlled by the political systems of its member states).	Union operates a Democratic Political System of its own)
How far, how equally and how securely do citizens enjoy rights of free speech, association and assembly?	National freedoms of speech, association and assembly are available to domestic publics in their control of powers delegated to the EU.	Union-wide guarantees of freedoms of speech, association and assembly in each Member State.
How far and how equally can citizens exercise public control through free and fair voting?	Free and fair elections to national executive and legislative offices which control delegations of power to the Union.	Free and fair elections to executive and legislative office at the Union level
How far can representatives elected by the people exercise ultimate controlling power over all public bodies on a day-to-day basis?	Effective national parliamentary scrutiny and control of powers delegated to the Union.	A European Parliament scrutinises and controls of the powers of other Union institutions.
How far do political parties structure voter choice in ways which help citizens exercise public control as equals?	National party competition allows citizens to exercise control over delegations of power to the Union	Parties structure voter choice so that elections to EU office can be used to exercise control over Union decisions
How plural and how independent is the range of social groups, organised interests and communications media that seeks to influence the polity?	Range and independence of the national civil society actors that seek to influence Union policy, and the equality of their access.	Range and independence of civil society actors organised to influence majority formation at EU level and equality of their access
How far are decisions deliberated within a public sphere that allows all matters to be justified and decided, free of inequalities in power and resources?	Each Member State is a well-formed public sphere where all points of view have equal access to national procedures for controlling delegations of power to the EU	The EU is itself a public sphere in which all views on the exercise of its powers are considered and justified in relation to one another on a basis of equality
How far and how equally do citizens enjoy civic capabilities needed for them to exercise public control over the polity?	Citizens are able to make informed and deliberated choices in selecting representatives who exercise national procedures for controlling delegations of power to the EU	Citizens are able to make informed and deliberated choices in elections to executive and legislative office at the Union level.
How far does the polity rest on a rule of law that itself encompasses no more and no less than those conditions required for citizens to author their own laws as equals?	National procedures for controlling delegations of power to the EU are covered by rule of law principles in all Member States	The European Union develops its own democratic rule of law controlled by majorities formed at the European level
How far is the polity accepted as a unit whose members can make decisions binding on all? How far can citizens exercise equal control over the design of the polity itself?	National control over delegations of power to the Union ground public acceptance of the EU polity itself. National procedures for bargaining and ratifying Treaty change allow citizens of all Member States to exercise public control over the design of the EU polity as equals	Majorities of voters and their representatives are widely accepted as having the right to make legally binding decisions in the exercise of powers assigned to the EU. Those majorities can also control the further development of the EU polity as equals in so far as changes affect powers already assigned to the Union.

References

Abromeit, H. (2004) 'Die Messbarkeit von Demokratie: zur Relevanz des Kontextes', *Politische Vierteljahrsschrift* 45 (1): 73-93.

Appiah, K. (2008) *Experiments in Ethics*, Harvard: Harvard University Press.

- Beetham, D. (1999) 'The Idea of a Democratic Audit in Comparative Perspective', *Parliamentary Affairs*,
- Beetham, D., Bracking, S., Kearton, I. and Weir, S. (2002) *International IDEA Handbook on Democracy Assessment*. The Hague, Kluwer Press.
- Bellamy, R. (2007) *Political Constitutionalism, A Republican Defence of the Consitutionality of Democracy*, Cambridge: Cambridge University Press.
- Bohman, J. (2007) *Democracy across Borders from Demos to Demoi*, Cambridge MA: MIT Press.
- Bollen, K. (1980) 'Issues in the Comparative Measurement of Political Democracy', *American Sociological Review* 45(3): 370-90.
- Boydston, J. (1979) (ed.) *The Middle Works of John Dewey*, Volume 8, South Illinois: South Illinois University Press.
- Brandt, R. (2008) *Between Saying and Doing. Towards an Analytic Pragmatism*, Oxford: Oxford University Press.
- Buchanan, A. (2002) 'Political Legitimacy and Democracy', *Ethics* 112 (4): 689-719.
- Bühlmann, M., Merkel, W., Müller, L., Giebler, H. Weßels, B., (2012) 'Demokratiebarometer: ein neues Instrument zur Messung von Demokratiequalität', *Zeitschrift für Vergleichende Politikwissenschaft*, Published on line at 6(1): 115–158. Online. Available at <<http://www.springerlink.com/content/c23n48n122116u03/>> (last accessed 19 November 2012).
- Casper, G. and Tufis, C. (2003) 'Correlation versus Interchangeability: the limited robustness of empirical findings on democracy using highly correlated data-sets', *Political Analysis* 11 (1): 1-11.
- Chong, D. and Druckman, J. (2010) 'Dynamic Public Opinion: Communication Effects over Time', *American Political Science Review*, 104 (4): 663-681.
- Coppedge, M., Alvarez, A. and Maldonado, C. (2008) 'Two Persistent Dimensions of Democracy: Contestation and Inclusiveness', *Journal of Politics* 70 (3): 335-50.
- Coppedge, M., Gerring J et al (2011) 'Conceptualising and Measuring Democracy: A New Approach', *Perspectives on Politics*, 9(2): 247-267.
- Dahl, R. (1971) *Polyarchy, Participation and Opposition*, New Haven CT, Yale University Press.
- Democratic Audit. "A Do it Yourself Guide to Democratic Auditing." from www.democraticaudit.com.
- Druckman, J., Green, D., Kuklinski, J., Lupia, A. (2006) 'The Growth and Development of Experimental Research in Political Science', *American Political Science Review*, 100 (4): 627-635.
- Druckman, J., Green, D., Kuklinski, J., Lupia, A. (2011) *Cambridge Handbook of Experimental Political Science*, Cambridge: CUP.
- Economist Intelligence Unit (2007). "The Economist Intelligence Unit's Index of Democracy." Online. Available at <http://www.economist.com/media/pdf/DEMOCRACY_INDEX_2007_v3.pdf> (last accessed 19 November 2012).

- Eriksen, E.-O. and Weingard, J. (2003) *Understanding Habermas: Communicative Action and Deliberative Democracy*, London: Continuum.
- Fishkin, J. (2009) *When the People Speak. Deliberative Democracy and Public Consultation*, Oxford: Oxford University Press.
- Føllesdal, A and Hix, S. (2006) 'Why there is a Democratic Deficit in the European Union, A Response to Majone and Moravcsik', *Journal of Common Market Studies*, 44 (3): 533-62.
- Forst, R. (2007) 'Zu einer kritischen Theorie transnationaler Gerechtigkeit' in R. Forst, *Das Recht auf Rechtfertigung: Elemente einer konstruktivistischen Theorie der Gerechtigkeit*, Frankfurt am Main, Surhrkamp.
- Freedom House (2012) www.freedomhouse.org
- Frohlich, N. and Oppenheimer, J. (1992) *Choosing Justice: An Experimental Approach to Ethical Theory*, Berkeley: University of California Press.
- Habermas, J. (1984) *The Theory of Communicative Action – Reason and Rationalization of Society*, vol. 1, Boston: Beacon Press
- Habermas, J. (1996) *Between Facts and Norms*, Cambridge: Polity
- Habermas, J. (2003) *Truth and Justification*, Cambridge: Polity
- Hadenius, A. (1992) *Democracy and Development*, Cambridge: Cambridge University Press.
- Knobe, J. and Prinz, J. (2008) 'Intuitions about Consciousness: Experimental Studies', *Phenomenology and Cognitive Science*, 7 (1): 67-83.
- Lord, C. (1998) *Democracy in the European Union*, Sheffield: Sheffield University Press/UACES.
- Lord, C. (2004) *A Democratic Audit of the European Union*, Basingstoke: Palgrave Macmillan.
- Lord, C. (2012a) 'A Democratic Achievement and not just a Democratic Deficit', in H. Zimmermann and A. Dür (eds) *Key Controversies in European Integration*, Basingstoke: Palgrave Macmillan.
- Lord, C. (2012b) 'A Democratic Audit Framework', in E.O. Eriksen and J.E. Fossum (eds) *Rethinking Democracy and the European Union*, London: Routledge.
- Lord, C. and Harris, E. (2006) *Democracy in the new Europe*, Basingstoke, Palgrave/Macmillan.
- Majone, G. (2005) *The Dilemmas of European Integration: The Ambiguities and Pitfalls of Integration by Stealth*, Oxford: Oxford University Press.
- Moravcsik, A. (2002) 'In Defence of the "Democratic Deficit": Reassessing Legitimacy in the European Union', *Journal of Common Market Studies* 40 (4): 603-24.
- Nichols, S. (2011) 'Experimental Philosophy and the Problem of Free Will', *Science* 331(6023): 1401-3.
- Ostrom, E. (1998) 'A Behavioural Approach to the Rational Choice of Collective Action', *American Political Science Review*, 92 (1):-22.
- Pettit, P. (2008) 'Three Conceptions of Democratic Control', *Constellations* 15 (1): 46-55.

- POLITY IV (2012) 'Polity IV Project. Political Regime Characteristics and Transitions 1800–2010.' Online. Available at < <http://www.systemicpeace.org/polity/polity4.htm>> (last accessed 19 November 2012).
- Puchala, D. (1972) 'Of Blind Men, Elephants and European Integration', *Journal of Common Market Studies* 10 (2): 267-84.
- Putnam, H. (2004) *The Collapse of the Fact-Value Dichotomy and other essays*, Harvard: Harvard University Press.
- Putnam, R. (1993) *Making Democracy Work. Civic Traditions in Modern Italy*, Princeton NJ: Princeton University Press.
- Rawls, J. (1993) *Political Liberalism*, New York: Columbia University Press.
- Rawls, J. (1999) *The Law of Peoples*, Cambridge Mass: Harvard University Press.
- Reif, K.-H. and Schmitt, H. (1980) 'Nine Second-Order National Elections, A Conceptual Framework for the Analysis of European Election Results', *European Journal of Political Research* 8 (1): 3-45.
- Saward, M. (2010) *The Representative Claim*, Oxford: Oxford University Press.
- Scharpf, F. (2009) 'Legitimacy in the Multilevel European Polity', *European Political Science Review* 1 (2): 173-204.
- Searle, J. (2010) *The Making of the Social World, The Structure of Human Civilization*, Oxford: Oxford University Press.
- Sibley, W. (1953) 'The Rational Versus the Reasonable', *The Philosophical Review*, 62 (4): 554-560.
- Vanhanen, T. (2008) *Vanhanen's Index of Democracy*. Oslo, International Peace Research Institute (PRIO).