Can Mini-Publics Make Wise Decisions?

Deliberative Democracy Meets Technocracy

Vincent Schuttevaar

Abstract

Deliberative democracy has graced us with a new democratic innovation: the *mini-public*. These are small representative groups of the population that deliberate together about political issues. The normative appeal of this innovation, is that it can avoid populism by deliberating towards carefully considered preferences, without admitting to the yoke of technocracy. Yet the visionaries of expertise want to strike back. Can mini-publics really make wise decisions? They think not. Cognitive bias and a lack of domain-specific expertise stand in the way. But do they? This thesis argues that the answer is "no, cognitive bias and ignorance are not impregnable obstacles", and "yes, mini-publics can make wise-decisions". The reason for this, is design, which, if done carefully, should transform mini-public's follies into wisdom.

Contents

1. Introduction

Can mini-publics make wise decisions? This thesis answers "yes". Mini-publics – small representative groups that deliberate on a political issue – *can* make wise decisions. This conclusion is derived from a set of design recommendations that should ensure that mini-publics reliably make wise decisions.

Very briefly, these mini-publics are associated with the deliberative democracy literature (c.f. Dryzek, 2002; Gutmann & Thompson, 2004; Steiner, 2012). In this literature, the key message is that democracy should focus more on talk, and less on votes. The reason, is that standard preferences and opinions are open to irrelevancies and manipulations, and that careful deliberation can solve these issues. This applies to both society at large (Parkinson & Mansbridge, 2012), as well as small scale democratic platforms. And among these small scale platforms, are mini-publics (Fishkin, 2011; Setälä, 2014; Warren & Pearse, 2008).

Deliberative democracy is not without criticism, however. One common challenge, is that well-ordered deliberation is not realistic (Somin, 2016; Caplan, 2007). The argument revolves around citizens being both ignorant and irrational, and that these issues do not simply dissolve under some pretense of deliberative exchange of reasons and information. By extension, this criticism also applies to mini-publics (c.f. Brennan, 2017). So, given these issues, why would mini-publics be able to make wise decisions?

The journey leading to the thesis' conclusion – that mini-publics *can* make wise decisions – is as follows. Chapter two introduces the general framework of the thesis, by showing how we get from deliberative democracy more generally, to non-ideal theory about mini-publics. These mini-publics are then evaluated according to two criteria: i) *representativeness*, because mini-publics are by definition supposed to be representative, and according to ii) *relevant expert endorsability*, which is used as the (epistemic) criterion for wisdom.

Chapter three continues with these epistemic evaluation criteria, by reviewing those used in the literature. One reason for why *relevant expert endorsability* is chosen, is because it is the only criterion that is sufficiently strong by itself, without becoming too strong (i.e. demanding the truth). Another reason, is that *expert endorsability* is also used to evaluate mini-publics by some critics discussed in the thesis. Therefore, if mini-publics can satisfy this criterion, then these critics should be persuaded. However, given the controversial nature of expert endorsability, I make the criterion more precise by narrowing the meaning of *relevant expert*, and widening conditions needed for *endorsability*. This makes the criterion of *expert endorsability* more robust to criticism.

Chapter four uses both of these criteria to evaluate actual mini-publics. After giving an example, reviewing the literature, and taxonomizing types of mini-publics, the chapter asks whether these empirics can show whether there is a mini-public design that can pass the stipulated criteria. The claim is that there is not. One reason for this, is that mini-publics tend to be complex interventions, thereby not allowing causal inferences about individual design features. Another reason, is that evaluating mini-public types in their entirety is problematic due to selection-bias, too many uncontrolled variables, and publication bias. So even though we can find cases in which both criteria were satisfied, no robust inferences can be made about types of mini-publics.

Chapter five therefore redirects its focus to more theoretical concerns. The question here, is whether human reason is an unassailable obstacle to mini-publics making wise decisions. The claim is that it is not. The support comes from two sides. One side reviews the case of epistocrats (i.e. those favoring rule by those with knowledge). They present a strong case for political bias and absence of coherent incentive schemes, but likely overstate how dire the prospects of education are. The other side reviews the case for blaming bad deliberation on information and the deliberation environment.

Namely, the argumentative theory of reasoning shows how even cognitively biased agents become wise, just by putting reason in its proper environment (e.g. argumentative settings, public justification, evaluation of arguments made by others). Additionally, the theory of information phenomena shows how even rational individuals can converge on unwise answers, when the flow of information is not properly controlled. The corresponding solution, is to implement design features that reduce issues with information-flow. Taken together, these two theories – the argumentative theory of reasoning and information phenomena – show that human reason need not prevent mini-publics from making wise decisions.

Chapter six then inquires into whether such a design can be coherent, without any glaring normative issues, and without any other epistemic downsides. The claim is that such a design exists. I discuss design solutions for both *representativeness* and for *wise decisions*. Design solutions for *representativeness* (i.e. forced participation, opportunity costs compensation, social status improvement, politically binding decisions) have their respective issues. The options are therefore to bite the bullet, or to accept that mini-publics shall likely not be fully representative. However, this conclusion need not be too severe, as some of these design solutions, at least do not have any glaring normative issues. Next, I discuss design solutions for *wise decisions* (i.e. public justification, speaking order randomization, pre-deliberation votes, identification and evaluation separation, politically binding decisions, group incentives, prolonged duration). Even though these design solutions are not without possible issues, many of them can be deflated by either a) issues being offset by other design solutions, or b) issues not being too severe.

2. General Framework

This chapter motivates the relevance of studying mini-publics, and builds the general framework of the thesis. The most important choices in this framework are as follows: i) a minimal definition of deliberation, such that we can differentiate between good and bad deliberation, ii) narrow-scope deliberative democracy, such that the thesis can focus on mini-publics, iii) neutrality between different generations of deliberative democracy, such that particular values are not adopted *a priori*, iv) non-ideal theory, such that the thesis can respond to the domain of its critics, who ask whether mini-publics can actually make wise decisions, v) *representativeness* as both a definitional criterion and normative criterion of mini-publics, and vi) *expert endorsability* as the epistemic criterion used to represent wise decisions.

2.1 Justification

So, why would a question about mini-publics – an innovation in democracy's structure – be relevant in the first place? The reason is that democracy has been getting bad press. Recently, some of that publicity has been in *university* presses. Consider the following books: "How Democracies Die" (Levitsky & Ziblatt, 2018), "the Fate of the West" (Emmott, 2017), "How Democracy Ends" (Runciman, 2018) , "The People vs. Democracy" (Mounk, 2018), and "Against Democracy" (Brennan, 2017). These books argue that democracy is not the stable end-point of society, as it has sometimes been portrayed (Fukuyama, [1992]; 2006). Worse, people are losing trust in democratic leaders. Polarization has increased. "Partyism" has in the U.S. become more commonplace than other undesirable "isms" (Sunstein, 2015). And it is unsure whether current institutions can deal with the issues of tommorow.

Some attempt to explain these charges by factors such as the 2008 financial crash, within-country increases in income and capital inequality, and an aging demographic. But often, the mentioned perpetrator is democracy's structure.

To solve these issues, the literature has provided recommendations for structural change. I will mention several of them. Some options are rigorous deviations from the current norm. Examples are *meritocratic authoritarianism*, which advocates benevolent authoritarian leadership by those with merit (Nathan, 2003; Bell, 2016[1]), and *epistocracy,* which says that rule by those with knowledge is justified because democracy allows incompetent voters to decide over the fate of others (Brennan, 2017; Caplan, 2011). There are also libertarian options, such as *localism*, which advocates local government, in which voters can easily "vote with their feet" by moving to a different community (Somin, 2016). All these options so far, contrast with "more of the same options", such as *competitive democracy*, which advocates making democracy even more competitive, such that parties and politicians need to compete for votes, thereby better serving their constituents (Shapiro, 2016; Schumpeter, 1942). And lastly, there is deliberative democracy, which advocates for more deliberation (Dryzek, 2002; Steiner, 2012)[2]. Given the widespread appeal, deliberative democracy makes for a sensible choice to start investigating structural innovation.

2.2 Deliberative Democracy

In short, deliberative democracy is talk-centric instead of vote-centric. To illustrate, whereas vote-centric theories of democracy emphasize aggregation of preferences into a social choice, a talk-centric theory of democracy emphasizes the talking that goes on prior to voting. Ideally, this talking will involve lively exchange of various arguments and points of view, complemented with some careful consideration. When this happens, we can speak of deliberation. The normative appeal of such deliberation, is that it can transform preferences for the better prior to voting. Such transformation can happen because people exchange information and thereby gain knowledge, but also because people change their values, for instance via new arguments or via exposure to unknown social groups.

Now, in order to understand what deliberative democrats think is wrong with prevalent competitive democracies and their populist tendencies, consider the reasons given by the influential political scientist James Fishkin (2011-2/3)[3]:

1. **Rational Ignorance**: it is difficult to effectively motivate citizens in mass society to become informed.
2. **Phantom Opinions**: the public has fewer "opinions" deserving of the name than are routinely reported in polls[4].
3. **Selectivity of Sources**: even when people discuss politics or policy they do so mostly with people like themselves—those from similar backgrounds, social locations, and outlooks.
4. **Vulnarability to Manipulations**: public opinion as we routinely find it in mass societies is.. [vulnerable] to manipulation.

The claim is that these reasons should sufficiently alarm us, and caution us not to unreflectively endorse democracies – such as populism and competitive democracy – that suffer from these limitations.

In response to this problem, one might be tempted to conclude that experts, allegedly unhindered by the above limitations, should be selected and given authority to make political decisions. And this is

---

[1] Although Bell, in his book "the China Model", calls it "political meritocracy".
[2] The literature also talks about participatory democracy, which has elements of competitive democracy, populism and deliberative democracy.
[3] Text original, listing mine (labels appear on page 122).
[4] "Respondents to polls do not like to admit that they "don't know" so they will choose an option, virtually at random, rather than respond that they have never thought about the issue" (ibid-2).

exactly what is desired by technocrats (i.e. those favoring rule by those with technical expertise) and epistocrats (i.e. those favoring rule by those with knowledge), whose arguments will be discussed in chapter three and five.

However, expertise-centric governance is definitely not favored by deliberative democrats:

> This expert/boss fallacy is tempting, but someone's knowledge about what should be done leaves completely open what should be done about who is to rule. You might be correct, but what makes you boss? (Estlund, 2009-3)

Put differently, the argument is that technocracy does not have the same liberal legitimacy as standard democracy. And given that such liberal legitimacy is normatively necessary, technocracy is unacceptable. Deliberative democrats thus want democracy without reverting to either competitive populism or elitist technocracy.

2.3 Definitions

As Thompson (2008-501) notes, many definitions of deliberation mesh together both descriptive and evaluative elements associated with a concept into a single definition. Thomspon gives the following example of deliberative democracy:

> . . .we have deliberative democracy when, under conditions of equality, inclusiveness and transparency, a communicative process based on reason. . .is able to transform individual preferences and reach decisions oriented to the public good (della Porta, 2005-340)

In definitions like these, deliberative democracy is good by stipulation. Critics, however, will be unconvinced.

In the study of rhetoric, definitions such as the above, are also called *evaluative-descriptive*. Terms that are evaluative-descriptive do not merely describe something, such as "she has green eyes". They also normatively evaluate, such as "she is untidy" (negative connotation), or "she is orderly" (positive connotation).

Such rhetoric is not limited to deliberative democracy. Cherished descriptive terms frequently gain positive evaluative connotations that stray from their original meaning. One example is *science*. Susan Haack (2012-78) describes it as follows: "*as the prestige of the sciences grew, words like "science," "scientifically," etc., took on an honorific tone: their substantive meaning tended to slip into the background, and their favorable connotation to take center stage*". Another example is philosophy. Allen Buchanan (2009-278) for instance states that: "*[i]f one employs a sanitizing, stipulative definition of 'philosopher', treating that term as an honorific title for those who meet standards of excellence that are currently endorsed by mainstream academic philosophers in developed countries, one will conclude that the influence of Philosophy on public policy is much more benign than if one operates with a broader definition*". Such attempts at sanitation can probably be repeated indefinitely[5].

The rhetorical appeal of this strategy for those inside the relevant community is obvious. However, for dispassionate inquiry, evaluative-descriptive terms do not lend themselves well to differentiating

---

[5] E.g. see Rodrik (2015) for a case in economics.

between good and bad instances. Granted, deliberation has a mildly positive connotation, as can be seen from the following standard definitions[6]:

(1) long and careful consideration or discussion
(2) an exchange of arguments for or against something

The issue is that critics point toward cases in which these definitions seem to be satisfied, but bad outcomes ensue nonetheless. Therefore, I concur with Mansbridge (2015-29), who advocates a minimal definition: "*neutral and minimalist definition of deliberation and a deliberative system allows users of these terms to specify the conditions of "good" and "bad" deliberation without having those standards built into the word deliberation itself*". So, for the purposes here, definitions (1) and (2) are already sufficiently neutral/minimal.

2.4 Distinctions

Besides general definitions, some important distinctions can be made in the deliberative democracy literature. Among them are i) scope, ii) generations, and iii) ideal versus non-ideal theory. I discuss these in order.

Firstly, scope. On a narrow scope, individual institutions and their deliberative qualities are studied. The institutions that are studied can be familiar political institutions, such as the House of Representatives (in which politicians deliberate), or new democratic innovations, such as internet democracy, in which citizens gain new ways to participate. However, there is also another kind of democratic innovation that is studied with frequency – the one of interest in this thesis – and that is the mini-public (c.f. Setälä, 2014).

In short, mini-publics are small representative versions of the larger public – hence, *mini-publics*. In order to make these mini-publics representative, standard self-selection of politicians and experts will not be sufficient. Therefore, some kind of (stratified) random selection is needed. Such a process is reminiscent of a period in Ancient Athens, in which leaders were chosen by lottery. The reason for this practice, as Barbara Goodwin ([1992]; 2013-46) notes in her book "Justice by Lottery", is that "*the choice of leader by lot averted the danger that power would go to the rich or to those who desired it*"; a justification that is not too far removed from reasons given today by those that distrust politicians and technocrats. The additional appeal of mini-publics, is thus the potential to be perceived as trustworthy. Citizens can rest assured that the decision has come from 'people like them', and not a distrusted elite.

With respect to political ambition, some propose a similarly radical reform as the system in Athens, in which mini-publics replace the now standard competitive elections system (e.g. Barnett & Carty, [1998]; 2017). But there are also many others who characterize mini-publics as more of a political supplement. One illustration of this, is by one of the visionaries of the field, Robert Dahl:

> Suppose an advanced democratic country were to create a "minipopulus" consisting of perhaps a thousand citizens randomly selected out of the entire demos. Its task would be to deliberate, for a year perhaps, on an issue and then to announce its choices…One minipopulus could decide on the agenda of issues, while another might concern itself with a major issue…A minipopulus could exist at any level of government—national, state, or local. It could be attended..by an advisory committee of scholars and specialists and by an

---

[6] These definitions are also close to for instance "*mutual communication that involves weighing and reflecting on preferences, values, and interests regarding matters of common concern*" (Bächtiger et.al., 2018-2).

administrative staff. It could hold hearings, commission research, and engage in debate and discussion. (Dahl, 1989-340)

By contrast to these views on narrow scope deliberative democracy, there is also a wide scope. The following is a similarly visionary description of such a wide scope:

> To understand the larger goal of deliberation, we suggest that it is necessary to go beyond the study of individual institutions and processes to examine their interaction in the system as a whole. We recognize that most democracies are complex entities in which a wide variety of institutions, associations, and sites of contestation accomplish political work – including informal networks, the media, organized advocacy groups, schools, foundations, private and non-profit institutions, legislatures, executive agencies, and the courts. We thus advocate what may be called a *systemic approach to deliberative democracy*. (Parkinson & Mansbridge, 2012-1/2).

This wide scope of deliberative democracy is also relevant for mini-publics, which eventually need to be linked to the deliberative system (c.f. Curato & Böker, 2016). Additionally, it is also relevant, because *"..while it is true that a large number of deliberative scholars research mini-publics, these studies are motivated by the desire to better understand how lessons learned from small-scale deliberative forums can be scaled up to mass democracies and enhance the quality of political participation*." (Curato et.al. 2017-32).

Secondly, there is the distinction between first- and second-generation deliberative democracy (Bächtiger et.al., 2018-4/5; Curato et.al. 2017). The difference between them are their standards for good deliberation. The first generation's ideals for good deliberation – associated for instance with Jürgen Habermas and Joshua Cohen[7] – are as follows (Bächtiger et.al., 2018-4/5):

1) Respect
2) Absence of power
3) Equality
4) Reasons
5) Aim at consensus
6) Common good orientation
7) Publicity
8) Accountability
9) Sincerity

Most of these ideals, will have intuitive appeal. Several might even be intuitively obvious. For instance, we require *respect*, because we cannot accept that our statements are viewed "as objects to be dismissed, demeaned, manipulated, or destroyed" (ibid-5); and we require *absence of power*, because "the threat of sanction or the use of force… should not play a role" (ibid-6). Some other ideals might also be intuitively plausible, such as requiring some *publicity*, which will encourage deliberators to think more before committing to a decision, and will also prevent closed-door politics. In turn, this raises *accountability*, which is core to most democracies. Another example that is intuitively plausible, is requiring the free exchange of *reasons*, which will increase the likelihood that the better argument will triumph. However, some ideals will be less intuitively obvious. One example of this, is *aim at consensus*. To see how the ideal could be desirable, consider a world in which everyone has a *common good orientation*. In this scenario, given that everyone has the same goal, no

---

[7] Although as Bächtiger et.al.(2018) show, Habermas' and Cohen's work changed through time, and did not always get interpreted properly.

concessions or clarifications of conflict are needed, and therefore an aim at consensus – together with free exchange of reasons — should lead to the most desirable decisions.

Now, the second generation's standards differ – except for the first two ideals – because they either make the ideals more precise, or stipulate exceptions. For instance, publicity is desirable under many circumstances, but not under conditions which require strategic interactions (under the assumption that corruption can be prevented), as publicity will then only prevent open discussion of important matters. To give another example, the second generation has given critique on the reason-emotion dichotomy. Namely, they argue that participants should not be required to give purely rational reasons for their claims, but that participants should also be allowed some rhetoric and storytelling, in order to convey all relevant considerations. Similarly, for the remaining ideals, the second-generation argue for comparable exceptions to first generation ideals, thereby not refuting prior ideals entirely, but rather refining them.

Thirdly, there is the distinction between ideal and non-ideal theory. For my purposes here, it is sufficient to regard ideal theory as a kind of vision, unconstrained by immediate empirical facts[8]. General justification for ideal theory, is its creativity and innovation potential, and perhaps the correct evaluative standards it can generate. In the deliberative democracy literature, much has been ideal theory. For instance, many of the ideals listed above have been inspired by visionaries who were not immediately concerned with empirical data on deliberation. Or to give another example of ideal-theory in deliberative democracy, consider the *wisdom of the crowds* literature, or at least parts of it. This literature shows under what conditions (e.g. no systematic bias, a sufficiently large group out of which some know the answer, cognitive diversity with learning diffusion, ideal deliberation) deliberative democracies make wiser decisions than small groups of experts (c.f. Landemore, 2014). These theories often focus on building a normative model, without arguing explicitly for the plausibility of these conditions obtaining for the empirical target (e.g. real mini-publics).

The above is in contrast to non-ideal theory, which tends to focus more on feasibility of ideals, and on what we should do in order to achieve our goals. A prime example, is recent work done in deliberative democracy, which performs real world experimentation with mini-publics (e.g. Fishkin, 2011). The focus here, differs in several ways from ideal theory. Namely, relevant questions here are: to what extent can mini-publics attain stated ideals, which mini-public design increases the likelihood of attaining these ideals, and what should we do to prevent observed issues, such as belief polarization and domination of the deliberation by those with higher status? Also relevant, is that questions like these, and this type of literature more generally – together with literature about human reasoning, education and political science – is the focus of those who doubt mini-publics can make wise decisions.

2.5 Locating the Thesis
So, to briefly illustrate mini-publics further (see 4.1 and A.1 for a more detailed account), consider the entire process that goes into a mini-public. Before deliberation begins, a group of staff needs to organize all the relevant material, ranging from setting an agenda, to collecting a group of competing experts who will prepare lectures and briefing material. Next, a randomly chosen sample of the population comes together to deliberate about the political issue. Participants experience several days (up to months) of plenary sessions and small group deliberation, after which they will finally make a decision, either by vote or by a recommendation report.

---

[8] See Valentini (2012) for finer distinction of ideal theory: i) full compliance vs. partial compliance, ii) utopian vs. realistic theory, and ii) end-state vs. transitional theory.

It is mini-publics such as these, that the following research question is about:

> **Research Question**: Can mini-publics make wise decisions?

Now, in light of what is said above, the thesis uses a minimal definition of deliberation:

> **Definition of Deliberation**: long and careful consideration or discussion, an exchange of arguments for or against something

And with respect to the distinctions, this thesis uses the following:

> **Scope**: narrow scope (mini-publics)
> **Generations:** a priori neutral between values of both generations
> **Type of Theory**: non-ideal theory

Neutrality between generations is chosen, because commitment to ideals of a specific generation is not necessary at this stage. Additionaly, non-ideal theory is chosen, because both the relevant design literature and critics this thesis focuses on, is primarily non-ideal theory. The thesis therefore amounts to a non-ideal inquiry into whether mini-publics can make wise decisions.

Now, the focus of the thesis is on epistemic success, but if the thesis is not to be normatively irrelevant, some ethical constraints need to apply. Among them are core deliberative values: respect, absence of power, and some normative version of equality.

Representativeness – a version of political equality – is especially relevant. The reason is that mini-publics are supposed to be representative of the population, as one of its prime justifications, is that decisions come from "people like us", not "people a little bit more eager and elite than us". By definition then, mini-publics "*are distinguished by their mode of selecting citizens, namely random selection*" (Smith, 2009-72), thereby trying to ensure representativeness. Or to put it differently: "*[to] put the whole country in one room*" (Fishkin, 2011-25)[9]. Therefore, the thesis also uses the following normative criterion:

> **Normative Criterion:** representativeness

The last ingredient that is needed for a well-defined research project, is an evaluation criterion for wise decisions. Here I am going to focus on the knowledge side of wise decisions. In philosophy, the study of knowledge is also called epistemology. In epistemology, the traditional criterion for knowledge is true, justified belief[10]. But since truth is not something that we can reliably obtain in decisions made by mini-publics, a weaker epistemic criterion is needed in order to evaluate epistemic success. In this thesis that criterion is the following:

> **Epistemic Criterion:** decisions *endorsable* by *relevant* experts

In the following chapter, specifically in sections (3.1) and (3.2), I shall further explain why this criterion is chosen, what it means, and what to do when the criterion cannot be met. After this, "wise decisions" and the epistemic criterion are used interchangeably.

---

[9] The ideal of representativeness seems standard, although some caution with a feasibility note: "still other deliberative democrats placed their hopes in "mini-publics." These are designed to be groups small enough to be genuinely deliberative, and representative enough to be genuinely democratic (though rarely will they meet standards of statistical representativeness, and they are never representative in the electoral sense) (Goodin & Dryzek, 2006-220).

[10] Although this account has been systematically challenged (e.g. Gettier cases).

3. Epistemic Criteria

This chapter argues that "decisions *endorsable* by *relevant* experts" is the correct epistemic success criterion for mini-publics. There are two reasons for this. The first reason, is that out of all epistemic criteria in the literature, no other single criterion is a sufficiently strong measure of epistemic success, without becoming too strong. The second reason, is that expert endorsability is also adopted by some critics of mini-publics (i.e. technocrats and epistocrats). Since these critics are relevant for this thesis (chapter three and five), expert endorsability is a more relevant criterion. However, given that the criterion is controversial, I shall narrow the definition of relevant experts and widen the criteria for expert endorsability, thereby salvaging the criterion against some of the criticism it has been subjected to.

3.1 Ranking Epistemic Evaluation Criteria

Currently, the literature has no systematic review of epistemic criteria on which mini-publics could (and should) be evaluated. To be sure, there *are* review papers discussing deliberation-measurements. One example is Estlund and Landemore (2018), who, in their section "measuring the epistemic quality of democratic deliberation", mention criteria five and six from the list below[11]. Another example, is Black et.al. (2011), who discuss criteria one to six from the list below[12]. In addition, they discuss qualitative case studies, which could be interpreted as a subset of what is listed below as criterion seven.

However, what these reviews lack most, is a systematic evaluation of the epistemic strength of these measures[13]. Such a ranking would help in evaluating mini-publics. Additionally, a ranking could help explain why some see mini-publics as an empirical success, whereas others do not. Namely, some academics might have stronger epistemic criteria for success than others, leading them to different conclusions[14]. Therefore, ranking criteria on epistemic strength seems fruitful.

The following is thus such a list of possible epistemic criteria, ranked based on ascending (categorical) order of epistemic strength:

    I.   Too Weak[15]

---

[11] In addition, they mention direct outcomes, such as the success of deliberative ancient Athens, and experiments which show that deliberating groups are good at solving logical and factual puzzles. These outcome measures are not applicable to mini-publics, as mini-publics do not have similar outcomes we could measure (i.e. such as the success of Athens or a true answers to a logic puzzle).

[12] Although consensus is only mentioned indirectly in the forming of consensus seeking, which is part of some measures.

[13] Caplan (2012-329) even complains that: "without a standard of competence, how can Landemore begin to judge the epistemic quality of democracy…?". Though it should also be noted that Black et.al. (2011) do critically discuss some of these measures (with a slightly more global focus than the subcategories of mini-publics and epistemology).

[14] I will not return to testing this explicitly, however, even though it seems plausible, as critics discussed in the thesis hold criterion seven, instead of the weaker criteria some other academics use.

[15] Single-peakedness are also sometimes used to measure epistemic improvement, as it can indicate internal (preference) coherence and helps prevent majority cycles. I am unsure whether internal-coherence indicators are strong enough epistemically (e.g. there is also conformation bias, among other issues with coherentism epistemology). And with respect to majority cycling (an epistemic ailment), the study of single-peakedness is based not on its pressing empirical presence, but rather on further understanding. For instance, List et.al. (2012-80/1) say that "[a] number of authors have recently argued that majority cycles are empirically rare (Miller 2000; List and Goodin 2001, Appendix 3; Tsetlin, Regenwetter, and Grofman 2003; Mackie 2004; Regenwetter et al. 2006). That may well be (the literature as a whole has not yet conceded the point), but we

1. Participant assessments
2. Preference/opinion change
3. Consensus
4. Currently endorsed mini-public design

II. **On the Border**

5. Epistemic quality of deliberation
6. Passing knowledge tests of relevant material

III. **Sufficient**

7. Decisions *endorsable* by *relevant* experts

IV. **Too Strong**

8. Truth

The ranking is not supposed to be an exact ascending list of epistemic strength, but the categories are supposed to be roughly correct. In what follows, I will first discuss criteria that are too weak. Then, I discuss criteria that are on the border, and comment on whether these criteria, if taken jointly, would not also be sufficient. Lastly, I comment on why truth is generally seen as too strong to use as an evaluation criterion in this literature.

Firstly, epistemic criteria that are too weak. Many of these epistemic criteria correlate with epistemic success, but they do not indicate epistemic success by themselves. Popularity ratings among non-experts, such as participant assessments and consensus, are examples of this. Assessments can be based on a host of other factors besides how much wiser someone has become, and popularity might be driven by groupthink, social pressure, or just a lack of understanding and information.

Similarly, preference- and opinion-change post-deliberation can be both good and bad (c.f. Baccaro, Bächtiger & Deville, 2016). We would want to know whether such preference change is driven by knowledge-gains (or by a shift in values). In the literature, a common measure is to let participants do knowledge tests, and see whether high scores on these tests correlate with opinion change (c.f. Fishkin et.al. 2010). The idea is that, if this relationship between opinion change and learning is robust enough, then opinion change can be used as a proxy measure for epistemic success. However, even if opinion change was a robust enough proxy, the actual epistemic criterion that is used, would be knowledge as judged by official facts or by expert evaluation, which are criteria six and seven. Therefore, if opinion change is a good proxy, then it is because these stronger criteria are satisfied, not because of opinion change itself.

Also too weak is 'currently endorsed mini-public design'. The idea of the criterion is that, if what theorists think is good design is implemented, then wise decisions follow. But this would only be a sufficient criterion if there were no (reasonable) sceptics about such design being sufficient to convince. Unfortunately, that is not the case.

Secondly, there are also epistemic criteria that are close to acceptable, and would be chosen if stronger criteria were not available. One of these, is criterion five: epistemic quality of deliberation. Currently, one way in which deliberation quality is operationalized, is the discourse quality index (Steenbergen et.al. 2003). In this index, there are coding rules for – among other non-epistemic factors – level of justification of one's claims, and of one's response to counterarguments[16]. Now,

---

should still wish to understand why they are not more frequent, and what helps minimize their frequency".
Lastly, criteria such as single-peakedness can be used by experts – criterion (7) – if they are important in certain circumstances.

[16] There are also other coding schemes of deliberation quality, such as the speech act analysis (Holzinger, 2004) and Stromer-Galley's coding scheme (2007). In principle, these indexes could be transformed to focus primarily on epistemic quality.

providing justifications and responding to counterarguments do indeed seem necessary conditions for epistemic success; ignoring each other hardly seems fruitful, even in heated arguments. However, such deliberation quality is not sufficient for epistemic success. The reason that deliberation quality is insufficient, is that even though participants are responding well to each others reasons and counterarguments, that does not ensure that the content of these reasons and counterarguments are satisfactory. Even worse, there is the potential for content of reasons and counterarguments to be actively misleading, thereby derailing the deliberation process. Now, to solve these issues, one could start evaluating the quality of the content. The problem with this, however, is that doing so requires experts to do that evaluation. And, if one starts using expert evaluation, then the criterion comes closer to something like "content of arguments endorsable by relevant experts", which blends into criterion seven. Therefore, epistemic quality of deliberation as used now, is not a sufficiently strong criterion for epistemic success.

Another criterion that is close to acceptable, is criterion six: passing knowledge tests of relevant material. Knowledge tests of official facts is one way to accomplish this. A problem with tests like these, however, is that they might be too weak for ensuring understanding of the material, and for ensuring a capacity to manipulate the material outside of the confines created by the test[17] (c.f. Vanlehn & Van de Sande, 2009). To solve this issue, one could make more elaborate tests that truly evaluate understanding. Examples are qualitative interviews, and case studies that need to be solved with conceptual understanding. The problem is that such understanding tests will be made and evaluated by experts. This would make criterion six come closer to something like "passing understanding tests endorsed by relevant experts", which also starts to blend into criterion seven. Therefore, passing knowledge tests is also not a sufficiently strong criterion for epistemic success.

So, individually, criteria five and six are judged to be insufficiently strong. But can they nevertheless be jointly sufficient? The answer seems to be "yes". One major reason for this, is that deliberation quality and knowledge tests seem to complement each other. For instance, a weakness of deliberation quality, is that it cannot guarantee that content is relevant. This is complemented by knowledge tests, as having basic knowledge makes it much more likely for content to be relevant. Conversely, a weakness of knowledge tests, is that it cannot guarantee that there is understanding in more complex contexts. This is complemented by the complex context of deliberation, in which confined knowledge is exposed to complex justifications and counterarguments. If this combination works well, then knowledge more likely indicates understanding. Therefore, admittedly, given that deliberation quality and basic knowledge complement each other, they could jointly be a sufficiently strong criterion for epistemic success. However, whether it *is* sufficient or not, the joint criterion of five and six will not be explicitly used in this thesis.

Thirdly, there is also criterion eight (i.e. truth) which is too strong, and therefore not often used to evaluate mini-publics. Currently, deliberative democrats that defend the use of truth, do so *not* because it is an attainable evaluation criterion. They do so because it is and ideal that implicitly hides in many normative accounts of deliberative democracy. For example, instead of accepting that there is nothing to epistemically attain beyond "reasonable disagreement", Landemore (2017-290) claims "*that there are better and worse answers to at least some political questions and that those can be fallibly approximated through democratic procedures, such as deliberation*". The claim is thus in the domain of ideal theory; truth is seen as an ideal, which democratic procedures should strive to achieve. That does not entail that we should use truth as an evaluation criterion of actual deliberation. Namely, the reason for being cautious with using truth, is that many political issues do

---

[17] Much like the issues with multiple choice exams. And similar to such exams, just passing tests can still mean having many mistakes, even within these built-in confines.

not have clearly true answers. At least not in the sense we would find in some domains of physics or mathematics. One reason for this, are the values that are involved in political decision-making. Another reason, is the comparatively weaker epistemic power of the social sciences. So, even if we accept an epistemology in which truth is an attainable goal, we still have reason to be careful with using truth as an evaluation criterion for mini-publics.

So, the only criterion that is judged as sufficient on its own, is therefore seven: decisions *endorsable* by *relevant* experts. Also relevant to note, is that an additional advantage of this criterion, is that it is congenial to sustaining dialogue between deliberative democrats, and some of its critics (e.g. epistocrats; Brennan and Caplan), who will show up next.

3.2 Decisions Endorsable by Relevant Experts

The criterion of choice is thus "decisions endorsable by relevant experts". However, given that using experts endorsability as an epistemic success criterion is controversial (c.f. Ingham, 2013; Reiss, 2019; Gunn, 2019), I will shortly explain what I mean by this criterion.

A basic definition of an expert is '*a person who is very knowledgeable about or skilful in a particular area'*. The takeaway from this definition, is that those labeled as "experts" are not the same thing as experts. For example, a university professor is labeled as an expert by the institution, but that does not always translate into actual expertise. One way to solve this issue, is by refering to "experts" and "true experts". In this discussion, however, I will stick to only using "experts" in its original definition.

So, why is endorsement by experts a good epistemic criterion? Some reasons were already given above. Caplan (one of the epistocrats), however, describes it as follows:

> My empirical approach does not rule out the possibility that the public is right; nor does it rule out the possibility that the experts are wrong. Its key assumption is simply that after controlling for a long list of possible confounding variables, any remaining lay–expert belief gaps are evidence of public bias (Caplan, 2012-328).

On the face of it, this seems highly plausbile, especially given the size of the group of experts.

So why is it controversial? The reason, Gunn says (2019-33), is that "Caplan does not demonstrate that economists are truly expert, and Brennan does not demonstrate that social scientists are". What Caplan demonstrates instead, is as follows. He first draws an analogy between economic experts and "math, science, history, and car repair" (Caplan, 2007-81). Where a mathematician is an expert about math, an economist is an expert about the economy. Next, Caplan (ibid-81/3) give some reasons to think so: common sense, control for demographic variables (e.g. wealth), inaccuracy of right-wing ideology accusations (economists in fact lean left), and correlation between education and economic thinking. These reasons all point towards economists being experts; not self-serving ideologues. The burden of proof – the argument goes – is on those criticizing economists' expertise about the economy.

For Reiss (2019) and Gunn (2019) this is not satisfactory. They point out issues such as the following:

1. Dissent in economics exists[18]

---

[18] Gunn (2019-8): "Brennan's position would be more compelling if it did not neglect the contested nature of free-market economics even among economists, not to mention the criticisms that have been made of economics as a whole for hundreds of years, from Karl Marx and the German Historical School to such recent philosophers of social science as Philip Mirowski (1991), Daniel Hausman (1992), Alexander Rosenberg (1992), Tony Lawson (1997), Geoffrey Hodgson (2001), and Julian Reiss (2008)".

2. Consensus is not satisfactory[19]
3. What economists think/publically pronounce, and what their work says, are different things[20]
4. Empirical generalisations in the social sciences hold only context-dependently[21]
5. Agreed-upon evidential standards are absent in the social sciences[22]
6. There is considerable fact-value entanglement in the social sciences[23]
7. Economists fail at prediction[24]

Therefore – the argument goes – the burden of proof is on experts to show that they can cope with these issues.

The point here is not to evaluate all these issues. It is rather to make my epistemic requirement more robust to many of them. One way to do that, is to stick to the "true expert" story. This strategy has a downside however, which is that Gunn and Reiss would put the true expert in the "too strong" section of my evaluation criteria list. Therefore, another way is needed.

There are two tactics that I use to accomplish this: i) widen the standards for "endorsable", and ii) narrow the standards for "relevant experts".

Firstly, decisions endorsable by experts need not mean that experts want option "A" and therefore they will only endorse "A". It could also mean that, given the particular values of the mini-public, experts will recognize that option "B" is a good decision for the participants. For instance, if experts on farming really like meat, and they think that farming cattle will help supply that meat, then they would want to farm cattle. However, if the mini-public, and therefore by extension the public at large, are vegetarians, they would rather focus on dairy farming. The experts, recognizing that the mini-public has different values, see that the mini-public decision is consonant with these values, and therefore endorse the decision.

Another option, is that experts know that option "C" is bad, but think that the other options are at least acceptable. To illustrate, suppose that the topic under deliberation, is free-trade for a developing country, and that a decision needs to be made about what sectors should be protected from foreign trade. Also suppose economists on average want option "A": free-markets. The reason they give, is that – although they are aware of the limitations of the law of comparative advantage and of various other market failures – free markets is still the best heuristic to apply. Now suppose that the mini-public instead wants to choose option "B": protect relatively new industries in the developing nation from foreign trade, at least for a certain period, before competing with established international industries. The reason they give, is that otherwise these new industries will be immediately outcompeted and have no chance of success. In other words, they use the *infant industry argument*. Now suppose economists know about the infant industry argument, and about

---

[19] Reiss (2019-8/9) talks about conformism. Gunn (2019-11) talks about "shared dogmas".

[20] This is the argument in a recent popular book by Dani Rodrik (2015). Gunn mentiones Rodrik, and Reiss gives similar examples as Rodrik.

[21] Reiss (2019) for instance gives the example of Ha-Joon Chang's work, who shows that today's rich countries have become rich with anti free-market policies. This goes against the 'all economists know free trade is good' type of contextless slogan.

[22] There are different kind of models (deductive and inductive) and even within these kinds of models there is disagreement on the proper method and answer; see for example the minimum wage debate.

[23] Reiss (2019-6) gives an example about GDP: "[v]alues are involved in measuring GDP (e.g. Stiglitz, Sen, and Fitoussi 2010), the theories of rational choice that are often employed in modelling economic relationships (Sen 1993), the testing and acceptance of scientific hypotheses (Rudner 1953)". Since technocracy/epistocracy is often thought to lose its legitimacy if the expert's values are imposed on the public, such fact-value entanglement is a serious problem for experts in public policy.

[24] Tetlock (2005 [2017]) is the common reference here.

the data used in the literature to support it. Also suppose that economists do *not* know for which industries and countries the argument applies, and that they are slightly sceptical about their chances to guess. Lastly, assume that the average economist thinks that applying foreign trade protection to some key infant industries, is – even if likely not optimal – at least not disastrous. This allows the economist to endorse "B". Namely, what the economist really wants to prevent, is the mini-public choosing option "C": protectionism of developed industries; those that are already competing at an international level[25].

To reiterate, experts are said to endorse a decision when either values of participants are aligned with the decision, or the decision avoids bad options, and is therefore acceptable.

Secondly, "relevant experts" does not mean that those who study the economy are necessarily experts. There are several reasons for this. One reason, is that there are other fields besides economics that study parts of the economy (e.g. political science, sociology). These fields can have different points of view. To solve this issue, an account is thus required that shows who the relevant experts are (and who the "true experts" are). Another reason, is that not all work done by economists can be used for policy, as some work will be on a level of abstraction – without obvious political implications – that is suitable for academia only.

Now, this only leaves a subset of economics that can be used for policy. However, even this subset needs to be subdivided when we want to talk about "relevant" expertise. For one, not all economists and their work are alike, and we need to know which we should trust as experts. And, if we do not want to blindly trust mainstream opinion, then we might want to hear some evidence; for example about cases of success and failure of policy economics, and what criteria are used to define that success. A good example of this practice, is by the economic historian Roger Backhouse (2010), who goes over several case studies (i.e. creating new markets, transitions from socialism to capitalism, globalization, financial markets). The following conclusion is derived from these cases:

> The conclusion that emerges from these chapters is that, where problems are narrowly and precisely defined, and where they involve agents whose motivations are well understood and who operate under well-understood constraints, economic analysis is remarkably powerful (Backhouse, 2010-16)

One of the problems, however, is that many policy issues do not satisfy these constraints. Backhouse (ibid-17) therefore also notes that "*[t]he main lesson from these various case studies is that, when judging success or failure, very careful attention needs to be paid to the criteria that are being used*". For example (ibid-27-35), in the case of auctioning off British 3G mobile phone network licenses, if our evaluation criterion is 'maximizing government revenue', then economist intervention was a great success. However, if our evaluation criterion is something like 'efficiency' or 'ensuring competition', then success is much more doubtful. Namely, winners of the license had large debt problems, British Telecommunications was broken up, and postponements were requested for setting up 3G. It therefore seems possible that firms overbid for these licenses. A possible reason for this, is the pressure on management to signal strength to shareholders by bidding high, even though they had private information that bids might be too high to sustain. This phenomenon is much like the traditional principal-agent problem, in which the agent (i.e. management) has an incentive to make decisions, often short-term, that are in the interest of the agent, and not in the interest of the principal (i.e. shareholder). The lesson is thus that one needs to be careful in judging whether the narrow criteria are met.

---

[25] Even for this option there might be exceptions, depending on what assumptions one is willling to make. An example is protectionism of domestic agriculture in order to prevent nations from being completely dependent on other countries, especially relevant during times of conflict.

So, a relevant expert about economics is someone who applies [relatively mainstream] economic analysis to problems which are i) narrowly and precisely defined, ii) motivations of agents are well understood and operate under well-understood constraints[26]. That means that for the cases in which these criteria are not satisfied, a weaker epistemic criterion should be used. Naturally, this case only involves economics, but similarly narrow criteria for "relevant" should be used for other fields[27].

Lastly, there remains the question of how "decisions endorsable by relevant experts" should be operationalized. Namely, if there is no way in which the criterion can be easily applied, there is the risk of sailing too close to ideal theory. Now, one way to select relevent experts, is to get reasonable stakeholders to agree that the above conditions have been met for a certain group of experts. For instance, by some supermajority or consensus. Another option, is to select a group of ad hoc experts from uncontroversial domains of science. Next, from this selection of experts, a qualitative statement should follow that the decision is endorsed. To be sure, there are likely to be many cases in which agreement on expertise will not be met. In that case, a more diverse group of experts can be selected, such that all stakeholders have a representative (group of) experts. The same qualitative statement of endorsement is then required from this more diverse group.

To reiterate, decisions need to be *endorsable* by experts, which means that the decisions are consonant with values of participants, or among the acceptable options from the viewpoint of an expert. The designated experts are also supposed to be *relevant* experts, which means that sufficiently narrow criteria need to be applied, in order to infer what their particular domain of expertise is. Lastly, to operationalize "decisions endorsable by relevant experts", stakeholder approved experts should provide statements of endorsement.

4. Mini-Publics: Empirical Evidence

This chapter reviews the empirical literature on mini-publics, and inquires into whether this literature can point towards a mini-public design that can reliably make wise decisions (i.e. pass the criteria of *representativeness* and *expert endorsability*). The claim is that the empirics cannot do so. The reason is twofold. Firstly, most mini-publics are complex interventions, which do not allow causal inference about individual design features. This prevents us from making an amalgamation of the most effective design features. And secondly, evaluating mini-publics in their entirety is problematic, due to selection-bias, too many uncontrolled variables, and publication bias. Therefore, even though there are mini-publics that satisfy the criteria, no inference can be made about a type of mini-public being able to reliably make wise decisions.

Before the chapter delves into these specific arguments, the chapter first illustrates the mini-public procedure with an actual example, reviews current perceptions of mini-public's success, and taxonomizes common types of mini-publics.

4.1 A Case of a Mini Public

So, how does actual deliberation in a mini-public look like? A good impression can be gained from going over a particular case. An example that is commonly used, is that of the British Columbia Citizens' Assembly (2004)[28]. In this Assembly, 160 people from the Canadian province came together to deliberate about a change in the voting system for their electives. Given that different voting

---

[26] It could of course be debated per case whether these constraints are met or not.
[27] I do not specify these criteria here.
[28] See also Warren and Pearse (2008) for an academic review of the Assembly.

systems genuinely matter for which officials get elected, the decision became of political interest. The norm is that these decisions are made by experts. But this time it was done by citizens.

So, over the course of eleven months, citizens came together to deliberative about a decision. These citizens came from an initial 23,034 (random) invitations to participate in the Assembly. Out of these, 1,715 citizens responded positively. Eventually, from these, 158 were chosen randomly (stratified by gender, district and age) in order to be representative of the population. Lastly, two aboriginals were added due to underrepresentation, making a total of 160 members. With respect to compensations, all members' travel, food and accommodation costs were covered, in addition to $150 per meeting day.

Before the members of the assembly started with the process, they had a social reception, and a discussion about what shared values they could accept as a basis for working together. The following is the list they agreed on:

- Respecting people and their opinions
- Open-mindedness - challenging ideas not people
- Listening to understand
- Commitment to the process
- Focus on the mandate – preparedness
- Simple, clear, concise communication
- Inclusivity - all members are equal
- Positive attitude
- Integrity

This list was prominently displayed in all discussion rooms.

After these agreements, the Assembly started with a learning phase of six weekends (one weekend break in between). The curriculum for the learning phase was assembled by academics, and was reviewed by a consultative committee made up of other experts. The following are the weekly topics of the learning phase:

1. Introduction to the Citizens' Assembly
2. Elections and Parliamentary Government
3. Democratic Electoral Systems (part 1)
4. Democratic Electoral Systems (part 2)
5. Changing Electoral Systems
6. Options for Public Discussion

Formal sessions began on Saturday morning at 9:00 am, ending that day at 5:00 pm, and beginning again on Sunday at 9:00 am and ending at noon. These sessions contained lectures (45-50 minutes), discussion between small groups of ten to fifteen people (45-60 minutes), a spokesperson to report group results, and informal breaks (30-45 min). Prior to each of these weekends, a draft of the agenda and briefing materials were provided.

After these lectures, a preliminary statement followed that outlined salient issues which should receive attention. This preliminary statement was drafted by the research staff, and modified by the Assembly. The document also contained an invitation to members of the public to participate. This participation was done in the public hearings phase, where 50 hearings took place in which 350 citizens presented to the Assembly across the province (which made attendance of Assembly members selective). There was also an opportunity for written submissions, which were curated by staff for reading digestibility.

Next came the deliberation phase. The objective here was to reach a consensus about a recommendation, and perhaps suggest alternative options. The deliberation was done in six weekends, in which pre-structured topics were asked to be deliberated about in order to work towards the objective. These discussions, like other discussions, were facilitated by a moderator. In addition to live deliberation, there was an online discussion forum which allowed further discussion between Assembly members. Lastly, the deliberation phase also contained some presentations chosen from public hearings, and plenary overviews by the research staff.

The first four sessions of the deliberation phase were supposed to achieve the following (ibid-89; list original, text compressed):

- Articulate underlying values
- Choose the type of electoral system that incorporate these values
- Construct detailed versions of possible alternative electoral systems that could incorporate these values
- Compare these alternatives against each other
- Decide on a recommendation

Lastly, session five and six focused on finalizing the process, and discussing the actual content and design of the Final Recommendation Report.

4.2 On the Success of Mini-Publics

Now that we have an idea of the nature of mini-publics, we can look at what the literature thinks about their success. Given that the literature has no agreed upon definition of success (epistemic or otherwise), no such criterion is adopted for this particular section.

So, currently, scholars disagree on what the literature says about the performance of mini-publics. Consider comments by mini-public **supporters**:

> The current weight of findings strongly supports the claims of those who see mini-publics as a site of democratic deliberation (Setälä & Smith, 2018-307).

Assuming this is true, that would suggest that current focus should be on institutionalizing mini-publics and on giving them a substantial role in our democracies.

Now contrast that with claims by for instance Brennan (2017), who after reviewing the empirical literature argues that we should consider being **detractors**:

> [A]dvocates of deliberative polling don't yet have sufficient evidence to proclaim it a solution to our troubles (Brennan, 2017-67).

> The available empirical evidence, however, more strongly supports the view that democratic deliberation tends to stultify and corrupt us (ibid-187).

> I doubt a couple days of deliberation can impart that knowledge— after a semester's worth of study, most undergraduates still don't understand, say, basic microeconomics (ibid-215).

So, which of the above accounts is most reasonable? A clear-cut victory for deliberative democracy? Or a gloomy view of its prospects? Should we be supporters? Or should we be detractors?

Perhaps there is a way out of this dichotomy. The heuristic that the author is inclined to use, is to look at recent meta-reviews by scholars that have a history in the field[29]. Two prominent ones are Thompson (2008) and Gastil (2018). Consider some of the comments, which overall are **balanced**:

> The general conclusion of surveys of the empirical research so far is that taken together the findings are mixed or inconclusive (Thompson, 2008-499).

> One common observation is that these gatherings vary widely in their ability to generate substantial changes in issue knowledge, policy preferences, and civic attitudes—with some events producing underwhelming levels of such changes—at least on some issues (Gastil, 2018-281).

> Deliberation scholars frequently call for more systematic comparisons of deliberative designs (e.g., Bachtiger et al. 2009), yet few studies have managed the kind of field experiments Carman et al. (2015) conducted (Gastil, 2018-281).

> Because unsuccessful deliberative events too often go unreported in the published scholarly literature, we cannot know for certain the rate of success for such endeavors. We do know, however, that successful events of many varieties abound (Gastil, 2018-280)[30]

Results are thus mixed on many standards, including epistemic ones. Yet deliberative democrats do not think that these results are definitive. The following is the reason they remain **hopeful**:

> The main reason for the mixed results is that the success or failure of deliberation depends so much on its context. The contingent character of these results may seem to give theorists hope. If only theorists can identify the right conditions, they can confidently continue to extol the virtues of deliberative democracy (Thompson, 2008-499/500).

So, can these hopes be cashed in? In what follows, the chapter shall try to show that the answer is "we do not know" from the empirical side, and shall try to explain why. It does so by taxonomizing the design features of different types of mini-publics, and looking at what inferences can be drawn from these taxonomies. The idea is that causal inference is difficult for individual design features, due to the nature of complex interventions. Moreover, for evaluating entire complex interventions, based on "decisions endorsable by relevant experts" and "representativeness", known issues are knowledge of mixed results, too many uncontrolled variables and publication bias.

4.3 A Taxonomy of Mini-Publics

Two key features of mini-publics are deliberation and small-scale representation of the public. Five common versions of democratic innovations *seem* to satisfy these conditions[31]. These are Deliberative Polls, Citizens' Assemblies, Citizens' Juries, Planning Cells, and Consensus Conferences. Most of these are still relatively young. Experimentation with planning cells, citizens' juries, and consensus conferences started in the 1970s or 1980s, Deliberative Polls started since 1994, and

---

[29] These tend to avoid statements that are too optimistic or too pessimistic. Although another heuristic that should be considered, is to not fully trust results from meta-reviews, given how frequently meta-reviews in the social science which are followed up over time attain different results (which careful meta-reviews, such as those listed above, tend to hedge against by adding cautionary notes).

[30] References removed.

[31] Setälä (2014) also mentions 21st century town meetings. These are like consensus conferences, in that they allow initial self-selection. Depending on how wide one wants the concept of mini-public to be, 21st century town meetings can be included.

citizens' assemblies since 2003/2004. However, despite their age, hundreds of mini-publics have already been completed.

In the following table, I taxonomize these common mini-publics based on salient design features. These design features are commonly discussed in the literature. For instance, Gastil (2017) covers – among other things – many of these design features[32]. Similarly, Smith (2009) discusses most of these design features for the mini-public types below (see also A.2)[33]. My contribution, is systematizing these design features, adding some missing information, and explicitly rating the design features for each type of mini-public.

Table 1. Taxonomy of Mini-Public Types[34]

	Deliberative Polls (James Fishkin)	Citizens' Assemblies (BCCA)	Citizens' Juries (Ned Crosby)	Planning Cells (Peter Dienel)	Consensus Conferences (Danish Board of Technology)
Representative Sample	Medium	Medium	Low	Medium	(very) Low
Expertise Balance	Medium-High	Medium-High	Medium-High	Medium-High	Medium-High
Political Linkage	Medium	Medium	Medium	Medium	Medium
Total Participants	100+	100+-	10-24	100+	10-16
Discussion Group Size	8-18	10-15	10 -24	25	10-16
Norm Enforcement (moderator)	Yes	Yes	Yes	Yes	Yes
Decision-Making Method	Questionnaire	Consensus	Consensus	Moderator Collation	Consensus
Secrecy of Decision	Yes	No	No	No	No
Output	Opinion Change Profiles	Recommendation report	Recommendation report	Recommendation report	Recommendation report
Incentives	$50-300	$150 per meeting day	?	?	?

---

[32] In addition to how narrow the topic for deliberation is (e.g. Gastil mentions that Deliberative Polls often have less rigidly defined topics, or a wider variety of topics), which I have not managed to incorporate into the table.
[33] Smith (2009) discusses design in the categories of inclusiveness, popular control, transparency, efficiency, and transferability.
[34] Those denoted with a "*" are unsure, because there are several sources that give slightly different answers. The differences are not subtantial substantial, however (except for perhaps costs, as acquiring original source material is difficult for costs). Questions-marks at incentives are uncertain values, although all mini-publics seem to get similar stipends (Smith, 2009). Lastly, under the mini-public names in the table above there is a name in brackets who is generally associated with the mini-public/was the first one to propose or use it.

Duration	3 days*	Months	2-5+ days*	2-7 days	Two weekends + 3-4 days*
Public Transparency	Yes	Yes	Yes	Yes	Yes
Teaching	Yes	Yes	Yes	Yes	Yes
Costs	£200.000*	$5.5 million (Canadian)*	£16,000-£30.000*	?	£100.000*

These design features (i.e. row headers) are thus chosen based on what overviews of mini-public design tend to mention, and on what could plausibly affect epistemic success[35]. That does not mean that all relevant design features are incorporated. For instance, a factor that is likely to affect epistemic success, is how politically sensitive the issue is on which is deliberated. The reason that such design features are missing here, is that they tend to vary within each mini-public type, which makes them difficult to rank. Furthermore, many of the categories are crude. For instance, there are many ways of performing norm enforcement, but this table only gives a "yes" or "no" answer.

Some of the evaluations in the table are fairly straightforward, such as factual information about total participants, disccussion group size, norm enforcement by moderators, decision-making method, secrecy of decision, output of the mini-public, incentives (i.e. payment beyond travel and lodging compensation), duration of the mini-public, whether there is teaching involved by experts or staff, and total costs of the mini-public.

And some only need a bit more nuance. For instance, public transparency, for which, even though all mini-publics have some of it, they differ in how they achieve it. Namely, whereas the British Columbia Citizens' Assembly (BCCA) had some audio recordings, online reports, online profiles of participants, and possibilities for media to contact participants, Deliberative Polls sometimes go as far as televising actual deliberation.

However, there are also some choices in the table that required actual judgment. The following three are those that needed most judgment: i) representativeness of sample, ii) expertise balance, and iii) political linkage. I will shortly explain why I made the choices shown in the table.

Firstly, the representativeness of the sample. All mini-publics – except consensus conferences – use unforced random selection. Due to selection being unforced, there is already a substantial degree of self-selection (e.g. of the opinionated). Karjalainen & Rapeli (2015) point towards such evidence for some of the mini-publics (i.e. deliberative polls, citizens' assembly, citizens' juries)[36]. Furthermore, due to the size of compensations, people with high opportunity costs are discouraged from participating. Therefore, the highest score is medium.

---

[35] Although, to be sure, the table could be extended to include agenda-setting potential, internal transparency, and more.

[36] Fishkin (2011-118/19) claims that these issues are not substantial for Deliberative Polls. Similarly, Mansbridge (2010), calls Deliberative Polls the "Gold Standard" of mini-publics. However, if one looks for instance at the Deliberative Poll in Farrar et.al. (2010-347) – like Karjalainen and Rapeli (2015) do – then one could hardly come away with the impression that the educated are only "somewhat more educated" (Fishkin, 2011-119). Perhaps this suggests a difference in standard, however. The standards that are used here are relative to ideal versions of mini-publics, not relative to empirical constraints of current researchers, no matter how well they do within those constraints. Another defense of Fishkin is that, attitudinal representativeness is present (even if there are demographic differences). This will be discussed in (4.5).

Citizens' juries and consensus conferences get even lower scores because of their small sample sizes, which makes the samples less representative. To be sure, they use stratified sampling (i.e. quotas in order to make sure salient social groups are represented), but this only partially reduces the issue with these sample sizes. The reason consensus conferences get an even lower score, is due to initial self-selection (i.e. participants need to apply). For comparison, general politics, which has no similar stratified random sampling, would get an "extremely low" on this rating.

Secondly, expertise balance. Mini-publics tend to let briefing materials and expert lectures be reviewed by peers and stakeholders (in addition to making the material public). As long as the correct stakeholders are identified, and they agree on the content, then expertise is sufficiently balanced. In practice, this might be difficult to achieve, however. Additionally, not all mini-publics have a stakeholder approved moderator, which can have substantial consequences (Humphreys, Masters & Sandbu, 2006; Spada & Vreeland, 2013). Therefore, a medium to high score is given.

Thirdly, political linkage. The reason all mini-publics get medium, is because they only give recommendations. To be sure, Planning Cells and Citizen Juries often use contractual agreement with officials to give a public explanation for why the recommendation was accepted or rejected. And other mini-publics also give recommendations directly to relevant officials; in the BCCA Citizen's Assembly there was even an official (binding) referendum based on the recommendations. But none of these recommendations are binding by themselves; and many do not get accepted in practice.

4.4 Individual Design Features

Now that we have a taxonomy of mini-public types and their individual design features (i.e. row headers in the table), it would be ideal if we knew how much each of these individual design features contribute to making wise decisions[37]. If we knew that, we could take the most effective design features, and build an improved mini-public. Unfortunately, the causal inference paradigm – randomized control trials (RCT) – faces a substantial problem: these mini-publics all come in "package treatments" (i.e. complex interventions), in which inferences can only be made about mini-publics in their entirety, not about individual design-parts.

To explain, the endorsed methodology for causal inference in the literature is that of RCT's[38]. In short, in RCT's, the experiments begin with two identical groups. One receives a treatment – in this case a type of mini-public design – and the other does not (or gets a placebo)[39]. Here, the "other group" – also called the control group – is often the general population. Next, after treatment is administered, any difference in outcome – which in this case is measured by the epistemic criterion "decisions endorsable by relevant experts" – between groups, is then, everything equal, caused by the treatment[40]. Ideally, multiple treatments (i.e. different mini-public designs) are compared, such that we can see which are more effective.

Now, the problem is that for each of the cases in the taxonomy table, the treatment is an entire mini-public; there is no type of mini-public which varies only single design features, keeping the rest

---

[37] And potential synergy effects between design features.
[38] See Esterling (2018).
[39] Different treatments can also be compared, in order to see which is better.
[40] There are conditions that need to be met in actual research. To just name some: i) randomization of sufficiently large groups in order to make it likely that the two groups are sufficiently equal, ii) make sure that the only difference is the treatment effect between groups; this can go wrong for instance if control groups take some treatment anyway or if the control group knows it is assigned to the control group and is therefore disappointed, iii) measurement instruments need to be accurate.

constant. Therefore, individual design features (i.e. row headers) cannot be evaluated with RCT methodology in a straightforward manner[41].

Now, one could switch causal inference paradigms, for instance by using data mining techniques to look for correlations between (clusters of) design features and outcome-success, adding some theory, and then try to make plausible causal inference. The problem with this strategy, however, is that, besides general issues with data mining, the same problems apply as evaluating mini-publics in their entirety, as shown in the next section (i.e. selection-bias, too many uncontrolled variables, publication bias).

To be sure, another strategy could be to look at lab-based studies that *do* manipulate only a single variable, at least that being the purpose. However, there are too few of these studies (i.e. no large meta-reviews possible for specific design features), and they are not without flaw (e.g. self-selection, participants evading treatments to which they have been assigned, external validity), which hinders drawing many relevant conclusions from them (c.f. Esterling, 2018). This thus leaves evaluation of complex interventions in their entirety as the remaining alternative.

4.5 Evaluation of Complex Interventions

So, do inferences based on entire complex mini-public interventions give an answer to our research question? We could try to answer this question by evaluating a particular case. Consider the BCCA example from earlier in the chapter. With respect to duration, the Assembly was one of the longer mini-publics. So, if there is a mini-publics in which we would expect to find learning effects and wise decisions, then this is the one. Fortunately, the conditions for "relevant experts" were met, as the BCCA had stakeholder approved experts, who likely passed the relevance test, due to their narrow domain of expertise[42]. And what these experts said was the following: *[f]ew electoral system experts would dispute the appropriateness of the criteria they considered, few could dispute that their decision was appropriate given the priorities they assigned to the criteria they applied* (Blais, Carty & Fournier, 2008-138/9). This amounts to an endorsement by relevant experts. Additionally, with respect to representativeness, the BCCA did not perform badly, but could have done better (James, 2008). As is quite typical, certain groups – the educated, high-status, and politically involved – were overrepresented. Therefore, overall performance could be called acceptable.

4.5.1 Issues

Unfortunately, that does not allow us to conclude that the Citizens' Assembly design is a good "package treatment" (i.e. complex intervention). The reasons for this, are threefold: self-selection, too many uncontrolled variables, and publication bias.

Firstly, there is the issues of self-selection. It is not only problematic because it makes mini-publics less representative, thereby not satisfying the normative criterion of "representativeness". The issue is also that self-selection makes the treatment group (i.e. participants in the mini-public) unequal to the control group (i.e. the general population). This is a problem, because unequal starting groups makes us unable to know whether the observed difference in outcome is caused by the treatment or by a difference in groups[43].

The problem is quite severe too, as "deliberation experiments generally must struggle with the issue of self-selection in a way that is often more severe than is found in survey research" (Esterling, 2018).

---

[41] That said, one could try to look for correlations between appearances of an individual design

[42] Although I am aware that there are differences between mathematicians and political scientists in what they think about electoral systems.

[43] Also called an "internal validity" problem.

People sometimes lead busy lives, and cannot afford to spend the time and effort. Others simply do not care. Karjalainen and Rapeli (2015-409), even find it "surprising how little attention the predictors of attrition have gained in recent scholarship in this particular field".

Now, it needs to be noted that representativeness in the literature comes in two versions: demographic and attitudinal. Fishkin (2011) claims that attitudinal representativeness is often overlooked. However, the distinction does change judgment about self-selection here. Namely, Fishkin claims that Delibertive Polls are highly representative on both dimensions[44]. Yet, as mentioned in (4.3) and footnote 36, whether one agrees depends on one's standards. On an absolute level, we saw that demographics were substantially far from fully representative. This is true even if relative to practical constraints, the results are impressive.

Furthermore, the importance of attitudinal differences are more pressing than is often assumed. To illustrate, it has been argued that if one slices and dices demographics carefully enough, and subsequently analyzes the data properly, then political issue attitudes do differ substantially on demographics that are frequently improperly controlled for (Weeden & Kurzban, 2017). This is not to discuss the arguments made for this particular claim, but it is to place at least a cautionary note on any claims about representativeness. Therefore, self-selection is a sufficiently serious issue for inference about package treatments.

Secondly, there are too many uncontrolled variables. Factors that potentially influence deliberation quality cover the entire social sciences: personality traits, intelligence, group composition, group size, discussion topic, expertise balance, moderator behavior, decision-rules, structure of discussions, sequence of events, expected impact of decisions, auditing institutions, publicity, monetary incentives, appetite, smell, etc. Controlling for all of these variables is not feasible[45]. A downside of this for evaluating variables, is that one cannot be sure that the mini-public design is causing the change, or one of these factors that vary across mini-publics which are not properly controlled for.

Thirdly, there is the problem of publication bias[46]. Generally, the idea of studying randomized control trials, is that single trials get replicated, and that all results get pooled together into meta-reviews. Such practice allows a community to be more sure that the treatment is driving the result, and not some contingencies in particular experiments. However, this practice does not function properly if not enough attention is paid to negative results (e.g. Ravazzi & Pomatto, 2014). Unfortunately, Spada and Ryan (2017) argue that this is what happens to evaluation of democratic innovations. The support for their claim partially comes from empirical journal publications about deliberation. They coded the publications from 2006-2016 in the top five political science journals, with the additions of the *Journal of Public Deliberation* and *Politics and Society*. The outcome is as follows (ibid-773):

1. Learning from best practices: 97 (64.2%)
2. Learning from the variety of quality of practices: 28 (18.5%)
3. Learning from failures: 7 (4.6%)
4. Other (e.g., Quality of deliberation metrics, mapping, surveys about citizens' propensity to deliberate): 19 (12.6%)

None of these "learning from failures" publications were in the top journals.

---

[44] Fishkin, in the Blackwell Guide to Social and Political Philosophy, says (pg. 232): "in every case thus far, the weekend microcosm has been highly representative, both attitudinally and demographically, as compared with the entire baseline survey and with census data about the population".

[45] This is recognized in the literature (see e.g. Gastil, 2018).

[46] See also Wuttke (2019), for some general scepticism about findings in political science.

One explanation for these findings, is that there are not many failures. The authors do not believe this to be the case however, and refer to the presence of various specialized monographs giving testimony to deliberative failures. The difference is that the results described in these monographs have not been published in the relevant journals.

Now, even if learning from best practices is a legitimate research strategy, publication bias, as mentioned above, does hinder systemic evaluation of a type of mini-public.

To conclude, evaluating these complex interventions is difficult. Perhaps the reason BCCA performed well on the criterion "decisions endorsable by relevant experts", is because of the overrepresentation of the educated (i.e. self-selection). Or because voting systems lack emotional salience (i.e. uncontrolled variable). Or perhaps even because we selected a cherry from the literature (i.e. publication bias), which in fact belongs to a random distribution of epistemic success (i.e. chance). Problems such as these, make it difficult to know which type of mini-public makes wise decisions.

So, given that we do not get an answer to the question of whether mini-publics can make wise decisions based on the empirics of individual design features, and on complex interventions, perhaps a new strategy is needed that focuses more on theoretical arguments. This is what the next chapter will do with respect to average human reason, by inquiring whether such reason is sufficient to reliably make wise decisions in mini-publics, or whether it is the root cause of empirical failures.

5. Human Reasoning

This chapter reviews whether average human reason prevents mini-publics from making wise decisions. My claim is that – even though deficiencies in human reasoning are a substantial challenge – human reasoning capacity is in principle sufficient to make wise decisions in mini-publics.

The support for this claim can be categorized into two separate parts. Firstly, I review the case for thinking that "the mind is to blame" for mini-public failures: i) voters are cognitively biased, and this is unlikely to change due to ii) lack of good incentive schemes, iii) robustness of cognitive biases, and iv) low promise of education. I argue that this case is inconclusive, despite agreeing that a proper account of incentive schemes are indeed lacking, as education likely has more potential than portrayed by these critics. Secondly, I review the case for believing "the information of the agent is to blame" for mini-public failure: a) the argumentative theory of reasoning, which shows why human reason is sufficient when put in its proper environment, and b) unwise decisions can be explained by rational agents finding themselves in information cascades. These theories taken jointly, should be sufficient to show that, in principle, human reason is not an impregnable obstacle to mini-publics making wise decisions.

However, before making these specific arguments, I shall briefly introduce reasons for why we should expect a lack of political expertise, and whether (and how) the literature thinks this problem can be solved for mini-publics.

5.1 Voters' Lack of Expertise

Concerned that the political domain insufficiently incentivizes voters to learn, Caplan (2007) writes:

> Political behavior seems weird because the incentives that voters face are weird….
> Economists should never have expected political behavior to parallel market behavior in the
> first place (ibid-141).

A recent overview paper in *behavioral political economy,* a field combining behavioral economics and political science, echoes these concerns:

> in the political arena, individual feedback and learning mechanisms are much weaker and more indirect than in the marketplace. Thus, we can expect cognitive biases to be at least as, and probably more important in politics than in the marketplace (Schnellenbach & Schubert, 2015)

So, as claimed in the quotes above, the political arena has weaker incentives and weaker learning mechanisms, which lead to greater cognitive bias.

To clarify, the claim here is not that agents in the marketplace are free from bias. Neither is the claim that specific markets in which agents show substantial bias do not exist. Rather, the claim is that bias is *generally* larger in the political domain.

A very similar story can also be found in Kahneman and Klein (2009), two giants in the field of decision theory, coming together to resolve their differences about intuitive expertise. Kahneman is the champion of intuition failure (e.g. professional prediction failure), and Klein is the champion of intuition success (e.g. firefighters' danger instinct). Despite their appearing differences, they came to "a failure to disagree" about the status of intuitive expertise. Their joint claim is that expert intuition develops under the following conditions:

> Two conditions must be satisfied for skilled intuition to develop: an environment of sufficiently high validity and adequate opportunity to practice the skill. (ibid-520)

Loosely, "high validity" means that the environment is sufficiently regular/stable, such that situational cues can be recognized. And "adequate opportunity to practice" means sufficient exposure to such an environment.

To illustrate, consider the difference between chess and economic forecasting. Chess masters have ample time to practice in an environment in which the rules do not change. Economic forecasters do not have the same luxury. These contrasting conditions, is what explains most of the difference in intuitive expertise in these domains[47].

Now, much of voters' decisions will be based on their intuition. And these intuitions are unlikely to be expert. Namely, the target environment – everything important to a political vote – is unlikely to be complete, as voters typically do not engage with the full spectrum of information and opinions. And neither do voters have ample time to practice in a good environment, as political discussion is neither well-practiced nor well-liked.

One prediction that follows from this, is that voters will likely lack knowledge about political issues. And this is indeed what political scientists find (c.f. Converse, 2006 [1964]; Achen & Bartels, 2017; Somin, 2016). Voters often do not even remotely know how much government spends on what sectors, what the main political ideologies stand for, and where politicians stand on relevant issues (even politicians they vote on). These findings have also been surprisingly robust.

However, commentators disagree on whether such lack of knowledge and expert intuition is problematic. Some remain cheerful. They claim that voters use simple heuristics that produce

---

[47] To be sure, there are other factors – such as talent and motivation – that matter, but these factors will be better at explaining within domain variance than across domain variance.

normatively sufficient choices (c.f. Lupia, McCubbins & Arthur, 1998[48]). Or they claim that mistaken answers are randomly distributed, such that, under certain conditions, the law of large numbers will ensure that aggregate answers are sufficient (c.f. Landemore 2012).

Yet many commentators also lament voters' lack of knowledge and expert intuition; among them deliberative democrats and epistocrats (i.e. those favoring rule by those with knowledge). They tend to think that heuristic voting is open to manipulation and irrelevancies, or argue that mistakes are systemically biased, nullifying any positive effect the law of large numbers has. Here I am going to assume that these commentators are correct, at least with respect to mini-publics.

5.2 What is to Blame?

The aforementioned lack of knowledge and intuitive expertise, do not simply go away in mini-publics. Unless, of course, these mini-publics provide the right environment, with sufficient time to learn from that environment; much like Kahneman and Klein would have it for creating intuitive expertise. However, it could also be that mini-publics are unable to provide such an environment, and that cognitive deficiencies are too much of a problem to overcome.

Another way to frame such a debate, is to ask who (what) is responsible (attributable) for mini-public's lack of wise decisions. Caplan (2007-97) – who's focus is politics more generally – distinguishes two of these options[49]:

1. blame the mind of the agent
2. blame the information of the agent

Now, if one blames the information of the agent, then the solution can be as simple as providing an environment with information to the agent; perhaps supplemented with a motivating incentive scheme. However, if you blame the mind of the agent, then the solution can be as severe as seeking exclusion of these lesser minds in the political process; which includes giving up hope on mini-publics being able to make wise decisions.

The second option – blaming the information of the agent -- is the option endorsed by most deliberative democrats. The majority accepts this, because general citizen cognition is seen as already sufficient (Chambers, 2009-330). One will see "learning phases" in mini-publics, but these are related to content on which must be deliberated (c.f Pearse, 2008), not skills to evaluate that content[50]. The mind therefore does not need to be blamed; the solution is to inform agents and to let them use their sufficient cognitive capacities to deliberate and reflect[51].

Other deliberative democrats agree, but are less positive about the mind. They think that individual cognition is unlikely to improve (Mercier & Landemore, 2012-254; Goodin & Spiekermann, 2018-94/5[52]). The solution is therefore to not focus too much on teaching agents, but instead to construct

---

[48] For an example, although not plain optimistic, see the overview of retrospective voting (Healy & Malhotra, 2013), which is a voting theory analogous to using heuristics like 'throw out poor performers'.

[49] Caplan, an economist, uses this distinction partially due to prominence in economics of rational expectations models (blaming information) and behavioral economics (frequently blaming the mind). Another thing to note, is that I use the reverse order of listing them compared to Caplan.

[50] Note that I am here adopting a narrow definition of information, such that cognitive evaluation skills are different from information (i.e. cognitive evaluation skills are not a subset of information).

[51] There are some deliberative democrats that see room for long-term improvements though: "publicly supported and publicly accredited schools should teach future citizens the knowledge and skills needed for democratic deliberation" (Gutman & Thompson, 2004-35).

[52] They are somewhat more hopeful of increases in individual competence due to practical experience and deliberation.

an environment in which the mind can function properly. A better characterization here would therefore be a*) blame the institutional environment that elicits voters' answers. To be sure, most deliberative democrats would likely insist on a combination of (a) and (a*). For the purposes here, however, I will simply lump them together.

That said, a sharper distinction can be drawn between those opting for blaming the information of the agent, and those that blame the mind of the agent. Many epistocrats (i.e. those favoring rule by knowledge) endorse this second option. Their claim is that the mind is flawed and unlikely to improve.

The rest of this chapter reviews the case for both of these sides. Should we blame the mind of the agent? Or should we blame the information of the agent (or its institutional environment)?

5.3 Blame the Mind

This section discusses the arguments in favor of blaming the mind of the agent. In short, the idea is that voters' minds tend to be biased, and that this is unlikely to improve.

5.3.1 Voters' Minds are Biased…

Mini-public designers are aware that deliberation can fail. For example, Fishkin (2011) engages with two distortions – domination of discussion by those with high socio-economic status, and group polarization (Sunstein, 2000) – and tries to argue that his type of mini-public is robust to both issues.

However, critics tend to have a much longer list of distorting factors (c.f Caplan, 2007). Additionally, these lists are not just naïve enumerations of cognitive biases, postulated to apply to politics. Namely, substantial empirical evidence supports the presence of many of these biases in politics; frequently in exacerbated form. Consider the following compiled list of Brennan's (2017) second chapter (sources show a political application):

- In-group/out-group bias (Cohen, 2003; Mutz, 2006)
- Confirmation bias and disconfirmation bias[53] (Kahan et.al. 2017)
- Availability bias
- Affective contagion and prior attitude effect (Erisen, Lodge & Taber, 2014)
- Framing effects (Kelly, 2012)
- Peer pressure and authority

To illustrate these biases, consider a hypothetical example of partisan members of the political spectrum. Both those on the left and right feel that justice is on their side, and banner together to form close-knit coalitions, in which one of their missions is to triump over their evil opposition (i.e. in-group/out-group bias). Both groups will interpret every new point of data that they can find in a favorable light (i.e. confirmation bias). Yet what they will remember and deem important, is the vivid memory of how the *other* group bungled their statistics (i.e. availability bias). An immediate strong sensation will prime partisans of their opponent's errors, and they will not be able to rid themselves of the idea (i.e. affective contagion). Unable to use shared language anymore, the left now refers to "lives lost in the war", whereas the right refers to "potential lives saved in the war" (i.e. framing effects). Some partisans are unsure about their group's actions, but that will quickly be forgotten, as the majority will briskly step in to ensure that dissenters fall in line (i.e. peer pressure).

Now, even though the illustration is about strong partisans, the claim is that these biases are common to all people. These biases are then brought into any deliberation, and will therefore heavily distort the process. To illustrate, Brennan (2017) describes what he thinks happens when cognitively

---

[53] Also called my-side bias.

biased members of different tribes – which he think constitute a substantial part of the population – come together:

> Hooligans would try to dominate the discussion. They would ignore, jeer at, and dismiss one another during disagreements. They would insult one another, or at least mutter insults under their breath. Hooligans would fail to both offer reasons for their views and accept others' reasons, even when they should. They would happily manipulate one another, use language in a deceitful way to confuse people, and lie, if doing so helps their side. In the face of contrary evidence, hooligans will just dig in their heels and get angry. When hooligans deliberate, the "force of the better argument" is impotent. What matters are rhetoric, sex appeal, and promoting the team. When hooligans deliberate, they get worse. (ibid-61)

To support this position, Brennan cites empirical work on deliberative democracy (e.g. Mendelberg, 2002; Pincock, 2012). However, as mentioned before, the empirical results are perhaps better interpreted as mixed[54]. That said, even assuming empirical results are mixed, some of the points remain relevant: the behavior illustrated above has been shown and can occur again; and so it is for the associated biases.

5.3.2 And Voter's Minds are Also Unlikely to Improve
The previous section showed the presence of cognitive bias in politics. However, this leaves the possibility that a well-constructed informative environment makes these biases disappear. Therefore, fully consistent with their claims, some critics of mini-publics (i.e. epistocrats) argue that these biases are unlikely to be alleviated.

For instance, Caplan (2007) argues that bias is a stable feature of the voter's mind, because agents are *rationally irrational*: agents maximize their utility – which is rational – by holding false epistemic beliefs; which is irrational. Put differently, holding pleasant but false beliefs for consumption value, is rational from an individual perspective. The idea is that benefits of holding false beliefs, is simply higher than the costs. Costs are low, because the average voter will never be punished for holding false political beliefs. And benefits are high, because it is pleasant to hold false beliefs. Who would not want to believe that your team is righteous, or to believe that you are an ill-understood hero that is on the verge of greatness, even though both might be questionable from an objective point of view?

Complementarily, Somin (2016) further stresses the prevalence of *rational ignorance*: a theory that states that citizens will remain ignorant because gaining political knowledge is too costly (Downs, 1957). Costs are high (i.e. effort to gain knowledge) and benefits are low (i.e. chance of being the decisive vote). Therefore, voters will likely maximize – which is rational – by remaining ignorant.

Now, the idea is that, if agents are rationally irrational and rationally ignorant, then substantial incentive schemes are needed to make it even plausible for agents to choose holding or acquiring true beliefs. Critics of mini-publics are sceptical about such an incentive scheme existing.

Unfortunately, not many incentive schemes for mini-publics exist, which makes it difficult to evaluate these hunches. However, to give one example, consider Ackerman and Fishkin's (2002) "deliberation day" proposal. They suggest the following incentive scheme: "*each deliberator will be paid $150 for the day's work of citizenship, on condition that he or she shows up at the polls the next week*" (ibid-129). Somin (2016) rejects this incentive scheme, because he thinks combatting rational ignorance

---

[54] It seems that the more literal Brennan agrees: "that said, the research on deliberative polling shows promise. It may be able to overcome many of the problems of mass participation and mass democracy" (ibid-67).

will need much more education than several days or weeks, which in turn requires a huge increase in the costs of the incentive scheme[55].

However, besides this objection, there are also more fundamental reasons to reject the incentive scheme. For instance, the scheme is not an incentive to *learn*, but an incentive to *show up*. Participants can simply show up and learn nothing. Agents who are rationally irrational and rationally ignorant — in the way described above – could take advantage of that. Therefore, what is needed instead, is a scheme that either raises benefits of learning (e.g. knowledge spillover effects to the labor domain, intrinsic motivation, social prestige), or a sufficiently reliable detection mechanism that rewards only those that put in effort. Such an account has not been given, however. This could be seen as a problem for deliberative democrats[56].

Lastly, there is an argument from Caplan (2018). He claims that education (secondary and tertiary level) is much less succesful than commonly thought. The reason is that formal education is primarily about signalling to future employers that one is intelligent, conscientious, and willing to conform; not about being educated. The support for this claim comes, for instance, from literature showing what little knowledge is retained from education, and data calculations that are supposed to show that learning can explain little of future income gains. Some might wonder how an equilibrium could exist in which students learn almost nothing. Caplan claims that it is possible for several reasons: most school-knowledge is also not valuable to the employer, vocational education is still valuable, and formal education reliably signals a student's qualities[57].

Therefore, students who have the desirable qualities simply have to jump through the hoops, signalling to the employer that they are suitable for the better jobs. Caplan sees this equilibrium as unlikely to change. Now, this does not mean that he believes that none of the students become educated; there are tangible cases in which they do. It is just that these are the minority of students who are naturally motivated, interested, or gifted in their area.

Supposing these claims about education are true, that would cast further doubt on whether agents can be informed sufficiently in deliberative environments. I am not going to fully evaluate the argument here. What I will note, however, is that even if Caplan is right about education *in general* being mostly signalling – which is controversial – then that does not necessarily imply much for the *particular* education in mini-publics. Namely, mini-publics differ in important ways from these results in formal education. Some examples are as follows:

i) Information in mini-publics needs to only be retained relatively short-term; not long-term, which the incriminating evidence suggests students are bad at.

ii) Practical political issues and discussions are likely to engage more than standard theoretical lectures, especially if competing experts do their best to convince deliberators.

---

[55] He also rejects the proposal due to risk of political capture of for instance source material and agenda setting.

[56] Although theories in the next section are largely unconcerned about monetary incentives, as they think that designing the institutional environment is sufficient.

[57] In signalling models there can be a separating equilibrium, in which those with the desirable qualities (i.e. intelligence, conscientiousness and willingness to conform) have to perform several years of effortful signalling – in which they learn nothing – in order to make it too costly to perform these actions for those without those qualities. Any less effort/time on part of the high-quality students would make it advantageous to low-quality students to also start jumping through effortful hoops. The advantage is being able to get a better job because employers think better of you on average.

iii)   Mini-publics can use financial incentive schemes, which – unlike formal education – give immediate rewards.

iv)   Even if Caplan is right about *current* education, that need not imply much about *future* education. For instance, some still have high hopes for explicit teaching of critical thinking (Halpern, 2013), and some promote evidence that, despite expectations, de-biasing the mind is possible even with short-term interventions (Morewedge et.al. 2015). To be sure, Caplan would likely be sceptical, but it seems premature to conclude that projects such as these are likely to fail.

The argument here is thus, that differences between these learning environments are sufficiently large to merit the claim that results about formal education are not generalizable to mini-publics.

To conclude, this section has discussed the presence of cognitive biases (e.g in-group out-group bias, confirmation bias) and what that implies for mini-publics' potential. Several arguments were considered that support the claim that this potential is low. These were as follows: both rational irrationality and rational ignorance make agents unlikely to learn, incentive schemes to combat these problems do not exist, and formal education has proven bad at educating, and is unlikely to improve. In response, the incentive argument was made stronger, by showing how necessary incentive schemes do not yet remotely exist. And by contrast, the education argument was weakened, by showing how not all results apply to mini-publics, and by showing future potential of education.

5.4 Blame the Information

This section reviews the case for blaming the information of the agent (or the institutional environment). My claim is that the information/environment is indeed to blame, since a proper argumentative environment, and an adequately controlled flow of information, are sufficient to conclude that humans can in principle make wise decisions in mini-publics. Moreover, well-functioning incentive schemes – even though they might contribute – are not necessary for wise decisions.

5.4.1 New Psychology of Reasoning

Psychology of reasoning used to be dominated by classical logic. Reasoning was seen as "*based on binary distinctions—between truth/falsity, consistency/inconsistency, and validity/invalidity—and focused primarily on drawing inferences from arbitrary assumptions according to the rules of extensional textbook logic*" (Elqayam & Over, 2013-249-250). These traditions have retrospectively received labels like *logicism,* the deduction paradigm, the binary paradigm, and the traditional paradigm (Elqayam, 2017). The tenability of this paradigm has proven doubtful. Mid-twentieth century showed some of the first experimental findings suggesting that logical reasoning is not a good *descriptive* theory. Nowadays, some even doubt whether logical reasoning is a good *normative* theory (ibid). Examples of normative doubts, are the relative absence of probability, degrees of belief, and contextual information.

These issues have led to a "new paradigm of psychology of reasoning" (ibid). Other frequently used labels for this paradigm are probabilistic, Bayesian, and decision-theoretic[58]. The reason for these labels, is that much of the new paradigm embraces probability and utility as core features of human reasoning.

To see an application of the new paradigm, consider the following example:

---

[58] These traditions have existed for longer, but Elqayam and Over (2013) show that only *after* the 1990s "probability" became frequently mentioned in the psychology of reasoning literature.

Slippery Slope Arguments Reloaded: A slippery slope argument (SSA) asserts that relatively small first steps will cause a chain of events leading to large (negative) effects. An example is: "*if voluntary euthanasia is legalised, then in the future there will be more cases of 'medical murder'*" (Corner, Hahn & Oaksford, 2011-133). SSA's have a bad reputation among logicians and philosophers, who tend to dub them as informal fallacies. But new paradigm reasoning – applying subjective probability and utility – casts SSA's in a different light. Namely, no longer are SSA's categorically rejected. Instead, they are admitted based on context. Namely, people tend to find SSA's more convincing when either they i) evaluate the outcome-event (i.e. medical murder) to likely follow the conditional (i.e. legalization of euthanasia), or when they ii) dislike the outcome (i.e. medical murder). The SSA is thus accepted not based on standard logical reasons, but instead on probabilistic plausibility, and on how much the agent gains from accepting the argument.

Such a pragmatic account of reasoning is a large departure from the traditional paradigm[59].

5.4.2 Argumentative Theory of Reasoning
The main focus here, is on a particular case of this new paradigm: *the argumentative theory of reasoning*. Similarly to the story above, psychology of argumentation used to be focused on logical reasoning. For instance, Piaget (1928-204), an influential child development psychologist, claimed that "*logical reasoning is an argument which we have with ourselves, and which reproduces internally the features of a real argument*". Reasoning was thus both logical and solitary.

In recent influential work, Mercier and Sperber (2011; 2017) have not just questioned the classical logic approach of the traditional reasoning paradigm, but also its individualism. They call their theory "an argumentative theory of reasoning" and an "interactionist model". Mercier (2018-401/2) gives the following account (text original, compiling and listing mine):

1. The argumentative theory relies on the framework of the evolution of communication.
2. For communication to be evolutionarily stable, it has to benefit both senders and receivers.
3. Humans rely on mechanisms of epistemic vigilance that evaluate communicated information to reject harmful messages and accept beneficial ones.
4. Two important mechanisms of epistemic vigilance are plausibility checking and trust calibration.
5. Both mechanisms – plausibility checking and trust calibration – ought to be conservative.
6. To achieve a finer-grained discrimination of messages, senders and receivers can rely on argumentation.
7. Reasoning would have evolved mainly to enable such argumentation: to allow senders to find arguments supporting their messages, and to allow receivers to evaluate these arguments.

So, people communicate to each other, partially in order to persuade. To avoid being gullible, however, people do not simply trust each other's claims, but rely on argumentation. For senders, the message gets more persuasive when arguments are accumulated for support. And given that receivers want to avoid being gullible followers, they evaluate these arguments.

Mercier and Sperber champion their theory for several reasons: its consistency with their intuitions about evolutionary plausibility, for the predictions it generates, and for the large variety of empirical evidence that confirms these predictions. Some of these predictions are as follows.

---

[59] To be sure, the new paradigm does not discard everything from the old paradigm, and there are prominent accounts that stay fairly close the the traditional paradigm, such as mental models, in which the mind simulates representational possibilities of the world, about which logical inferences are made.

      I.   Isolated Individual Reasoning
            a.   Agents will tend to accumulate arguments that support their message, leading to my-side bias (i.e. confirmation and disconfirmation bias).
            b.   Agents will be better at evaluating arguments made by others than arguments made by themselves.
     II.   Social Interaction Reasoning
            a.   Argumentative dialogic contexts should give better epistemic results than aggregated individual reasoning, as isolated individuals do not feel the same urge to justify their arguments, and therefore stick closer to their initial generic argument.

Lastly, for both the isolated individual and the social interactions, these features are difficult to change due to their evolutionary origin[60].

However, if the argumentative theory of reasoning is correct, then why are the empirical results of mini-publics — an ideal case of dialogic context with argumentation – so mixed? Mercier and Landemore (2012) come up with a prominent explanation. Their claim is that evidence shows that bad results are caused by bad deliberative environments. Examples of bad environments they give are i) isolated individual reasoning; causing my-side bias and satisfaction with generic arguments, and ii) reasoning with like-minded people (causing group polarization). They also claim that easy design solutions are available. The first problem, isolated individual reasoning – which is hard to avoid entirely — can be improved by public accountability; "*having to defend one's opinion in front of an audience*" (ibid -254). This would make isolated reason ensure to prepare for public justification. The second problem, reasoning with like-minded people, can be improved by skipping discussion, or by introducing a devil's advocate[61]. Now, if such design solutions are implemented, then deliberation "*should yield good results*" (ibid-250).

As the argument stands, however, one might wonder whether "public justification of one's opinion" can explain the difference between findings of good and bad mini-public results. Especially because ideals of deliberative democracy have included publicity and accountability from the start, and because actual mini-publics tend to take these factors into account. So perhaps we must tailor our conclusion somewhat. One option is to conclude that public justification contributes to good performance, but that, on its own, it is not sufficient for good performance. Another option is to conclude that actual mini-publics should more fully exploit the potential of public justification. For instance, some mini-publics have elements of public justification, such as regular group deliberation, spokespersons summarizing group deliberations, public recommendation reports, and optional media interaction. But these tactics could be made more extreme, such as requiring each individual to publicly justify their decision in front of a live audience. These options seem plausible given the argumentative theory of reasoning.

Moreover, recommendations for mini-publics, based on the argumentative theory of reasoning, can be extended. For instance, one of the predictions above states that 'agents will be better at evaluating arguments made by others than arguments made by themselves'. A design implication that follows is that we should let different people identify reasons and arguments than those who evaluate them. One way to do this, is by letting experts come up with reasons and arguments, and letting the mini-public evaluate them. And another way to do this, is by letting small groups come up

---

[60] This could also imply that incentive schemes will not perform as well as expected.
[61] It should be noted that reasoning with like-minded people is not as likely to occur in mini-publics, given that they are supposed to be representative of the population.

with arguments, exchange these arguments with other groups, and then letting groups evaluate these exchanged arguments[62].

5.4.3 Information Phenomena

Previously, we have assumed that an obvious explanation for systemically unwise answers, is that agents' minds are biased. But that need not be so. An alternative explanation put forward here, is that of information processing. Or rather, "*socio-epistemic information phenomena resulting from processes that track truth imperfectly" (Hansen, Hendrick & Rendsvig, 2013-301).* These are phenomena in which epistemically rational individuals can converge on wrong answers, due to the way in which information is processed.

Consider for example information cascades:

> **Information cascades**: a phenomenon in which early public information has disproportionate influence on later public and private information. The mechanism by which these cascades occur is social proof[63].

One feature of this phenomenon is that public information accumulates over time, as private information is revealed sequentially. This private information, however, can be discarded at some point in the sequence, in favor of public consensus. To intuitively grasp why, consider a rationale an agent might use, such as 'if the previous four people think that life is found on Mars, then that is probably correct, even though I initially thought otherwise. Or 'If I were pressed, I would guess that there is life on Mars; otherwise chances are that I am wrong/not win a reward/look foolish'. These rationales can lead to cascade effects.

Such information cascades are a good reason to carefully consider design in cases with sequential release of information. For an example of such design, consider a group of judges. When judges need to reveal their opinion sequentially, it is best that the process occurs in order of ascending formal rank. Otherwise, lower ranked judges might discard their private information in favor of higher ranked opinion, in order not to look foolish[64].

These information cascades can also be an issue for mini-publics. For instance, early speakers – who tend to be more radical and of higher socio-economic status – can have disproportionate influence on subsequent deliberation. Similarly so for early voters. These issues have been a major concern for deliberative democrats. Fortunately, the design lessons from the previous paragraphs can also be applied here. For instance, one could randomize speaking order to avoid certain speakers from starting each discussion. Or one could make deliberators cast a preliminary vote on the issue, of which the results will be made public at a later stage. This would ensure that early speakers cannot cascade a discussion[65]. The hypothesis here, is thus that such design features can partially explain the difference between good and bad epistemic deliberation in mini-publics.

5.5 Conclusions

Those that blame the mind for failure in mini-public performance, draw from the cognitive bias literature and its applications in political science. Agency models (i.e. rational ignorance, rational

---

[62] This is similar to the recommendation to separate the *identification* stage from the *selection* stage (Sunstein & Hastie, 2015; chapter 7).
[63] For a more precise game theoretic account of this phenomenon see Hansen, Hendrick and Rendsvig (2013).
[64] See Suntein & Hastie (2015) for many other examples.
[65] Interestingly, some of these recommendations already exist in the business literature (e.g. in the form of the Delphi Method).

irrationality) show that substantial incentive schemes are needed to make agents change their behaviors. I discussed the lack of such incentive schemes for mini-publics. One example – Ackerman and Fishkin's Deliberation Day proposal – is a scheme that rewards showing up; not actual learning. The agency models above will take advantage of that. Better incentive schemes are therefore needed. Besides incentive schemes, pessimists are also suspicious of formal education. The claim is that formal education is mostly signalling, not actual learning, and that this is unlikely to change due to being a fairly stable equilibrium. In response to this claim, I argued that education for mini-publics has more potential than can be inferred from a general impression of current formal education. Assuming this to be true, the main worry for optimists seems to be a lack of well-designed incentive schemes that is able to combat rational irrationality and rational ignorance.

By contrast, those that blame the information of the agent or the deliberative environment, do not focus on incentive schemes. They think that putting reason in its proper environment is already sufficient to improve mini-publics. Support for this claim comes from two places. Firstly, the argumentative theory of reasoning, which predicts that applying its design recommendations (e.g. public accountability) will remove ailments of cognitive bias, and thereby ensure wise mini-public decisions. And secondly, the theory of information phenomena, which shows how rational individuals can converge on unwise decisions due to bad information sequences, and how design solutions (e.g. sequencing information) can solve these problems. Taken together, these theories show that – everything equal – both cognitively biased and rational agents, can make wise decisions. Human reason is therefore not the obstacle.

6. Mini-Publics: Design

This chapter reviews whether mini-publics which are designed to alleviate reasoning imperfections, can be coherent, without unintended epistemic downsides, and without any glaring normative issues. My claim shall be that "yes, this is possible". And the conclusion shall therefore be that "yes, mini-publics can make wise decisions".

The chapter comes to these conclusions as follows. Since we are interested in mini-publics making wise decisions, the chapter is split into two sections: *representativeness* and *wise decisions*. I argue that there are several design solutions for *representativeness* (i.e. forced participation, opportunity costs compensation, social status improvement, politically binding decisions), but that these all have their respective issues. One could either bite the bullet, or accept that mini-publics will not be fully representative. However, at the very least, not all of these design solutions have glaring *normative* issues, which is the constraint that needed to be satisfied.

Next, I also discuss several design solutions for *wise decisions* (i.e. public justification, speaking order randomization, pre-deliberation votes, identification and evaluation separation, politically binding decisions, group incentives, prolonged duration). Despite some of these design features having possible issues, many of them can be deflated either because these issues are offset by other design solutions, or because the issues are not too severe.

6.1 Representativeness
This section reviews design solutions for *representativeness*. In short, the strategy is to look at motivations, and seeing whether higher motivation can be achieved through design.

6.1.1 Incentives to Accept Invitations
Currently, motivations provided by mini-publics are as follows:

Monetary	Non-Monetary[66]
Direct (money)	Self-Respect
Indirect (human capital gains)	Duty-Fulfillment
	Social Engagement
	Image Concerns
	Intrinsic Motivation

Current non-monetary rewards are insufficiently motivating. By how much exactly is unsure, as non-monetary rewards are difficult to quantify. However, what we do know from the empirical evidence, is that these motivations are insufficient to gain even close to full acceptance of invitations. Conceivably, this could change. For instance, if mini-publics were more publically known, then image concerns would be more substantial. And if decisions in mini-publics had more political influence, then there could be a higher sense of duty and social engagement. This could be achieved, for instance, by making sure that mini-public decisions are politically binding, instead of the advisory role they have now. However, all of these changes are difficult to achieve in practice. Furthermore, more influence and publicity could also lead to shy and anxious people declining to participate (i.e. self-selection), which would result in lower representativeness.

Similarly, for monetary incentives, human capital gains – in the form of acquiring new knowledge and relationships – have also proven to be insufficiently motivating. And the same applies to direct monetary incentives.

Out of these motivations, the easiest to manipulate are direct monetary incentives. Now, in economics, in order to make a person accept a certain contract, it is common to pay at least the following:

> **Opportunity Costs:** costs incurred from not gaining the (net) benefits from the best alternative choice

These benefits ecompass enjoying leisure or wages received from work. Now, as appendix (A3) shows, paying for opportunity costs seems feasible; current mini-publics are not too far off from paying opportunity costs, and total costs of mini-publics will not rise by an order of magnitude.

So, given that paying opportunity costs is feasible, is it also desirable? One major concern, is whether paying opportunity costs is fair. For instance, presumably, joining a mini-public costs equal amounts of effort for every participant. Why should some get paid more for that effort than others?[67] Intuitively, that is unfair, and seems to go against the spirit of political equality that is core to deliberative democracy. Additionally, besides the matter of whether it *is* unfair, it can also be *perceived* as unfair. And if this is so, cooperative behavior might break down, which will likely lead to less wise decisions.

Moreover, we also cannot be sure that rewarding people with monetary incentives will make them perform better according to the standards we set. At least, that is the main message of the *crowding out* literature, which emphasizes the decrease of intrinsic motivations due to extrinsic motivations

---

[66] Some of these are difficult to separate. For instance, intrinsic motivations and doing some out of sense of duty seem very similar (intrinsic motivations being the umbrella-term), as duty is not done for external rewards or prodding (unless duty is seen as coming from an external source).

[67] This refers to fairness as desert. Given that there are stronger theories of fairness than desert, failing to satisfy this theory implies likely failure from the perspective of other theories too.

such as money (c.f. Bowles, 2016). To clarify, the following is the standard definition of intrinsic motivation:

> Intrinsic motivation is defined as the doing of an activity for its inherent satisfactions rather than for some separable consequence. When intrinsically motivated a person is moved to act for the fun or challenge entailed rather than because of external prods, pressures, or rewards (Ryan & Deci, 2000-56)

So, given these intrinsic motivations, one could accept a mini-public invitation for the inherent satisfaction of the experience. Or consider a closely related non-monetary motivation: duty-fulfillment. One might accept a mini-public invitation out of a sense of civic duty, but would refuse if the invitation was framed as a good financial offer; one can decline a financial offer without repercussions to one's virtue, but not so for a demand on one's sense of civic duty.

Additionally, similar to crowding out intrinsic motivations, there is a case to be made for crowding out image-concerns. Namely, by providing monetary incentives, people will be more inclined to believe that pro-social acts are done for the sake of money, not some kind of benevolent motivation. Thereby, the status of such an act is reduced, which causes people to perform less of these acts, given that they no longer benefit from the enhanced image (c.f. Benabou & Tirole, 2010). To illustrate further, consider a case of crowding out – of instrinsic motivation and image concerns – in mini-publics:

> **Crowding Out Jenny:** Mini-public invitations are sent out, promising only small compensations. Jenny receives an invitation and wants to accept immediately. She would do so out of a sense of civic duty, but she would also enjoy telling her friends about the experience. However, a mistake was made by the organization in calculating compensations. These mistakes are corrected briskly, and public announcements are made that opportunity costs will now be compensated instead. Jenny ponders the offer and declines. She is not motivated by the money, and instead chooses to forego the financial boon in order to help out at work. She is also slightly worried that, if she were to accept, her friends would think she took the offer because she dislikes her current job.

Crowding out possibilities such as these, is why we need to be careful with forcing or incentivizing people to do something.

So, we have now seen multiple solutions for achieving representativeness, but all have their issues. The following is an overview of the solutions discussed so far:

Design Solution (Representativeness)	Possible Issues
Forced Participation	Paternalism, Crowding Out
Opportunity Costs Compensation	Fairness, Crowding Out
Social Status Improvement	Feasibility, Self-Selection
Politically Binding Decisions	Feasibility, Self-Selection

Here I will not argue for any of these solutions (or any other for that matter). The point here is rather that, if one wants to achieve representativeness, then according to these solutions, one has to bite the bullet and accept potential trade-offs. That said, the upshot is that improvement of representativeness is possible. Moreover, especially if one looks at the latter two design solutions, improving representativeness can be done without glaring normative issues.

6.2 Wise Decisions

Now that we know that mini-publics *can* be designed to be representative, there is one final hurdle. Can mini-publics be designed to make wise decisions? Previously, chapter five has shown several design solutions: i) identification and evaluation separation, ii) public justification, iii) speaking order randomization, and iv) pre-deliberation vote, post-deliberation reveal. After reflection, the first solution – identification and evaluation separation – seems to be the only design feature that does not require urgent further justification; it is easy to implement, arguments become better because we are better at evaluating other people's arguments, and there do not seem to be pressing drawbacks. However, the others design solutions – in addition to some new ones – do need further inspection.

In this section, I shall begin by discussing (once more) incentives, but also self-silencing effects. Taken together, these lead to a recommendation for group rewards. Next, I discuss public justification and its potential for self-silencing. After going through potential factors that could offset the danger of self-silencing, I conclude that public justification is still likely to be a good design choice. Then, I discuss sequencing; design solutions (iii) and (iv) above. Even though these solutions have some issues, jointly they can still prevent information cascades and other self-silencing effects. Lastly, I discuss the desirability of prolonged duration, and how such duration has possible feasibility and self-selection issues.

6.2.1 Learning Incentives & Self-Silencing

The rationally ignorant agent of chapter five, has so far done nothing more than show up, as more has not been necessary. Doing stuff costs effort. And the rationally ignorant agent does not want to spend much effort when benefits of their decision-making are low, which is often the case in mini-publics with over 100 participants. To be sure, real people will have non-monetary incentives to learn, and will have dispositions to participate with some engagement. Additionally, design solutions, such as making decisions politically binding, will enhance these motivations. However, if one is sceptical about how much is learned during deliberation, then these facts do not make the problem go away entirely. Especially not if one believes that the rationally ignorant agent shows lackluster learning due to a lack of effort.

An easy way to increase this effort for learning, is rewarding success on knowledge tests, much like those that are already sometimes administered. Unfortunately, the same problems with monetary incentives arise here as before (i.e. crowding out, fairness).

Fortunately, there is a different solution to paying individual incentives, and that is group incentives. Similar to the individual case, knowledge tests could be administered, but now instead of paying for individual performance, every participant gets paid according to the average group performance. A major advantage of rewarding group performance, is that non-monetary motivations are much less likely to be crowded out, as many of these motivations – such as social engagement and image-concerns – are associated with being a team-player; something that rewarding group performance stimulates. Additionally, besides a reduced chance of crowding out, paying group performance also dissolves concerns about fairness, as everyone is paid equal.

Another benefit of rewarding group performance, is a reduction in self-silencing. Sunstein and Hastie (2015; chapter 1) idenitify two sources of self-silencing[68]:

1. **Informational signals**: lead people to fail to disclose what they know out of respect for the information publicly announced by others

---

[68] Text original.

2. **Social pressures**: lead people to silence themselves to avoid various penalties

Informational signals were already discussed before in the section about information phenomena. Less explicitly discussed were social pressures, an additional source of self-silencing. To see how this source is different, consider once more early speakers. Eager and high-status individuals often speak early. Due to respect and early public information, this can lead to informational cascades. However, not only do higher-status and eager individuals speak *earlier*, they also speak *more often*. Sometimes to the point that lower-status individuals hardly speak or engage at all. Such self-silencing can be explained by social pressures.

Now, although the example is about individuals, the same reasoning also applies to unpopular arguments, and information that is not shared by the group. And just like how informational signals can lead to informational cascades, social pressures can lead to reputational cascades. Neither are desirable from an epistemic perspective.

One solution Sunstein and Hastie (ibid-chapter 6) give to self-silencing, is group rewards: "[c]ascades are far less likely when each individual knows that he has nothing to gain from a correct *individual* decision and everything to gain from a correct *group* decision". The reason is that individuals will start revealing private information, without anybody trying to punish them anymore for doing so.

Group rewards thus incentivize *against* self-silencing and *in favor* of participants trying to engage everyone, making sure that everyone is on the same page, and that everyone has understood the various arguments. This makes rewarding group performance an attractive option, both from the perspectives of crowding out and fairness, as well as self-silencing.

6.2.2 Public Justification

The argumentative theory of reasoning from chapter five, showed how public justification could enhance reasoning performance. Namely, by preparing for public justification, groups and individuals are much less likely to accept generic reasoning, are more likely to work harder to attempt to convince their audience – especially if they are as critical as political audiences will be – and are more likely to imagine having a dialogue with their future audience, thereby putting reasoning in its more natural habitat (i.e. social argumentative contexts).

Yet it ignored a glaring issue: all modern democracies vote in secrecy. Public transparency sounds nice, but honesty is better. Why is this analogy with voting not sound? And how do we know that public justification will not simply make deliberators succumb to social pressures and thereby not reveal their private information and opinions? Or scare away possible deliberators from participating at all?[69]

One strategy for answering these questions, is to look at what politicians do. Current best practice is to have a combination of closed-doors discussions – in which honest exchanges can occur freely – and public discussions, in which rhetoric and finalized arguments take center stage. In how far this is copied by mini-publics, depends on the type. In any case, the upshot is that such a model could also work for mini-publics. Unfortunately, protecting only parts of the mini-public from public justification does not solve the issue of self-silencing. Participants will still self-silence if their final decisions need to be made public, just like how democracies would become self-silencing if voting had to be done out in the open. A different strategy is thus needed than copying best practice in politics.

---

[69] See also Chamber (2004), for an account of how secrecy could generate better deliberative performance at times.

Another promising strategy, is to look at factors that could potentially offset the self-silencing effects of public justification. Here I briefly mention four of these factors. Firstly, there are the group rewards discussed before, which lead to cooperative group dynamics that reduce self-silencing. Secondly, there are social norms that go against self-silencing. An example is how some top businesses frequently remind employees that they have an "obligation to dissent" whenever they judge something to be incorrect. Perhaps similar reminders could work for mini-publics. Thirdly, mini-publics can also use a combination of public justification *and* a secret vote, for instance by having both a consensus recommendation and a private questionnaire. Much of the discrepancy between the two can then be attributed to self-silencing[70]. In such a case, organizers would know when self-silencing has occurred, possibly leading to the discounting of consensus recommendations. Assuming participants will be aware of this, they might use less social pressure, as as they cannot control private votes as well as they can control group decisions. Fourthly, there could also be a platform for confessing anonymized opinions and arguments to the organizers of the mini-public, in which participants would be ensured anonimity, thereby reducing self-silencing effects. These admissions could then still get a public hearing, despite being anonymized.

So, will these offsetting factors sufficiently reduce self-silencing? The honest answer is that I do not know. But what I can say, is that, if reasoning-powers go up as much as the argumentative theory of reasoning suggests, and there are considerable factors able to reduce the downsides of self-silencing, then public justification seems like a good design choice.

6.2.3 Sequencing

As argued in chapter five on information phenomena, at least two design solutions seem to help against information cascades: iii) speaking order randomization, and iv) pre-deliberation votes, which are revealed post-deliberation (such that more deliberation can occur in a subsequent session, with initial votes in mind). By extension, these design solutions also apply to self-silencing due to social pressures. Now, since these design solutions do not seem to have obvious normative problems, the remaining question here, is whether these solutions have potential downsides, making them unfit for a coherent mini-public design that desires wise decisions.

Firstly, I shall discuss randomizing speaking order. Taken to the extreme, speaking order randomization definitely has a problem: decreased relevance. Consider only a seminar room in which immediate responses and back-and-forths are not allowed, because it violates rules about speaking order randomization. Deliberation would stagnate, and people would impatiently skip turns to hear what the relevant interlocutor has to say.

Fortunately, randomization need not be taken to the extreme. One example of a more moderate solution, is what are sometimes called "Round Robin" discussions, in which there are initial rounds in which each speaker gets a turn. After these rounds are finished, deliberation can be opened to a more plenary session. A similar kind of midway solution could also be used for mini-publics.

Moreover, a slightly different solution to randomization, is putting a cap on speaking turns. For instance, a fixed amount of tokens – representing speaking turns – could be distributed in equal amounts to each member, ensuring equal contribution from all members. Even eager participants might then think twice before spending all their tokens early on, thereby reducing cascade effects. Also, by requiring that all tokens are used during deliberation, self-silencing is reduced.

Secondly, there are the pre-deliberation votes that are revealed at a later stage. One worry a deliberative democrat might have, is that pre-deliberation votes will bias subsequent deliberation.

---

[70] Or genuine compromise in a consensus. However, decisions by compromise can just be written in the consensus recommendation.

For instance, consider a participant who is unsure of their preference. They are asked to vote before deliberation, which makes them have to think of a preference and state it (even if anonymous). At this point, certain cognitive biases could come into play, such as confirmation bias, making the participant less likely to deviate from initially stated preferences. However, when we look at Fishkin's Deliberative Polls, which frequently have participants take the same questionnaire before and after the mini-public, we do not see such a lack of preference change[71]. Therefore, initial voting might not bias deliberation as much as one might think. That said, if the worry is sufficiently large nonetheless, then this particular design feature could also be forfeited, as randomizing speaking order has many of the same benefits as these pre-deliberation votes (i.e. they prevent self-silencing and its negative effects).

6.2.4 Duration

Learning can take long a time. This make duration an important factor for mini-publics. Consider also once more the challenge of the mini-public critic:

> I doubt a couple days of deliberation can impart that knowledge— after a semester's worth of study, most undergraduates still don't understand, say, basic microeconomics (Brennan, 2017-215).

Abstracting from the example about microeconomics, the statement seems to be intuitively obvious. Even if people were to use their full brain capacity, only focus on the most important issues, and perfectly evaluate the plausibility of lectures given by competing expert, then some subject material will still be too difficult to expect deliberators to grasp sufficiently in several weeks, let alone days. This applies to both Einstein and Average Joe.

Therefore, there is a simple solution: make the duration of mini-publics longer. How long this should be in practice, shall depend on contextual issues, such as the subject, importance of the decision, diversity of experts, and so on. In any case, prolonging duration is likely to lead to wiser decisions.

Downsides of extending the duration of mini-publics, are the additional costs (i.e. feasibility), and potential self-selection effects, as those that have long-term obligations are more likely to decline invitations. Whether the additional costs are worth it, I do not determine here. And with respect to selection-effects, it is once more unclear how large these effects are. That said, it is at least plausible that these self-selection effects can be offset by design solutions for *representativeness* (if that is desirable).

6.2.5 Conclusions

The following table contains the design solutions of this section, with their possibly associated issues, for wise decisions:

Design Solutions (Wise Decisions)	Possible Issues
Public Justification	Self-Silencing, Self-Selection
Speaking Order Randomization	Relevance
Pre-Deliberation Vote, Post-Deliberation Reveal (Repeat)	Confirmation Bias
Identification and Evaluation Separation	–
Politically Binding Decisions	Feasibility, Self-Selection
Group Performance Incentives on Knowledge Tests	–
Prolonged Duration	Feasibility, Self-Selection

---

[71] Although perhaps there would be more preference change without the initial questionnaire.

Out of those that have possible issues, some are not too problematic, such as decreased relevance due to randomizing (to an extent) speaking order. Next, some design solutions are not necessary, such as pre-deliberation votes, and can be discarded if their downsides are too large. And some require more money and political will, such as making decisions politically binding, and prolonging mini-public duration. Furthermore, some designs potentially increase self-selection, yet it is unclear how large these effects will be. That said, it is plausible that such self-selection is more than offset by design solutions trying to attain *representativeness*. Lastly, the self-silencing effects of public justification are likely to be offset by group incentives. However, on the assumption that group incentives are insufficient, additional options exist, in the form of trying to create social norms against self-silencing, or ensuring that there is an additional anonymous platform available, in which participants can express themselves (6.2.2; not listed in table above). Therefore, given that there are good design solutions without too much epistemic downside, and without glaring normative issues, the conclusion is that mini-publics can make wise decisions.

6.3 Conclusions & Recommendations

Now that the thesis is at the end of its journey, it is appropriate to make some final comments on recommendations, unexplored territory, and implications for deliberative democracy.

6.3.1 Recommendations

To combine all discussed design features, the following is a table of design options (representativeness) and recommendation (wise decisions):

Representativeness	Wise Decisions
Forced Participation	Public Justification
Opportunity Costs Compensation	Sequencing
Social Status Increase	Identification and Evaluation Separation
Politically Binding Decisions	Politically Binding Decisions
	Group Performance Incentives
	Prolonged Duration

As discussed before, representativeness can be bought, but could come at a price. Wise decisions, however, should reliably follow from these design solutions.

That does not imply that these recommendations for wise decision-making should be set in stone, but rather that these design solutions should be corroborated by new experimental evidence. Some important reasons for this, are that some design solutions need fine-tuning (e.g. public justification, duration), and that we cannot be certain of the effect sizes of most of these design solutions, especially not given that some have complicated interactions (e.g. group performance incentives offsetting downsides of public justification).

6.3.2 Unexplored Territory

Also good to note, is that while this thesis has attempted to answer some questions, many still remain. For instance, can we make sure that all areas of expertise – agenda setting, briefing materials, teaching material, expert witnesses, and moderators – are properly balanced by stakeholder approval? And if so, how? What is the best decision rule? And under what circumstances? Is there a chance of political capture? And how can we prevent it?

Furthermore, the design here is not complete, as, for example, much inspiration can still be drawn from best practice. For instance, in the BCCA, values which were agreed upon, were given a

prominent place in each deliberation room. Now, if this practice is effective due to priming, then it could be extended to prominently placing heuristics for proper reasoning in each deliberation room. And if highlighting values is effective due to the memory cue, then the practice can be extended to prominently placing common reasoning fallacies in each deliberation room[72]. Or consider another example of best practice: using round tables to symbolize equality. For both the list of values and the round table, additional theory and experiments could prove fruitful to further enhance decision-making wisdom.

6.3.3 Implications for Deliberative Democracy

Yet even without these unanswered questions and unexplored opportunities, the conclusion is that mini-publics can make wise decisions. Human reason is not the obstacle, self-silencing can be reduced, motivation can increase, chances for learning can be prolonged, and reasoning can be put in its natural habitat. Therefore, the design solutions suggested above – despite being imperfect and incomplete – should ensure that wise decisions reliably follow.

Relatively controlled environments, such as the mini-public, can thus attain deliberative democracy's ideals. The main implication is therefore – assuming these findings are robust to empirical evidence – that mini-publics should play a more substantial role in democracies. Yet that need not imply that aspirations for deliberation in society at large should be renounced. Especially because these insights into group deliberation are quite transferable, whether that is to business or government institutions. The upshot for deliberative democracy is therefore, at minimum, rejuvinated hope for mini-publics making wise decisions, and at best, an inspiration to adopt more of these mini-public designs, and to transfer its design lessons to other deliberative environments in society.

---

[72] In the case of reasoning fallacies, with a cautionary note that many of them can be legitimate under certain conditions.

References

ACHEN, C. H., & BARTELS, L. M. (2017). DEMOCRACY FOR REALISTS: WHY ELECTIONS DO NOT PRODUCE RESPONSIVE GOVERNMENT (VOL. 4). PRINCETON UNIVERSITY PRESS.

ACKERMAN, B., & FISHKIN, J. S. (2002). DELIBERATION DAY. JOURNAL OF POLITICAL PHILOSOPHY, 10(2), 129-152.

BACCARO, L., BÄCHTIGER, A., & DEVILLE, M. (2016). SMALL DIFFERENCES THAT MATTER: THE IMPACT OF DISCUSSION MODALITIES ON DELIBERATIVE OUTCOMES. BRITISH JOURNAL OF POLITICAL SCIENCE, 46(3), 551-566.

BÄCHTIGER, A. (2018). A PREFACE TO STUDYING DELIBERATION EMPIRICALLY. IN THE OXFORD HANDBOOK OF DELIBERATIVE DEMOCRACY.

BÄCHTIGER, A., DRYZEK, J. S., MANSBRIDGE, J., & WARREN, M. E. (EDS.). (2018). THE OXFORD HANDBOOK OF DELIBERATIVE DEMOCRACY. OXFORD UNIVERSITY PRESS.

BARNETT, A., & CARTY, P. (2017). THE ATHENIAN OPTION: RADICAL REFORM FOR THE HOUSE OF LORDS (VOL. 5). ANDREWS UK LIMITED.

BELL, D. A. (2016). THE CHINA MODEL: POLITICAL MERITOCRACY AND THE LIMITS OF DEMOCRACY. PRINCETON UNIVERSITY PRESS.

BÉNABOU, R., & TIROLE, J. (2010). INDIVIDUAL AND CORPORATE SOCIAL RESPONSIBILITY. ECONOMICA, 77(305), 1-19.

BLACK, L. W., BURKHALTER, S., GASTIL, J., & STROMER-GALLEY, J. (2011). METHODS FOR ANALYZING AND MEASURING GROUP DELIBERATION. SOURCEBOOK OF POLITICAL COMMUNICATION RESEARCH: METHODS, MEASURES, AND ANALYTICAL TECHNIQUES, 323, 345.

BLAIS, A., CARTY, R. K., & FOURNIER, P. (2008). DO CITIZENS' ASSEMBLIES MAKE REASONED CHOICES. DESIGNING DELIBERATIVE DEMOCRACY: THE BRITISH COLUMBIA CITIZENS' ASSEMBLY, 127-44.

BOWLES, S. (2016). THE MORAL ECONOMY: WHY GOOD INCENTIVES ARE NO SUBSTITUTE FOR GOOD CITIZENS. YALE UNIVERSITY PRESS.

BRENNAN, J. (2017). AGAINST DEMOCRACY: NEW PREFACE. PRINCETON UNIVERSITY PRESS.

BUCHANAN, A. (2009). PHILOSOPHY AND PUBLIC POLICY: A ROLE FOR SOCIAL MORAL EPISTEMOLOGY. JOURNAL OF APPLIED PHILOSOPHY, 26(3), 276-290.

CALUWAERTS, D., & UGARRIZA, J. (2012). FAVORABLE CONDITIONS TO EPISTEMIC VALIDITY IN DELIBERATIVE EXPERIMENTS: A METHODOLOGICAL ASSESSMENT. JOURNAL OF PUBLIC DELIBERATION, 8(1).

CAPLAN, B. (2007). THE MYTH OF THE RATIONAL VOTER: WHY DEMOCRACIES CHOOSE BAD POLICIES. PRINCETON UNIVERSITY PRESS.

CAPLAN, B. (2012). THE MYTH OF THE RATIONAL VOTER AND POLITICAL THEORY. COLLECTIVE WISDOM: PRINCIPLES AND MECHANISMS, 319.

CAPLAN, B. (2018). THE CASE AGAINST EDUCATION: WHY THE EDUCATION SYSTEM IS A WASTE OF TIME AND MONEY. PRINCETON UNIVERSITY PRESS.

CARMAN, K. L., MALLERY, C., MAURER, M., WANG, G., GARFINKEL, S., YANG, M., ... & GOLD, M. (2015). EFFECTIVENESS OF PUBLIC DELIBERATION METHODS FOR GATHERING INPUT ON ISSUES IN HEALTHCARE: RESULTS FROM A RANDOMIZED TRIAL. SOCIAL SCIENCE & MEDICINE, 133, 11-20.

CHAMBERS, S. (2004). BEHIND CLOSED DOORS: PUBLICITY, SECRECY, AND THE QUALITY OF DELIBERATION. JOURNAL OF POLITICAL PHILOSOPHY, 12(4), 389-410.

CITIZENS'ASSEMBLY, B. C. (2004). MAKING EVERY VOTE COUNT: THE CASE FOR ELECTORAL REFORM IN BRITISH COLUMBIA.

COHEN, G. L. (2003). PARTY OVER POLICY: THE DOMINATING IMPACT OF GROUP INFLUENCE ON POLITICAL BELIEFS. JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, 85(5), 808.

CONVERSE, P. E. (2006). THE NATURE OF BELIEF SYSTEMS IN MASS PUBLICS (1964). CRITICAL REVIEW, 18(1-3), 1-74.

CORNER, A., HAHN, U., & OAKSFORD, M. (2011). THE PSYCHOLOGICAL MECHANISM OF THE SLIPPERY SLOPE ARGUMENT. JOURNAL OF MEMORY AND LANGUAGE, 64(2), 133-152.

CURATO, N., DRYZEK, J. S., ERCAN, S. A., HENDRIKS, C. M., & NIEMEYER, S. (2017). TWELVE KEY FINDINGS IN DELIBERATIVE DEMOCRACY RESEARCH. DAEDALUS, 146(3), 28-38.

CURATO, N., & BÖKER, M. (2016). LINKING MINI-PUBLICS TO THE DELIBERATIVE SYSTEM: A RESEARCH AGENDA. POLICY SCIENCES, 49(2), 173-190.

DOWNS, A. (1957). AN ECONOMIC THEORY OF DEMOCRACY.

DRYZEK, J. S. (2002). DELIBERATIVE DEMOCRACY AND BEYOND: LIBERALS, CRITICS, CONTESTATIONS. OXFORD UNIVERSITY PRESS ON DEMAND.

ELQAYAM, S., & OVER, D. E. (2013). NEW PARADIGM PSYCHOLOGY OF REASONING: AN INTRODUCTION TO THE SPECIAL ISSUE EDITED BY ELQAYAM, BONNEFON, AND OVER. THINKING & REASONING, 19(3-4), 249-265.

ELQAYAM, S. (2017). THE NEW PARADIGM IN PSYCHOLOGY OF REASONING. IN INTERNATIONAL HANDBOOK OF THINKING AND REASONING(PP. 152-172). ROUTLEDGE.

ERISEN, C., LODGE, M., & TABER, C. S. (2014). AFFECTIVE CONTAGION IN EFFORTFUL POLITICAL THINKING. POLITICAL PSYCHOLOGY, 35(2), 187-206.

ELSTUB, S. (2010). THE THIRD GENERATION OF DELIBERATIVE DEMOCRACY. POLITICAL STUDIES REVIEW, 8(3), 291- 307.

EMMOTT, B. (2017). THE FATE OF THE WEST: THE BATTLE TO SAVE THE WORLD'S MOST SUCCESSFUL POLITICAL IDEA. THE ECONOMIST.

ESTERLING, K. (2018). THE OXFORD HANDBOOK OF DELIBERATIVE DEMOCRACY. OXFORD UNIVERSITY PRESS.

ESTLUND, D. M. (2009). DEMOCRATIC AUTHORITY: A PHILOSOPHICAL FRAMEWORK. PRINCETON UNIVERSITY PRESS.

ESTLUND, D., & LANDEMORE, H. (2018). THE EPISTEMIC VALUE OF DEMOCRATIC DELIBERATION. THE OXFORD HANDBOOK OF DELIBERATIVE DEMOCRACY, 113.

Farrar, C., Fishkin, J. S., Green, D. P., List, C., Luskin, R. C., & Paluck, E. L. (2010). Disaggregating deliberation's effects: An experiment within a deliberative poll. British journal of political science, 40(2), 333-347.

Fishkin, J. S., He, B., Luskin, R. C., & Siu, A. (2010). Deliberative democracy in an unlikely place: Deliberative polling in China. British Journal of Political Science, 40(2), 435-448.

FISHKIN, J. S. (2011). WHEN THE PEOPLE SPEAK: DELIBERATIVE DEMOCRACY AND PUBLIC CONSULTATION. OXFORD UNIVERSITY PRESS.

Fishkin, J. S. (2011b). Deliberative polling. Consulting the people is not as easy as it sounds.

FUKUYAMA, F. (2006). THE END OF HISTORY AND THE LAST MAN. SIMON AND SCHUSTER.

GASTIL, J. (2017). DESIGNING PUBLIC DELIBERATION AT THE INTERSECTION OF SCIENCE AND PUBLIC POLICY. THE OXFORD HANDBOOK OF THE SCIENCE OF SCIENCE COMMUNICATION, 233-242.

Gastil, J. (2018). The lessons and limitations of experiments in democratic deliberation. Annual Review of Law and Social Science, 14, 271-291.

Goodin, R. E., & Dryzek, J. S. (2006). Deliberative impacts: the macro-political uptake of mini-publics. Politics & society, 34(2), 219-244.

Goodwin, B. (2013). Justice by lottery. Andrews UK Limited.

Gutmann, A., & Thompson, D. (2004). Why deliberative democracy?. Princeton University Press.

Haack, S. (2012). Six signs of scientism. Logos & Episteme, 3(1), 75-95.

HALPERN, D. F. (2013). THOUGHT AND KNOWLEDGE: AN INTRODUCTION TO CRITICAL THINKING. PSYCHOLOGY PRESS.

HANSEN, P. G., HENDRICKS, V. F., & RENDSVIG, R. K. (2013). INFOSTORMS. METAPHILOSOPHY, 44(3), 301- 326.

Healy, A., & Malhotra, N. (2013). Retrospective voting reconsidered. Annual Review of Political Science, 16, 285-306.

Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. Academic medicine, 73(5), 538-540.

Hoffrage, U., Krauss, S., Martignon, L., & Gigerenzer, G. (2015). Natural frequencies improve Bayesian reasoning in simple and complex inference tasks. Frontiers in psychology, 6, 1473.

Holzinger, K. (2004). Bargaining through arguing: an empirical analysis based on speech act theory. Political Communication,, 21(2), 195-222.

Humphreys, M., Masters, W. A., & Sandbu, M. E. (2006). The role of leaders in democratic deliberations: results from a field experiment in São Tomé and Príncipe. World Politics, 58(4), 583-622.

Ingham, S. (2013). Disagreement and epistemic arguments for democracy. Politics, Philosophy & Economics, 12(2), 136-155.

James, M. R. (2008). Descriptive Representation in the British Columbia Citizens' Assembly. Designing deliberative democracy: The British Columbia citizens' assembly, 106-126.

Joss, S., & Durant, J. (Eds.). (1995). Public participation in science: The role of consensus conferences in Europe. NMSI Trading Ltd.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. Behavioural Public Policy, 1(1), 54-86.

Karjalainen, M., & Rapeli, L. (2015). Who will not deliberate? Attrition in a multi-stage citizen deliberation experiment. Quality & Quantity, 49(1), 407-422.

Kelly, J. T. (2012). Framing democracy: A behavioral approach to democratic theory. Princeton University Press.

Landemore, H. (2014). Yes, we can (make it up on volume): Answers to critics. Critical Review, 26(1-2), 184-237.

Landemore, H. (2017). Beyond the fact of disagreement? The epistemic turn in deliberative democracy. Social Epistemology, 31(3), 277-295.

Levitsky, S., & Ziblatt, D. (2018). How democracies die. Crown.

List, C., Luskin, R. C., Fishkin, J. S., & McLean, I. (2012). Deliberation, single-peakedness, and the possibility of meaningful democracy: evidence from deliberative polls. The journal of politics, 75(1), 80-95.

Lerner, J. A. (2014). Making democracy fun: How game design can empower citizens and transform politics. MIT Press.

Lupia, A., McCubbins, M. D., & Arthur, L. (1998). The democratic dilemma: Can citizens learn what they need to know?. Cambridge University Press.

López-Guerra, C. (2014). Democracy and disenfranchisement: The morality of electoral exclusions. Oxford University Press, USA.

Mansbridge, J. (2010). Deliberative polling as the gold standard. The Good Society, 19(1), 55-62.

Mendelberg, T. (2002). The deliberative citizen: Theory and evidence. Political decision making, deliberation and participation, 6(1), 151-193.

Mercier, H., & Landemore, H. (2012). Reasoning is for arguing: Understanding the successes and failures of deliberation. Political Psychology, 33(2), 243-258.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. Behavioral and Brain Sciences, 34(2), 57-74.

Mercier, H., & Sperber, D. (2017). The enigma of reason. Harvard University

Mercier, H. (2018). Reasoning and argumentation. In V. A. Thompson & L. J. Ball (Eds.), International handbook of thinking & reasoning. Hove, UK: Psychology Press

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. Policy Insights from the Behavioral and Brain Sciences, 2(1), 129-140.

MOUNK, Y. (2018). THE PEOPLE VS. DEMOCRACY: WHY OUR FREEDOM IS IN DANGER AND HOW TO SAVE IT. HARVARD UNIVERSITY PRESS.

MUTZ, D. C. (2006). HEARING THE OTHER SIDE: DELIBERATIVE VERSUS PARTICIPATORY DEMOCRACY. CAMBRIDGE UNIVERSITY PRESS.

PARKINSON, J., & MANSBRIDGE, J. (EDS.). (2012). DELIBERATIVE SYSTEMS: DELIBERATIVE DEMOCRACY AT THE LARGE SCALE. CAMBRIDGE UNIVERSITY PRESS.

PEARSE, H. (2008). INSTITUTIONAL DESIGN AND CITIZEN DELIBERATION. DESIGNING DELIBERATIVE DEMOCRACY: THE BRITISH COLUMBIA CITIZENS' ASSEMBLY, 70-84.

PIAGET, J. (1928). JUDGMENT AND REASONING IN THECHILD. LONDON: KEGAN PAUL, TRENCH, TRUBNER.

PINCOCK, H. (2012). DOES DELIBERATION MAKE BETTER CITIZENS. DEMOCRACY IN MOTION: EVALUATING THE PRACTICE AND IMPACT OF DELIBERATIVE CIVIC ENGAGEMENT, 135-162.

RAVAZZI, S., & POMATTO, G. (2014). FLEXIBILITY, ARGUMENTATION AND CONFRONTATION. HOW DELIBERATIVE MINIPUBLICS CAN AFFECT POLICIES ON CONTROVERSIAL ISSUES.

RODRIK, D. (2015). ECONOMICS RULES: THE RIGHTS AND WRONGS OF THE DISMAL SCIENCE. WW NORTON & COMPANY.

RUNCIMAN, D. (2018). HOW DEMOCRACY ENDS. BASIC BOOKS.

RYAN, R. M., & DECI, E. L. (2000). INTRINSIC AND EXTRINSIC MOTIVATIONS: CLASSIC DEFINITIONS AND NEW DIRECTIONS. CONTEMPORARY EDUCATIONAL PSYCHOLOGY, 25(1), 54-67.

SETÄLÄ, M. (2014). DELIBERATIVE MINI-PUBLICS: INVOLVING CITIZENS IN THE DEMOCRATIC PROCESS. ECPR PRESS.

SIMON, R. L. (ED.). (2008). THE BLACKWELL GUIDE TO SOCIAL AND POLITICAL PHILOSOPHY (VOL. 20). JOHN WILEY & SONS.

SMITH, G. (2009). DEMOCRATIC INNOVATIONS: DESIGNING INSTITUTIONS FOR CITIZEN PARTICIPATION. CAMBRIDGE UNIVERSITY PRESS.

SOMIN, I. (2016). DEMOCRACY AND POLITICAL IGNORANCE: WHY SMALLER GOVERNMENT IS SMARTER. STANFORD UNIVERSITY PRESS.

SPADA, P., & RYAN, M. (2017). THE FAILURE TO EXAMINE FAILURES IN DEMOCRATIC INNOVATION. PS: POLITICAL SCIENCE & POLITICS, 50(3), 772-778.

STEENBERGEN, M. R., BÄCHTIGER, A., SPÖRNDLI, M., & STEINER, J. (2003). MEASURING POLITICAL DELIBERATION: A DISCOURSE QUALITY INDEX. COMPARATIVE EUROPEAN POLITICS, 1(1), 21-48.

STEINER, J. (2012). THE FOUNDATIONS OF DELIBERATIVE DEMOCRACY: EMPIRICAL RESEARCH AND NORMATIVE IMPLICATIONS. CAMBRIDGE UNIVERSITY PRESS.

STROMER-GALLEY, J. (2007). MEASURING DELIBERATION'S CONTENT: A CODING SCHEME. JOURNAL OF PUBLIC DELIBERATION, 3(1), 12.

SUNSTEIN, C. R. (2000). DELIBERATIVE TROUBLE? WHY GROUPS GO TO EXTREMES. THE YALE LAW JOURNAL, 110(1), 71-119.

SUNSTEIN, C. R. (2015). PARTYISM. U. CHI. LEGAL F., 1.

SUNSTEIN, C. R., & HASTIE, R. (2015). WISER: GETTING BEYOND GROUPTHINK TO MAKE GROUPS SMARTER. HARVARD BUSINESS PRESS.

THOMPSON, D. F. (2008). DELIBERATIVE DEMOCRATIC THEORY AND EMPIRICAL POLITICAL SCIENCE. ANNU. REV. POLIT. SCI., 11, 497-520.

VALENTINI, L. (2012). IDEAL VS. NON-IDEAL THEORY: A CONCEPTUAL MAP. PHILOSOPHY COMPASS, 7(9), 654-664.

VANLEHN, K., & VAN DE SANDE, B. (2009). ACQUIRING CONCEPTUAL EXPERTISE FROM MODELING: THE CASE OF ELEMENTARY PHYSICS. THE DEVELOPMENT OF PROFESSIONAL PERFORMANCE: TOWARD MEASUREMENT OF EXPERT PERFORMANCE AND DESIGN OF OPTIMAL LEARNING ENVIRONMENTS, 356-378.

WARREN, M. E., & PEARSE, H. (EDS.). (2008). DESIGNING DELIBERATIVE DEMOCRACY: THE BRITISH COLUMBIA CITIZENS' ASSEMBLY. CAMBRIDGE UNIVERSITY PRESS.

WEEDEN, J., & KURZBAN, R. (2017). SELF-INTEREST IS OFTEN A MAJOR DETERMINANT OF ISSUE ATTITUDES. POLITICAL PSYCHOLOGY, 38, 67-90.

WUTTKE, A. (2019). WHY TOO MANY POLITICAL SCIENCE FINDINGS CANNOT BE TRUSTED AND WHAT WE CAN DO ABOUT IT: A REVIEW OF META-SCIENTIFIC RESEARCH AND A CALL FOR ACADEMIC REFORM. POLITISCHE VIERTELJAHRESSCHRIFT, 60(1), 1-19.

Appendix

A1. Schema of a Mini-Public (illustration)

1. Pre-deliberation
   a. Agenda Setting
      i. Experts
      ii. Participants
   b. Public Consultation
      i. Identify stakeholders
      ii. Assemble evidence from stakeholders
   c. Experts
      i. Agreed on by stakeholders
   d. Briefing Material
      i. Assembled by academics
      ii. Agreed on by stakeholders
   e. Selection of Participants
      i. Random inivitations
      ii. Randomly (stratified by salient social groups) select from those that accept invitation
2. Deliberation
   a. Learning Phase
      i. Briefing materials
      ii. Lectures (by experts)
      iii. Small Group Deliberation

b. Deliberation Phase
            i. Small Group Deliberation
            ii. Make Recommendation

A2. Sources of Table[73]

- Discussion Group Size
    o Deliberative Polls (Fishkin, 2011)
- Norm Enforcement by moderator (Smith, 2009-85/88)
- Duration
    o Deliberation Polls (Fishkin, 2011)
    o Other (Smith, 2009)
- Transparency
    o General (Smith, 101-105)
    o Concensus conferences (Joss & Durant, 1995-100)
- Incentives
    o Deliberative Polls (Fishkin, 2011)
- Costs (Smith, 2009-105)
- Citizens' Assembly Information (Citizens' Assembly, 2004)

A3. Opportunity Costs

Current direct monetary incentives in mini-publics are stipends (i.e. small compensations for time, effort, and travelling costs). How large this stipend is, will determine whether this compensation is sufficient to show up.

One possible way to analyze this problem, is to draw an analogy with other forms of labor contracts. Economic orthodoxy uses the concept of *reservation wage* (i.e. the lowest wage which an agent is willing to accept for their labor) for this purpose. In turn, the reservation wage depends on the following:

> **Opportunity Costs:** costs incurred from not gaining the (net) benefits from the best alternative choice

In job-related terms, these opportunity costs refer to missing wages, and forfeiting the opportunity to gain extra experience or credentials on the job. In more non-monetary terms, it could also refer to missed opportunities for vacation, resting from injuries, or time away from family. Here, I will focus only on the opportunity costs of missing wages. The idea being that this will be a lower bound of what people will accept based on opportunity costs.

Now, in order for us to calculate amount, let's start to make some assumptions. So far, we have the following assumptions:

(1) If we want (non-mandatory) full acceptance of mini-publics, then we need to compensate the opportunity costs of missing wages[74]
(2) We want (non-mandatory) full acceptance of mini-publics

---

[73] If unspecified, then from Smith (2009).

[74] I am also assuming in the exercise here, that paying people unevenly based on previous wages, is normatively acceptable.

From this, we can conclude that we should compensate the opportunity costs of missing wages.

We also need to know who is paying:

(3) Compensation is provided by government[75]

And we need to know how much the opportunity costs are. For the purposes here, I will us some simplifying assumptions that will make calculation easier:

(4) Agents can work as many hours as they like[76]
(5) Agents get the same wage as their current contract for all additional hours[77]
(6) Agents are indifferent between working more hours and having free time[78]

Lastly, in order to focus on opportunity costs of wages only, some additional simplifying assumptions:

(7) There are no non-monetary benefits
(8) There are no fringe benefits
(9) Travel time is not paid
(10) There are no job-experience benefits
(11) The are no additional costs incurred by firms from missing an economic agent

Now, in order to calculate costs, we also need gross wage and a tax rate. The (mean) average hourly wage in the Netherlands in 2017 was €22.23[79]. And the tax rates in 2019 range from 36.65% to 51.75% (depending on the height of the income). Given that the example is not supposed to be nation specific – or precise – let's assume an hourly wage of €22 and a tax rate of 40% on wages.

Lastly, let's assume that we calculate for a three day mini-public (e.g. Deliberative Poll) in which there are an average of seven hours a day of participation, which makes 21 hours of total mini-public participation[80].

These assumptions lead to the following calculation:

(Mean) Opportunity costs: 22 (hourly wage) x 21 (mini-public hours) x 0.6 (1 - tax rate) = €277.20

So, €277.2 is the amount that needs to be paid to participants (on average). And given that this calculation foregoes tax revenue, costs for the government are €462 per participant[81].

Now let's assume that there are 200 participants. This gives the following compensations costs:

Total Compensation Costs: 22 (hourly wage) x 21 (mini-pulic hours) x 200 (participants) = €92,400

---

[75] Given that government is financing mini-publics.
[76] Such that everyone agent can be a wage earner (which is of course not the case). And such that free time could be traded off against against paid work hours without limit (which is also not the case).
[77] Agents without current contracts get perhaps minimum wage or their previous wage. Here, however, I am assuming everyone has a contract.
[78] Generally, agents value a single unit of free time more when their working hours are long. This assumption therefore helps, because otherwise we would have to specify for individuals how much they value additional units of free time versus working.
[79] Source: Centraal Plan Bureau.
[80] Hours are chosen fairly abitrarily.
[81] 22 x 21 = 462

Given these calculations, how do mini-publics compare to these numbers? Firstly, let's look at Deliberative Polls. According to table 1, Deliberative Polls pay $50-300 for completing the poll. The high end of this spectrum compares to the €277.20 as calculated above. Moreover, total costs of a Deliberative Poll are around $200.000. Total costs would thus not go up by an order of magnitude.

Secondly, let's look at the British Columbia Citizens' Assembly (BCCA). The compensation here was $150 for each meeting day. Given that for three days, average opportunity costs are €277.20, the $450 compensation of the assembly is already sufficiently high for a substantial proportion of citizens. Total costs of the assembly would thus also not go up.

One thing to note, however, is that many participants did not accept the invitation in these mini-publics, despite compensations being close to calculated opportunity costs. In the model here, that would mean that the lower bound of opportunity costs calculated above, leaves out a substantial part of the actual opportunity costs. Or it could be that some citizens attach additional costs to participating in mini-publics, perhaps because they dislike participating in politics[82]. It is thus good to emphasize that these calculations are lower bounds, and that actual compensation might therefore need to be higher.

---

[82] But also other factors, such as citizens not understanding to invitation, or caring more about their free time.