

Counting *and* Talking:
a Benthamite view of Public
Reasoning

James Shafe (721635)

PhD, Political Theory

University College London

“No Man...can be more sensible than I am to the abuses to which the maxim that private interest should give way to public, is liable...as if the public were made up of anything but individuals”
(Bentham, 1981: 472)

I, James Shafe, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis considers two separate areas of debate in Bentham scholarship and political theory to discern what the latter might learn from the former.

The concept of “public reason” has been developed largely by contemporary deliberative democrats (Habermas, 1997). These thinkers commonly believe that talking between voters holds a special value because it allows citizens to engage with one another in a particular way. This mode of participation or “deliberation” is often designed to provoke individuals to move beyond their “mere” self-interest, and instead consider what is best for the community before they vote. Rawls (1997), for example, believes that this type of talking between citizens should be subject to certain constraints: citizens should appeal only to shared or “public” reasons when justifying arguments on fundamentally important matters. This basic idea has many different formulations, with different implications for how public reasoning within democracy is viewed. In contrast, aggregative democratic theorists are generally thought to contribute less to this debate. They insist that individuals’ basic preferences are fixed, and do not transform during discussion.

I aim to establish how existing—“conventional” and “participatory”—interpretations of Bentham fit into this debate, and to argue that there is an alternative. This alternative “hybrid” view reflects Bentham’s work on publicity and official aptitude. It casts public reasoning as a way of securing the public against the misconduct of officials. This interpretation clarifies our understanding of Bentham in relation to contemporary writing on public reason. It strengthens the view that aggregative and deliberative democracy should be seen as compatible, rather than opposing approaches. Further, it does not expect or require consensus and is based on participants’ self-interest. As a result, it is less susceptible than contemporary Rawlsian and Habermasian views to critiques based on false consensus and exclusion.

Contents

Introduction		<u>p.5</u>
Chapter 1.	Two challenges for an account of public reasoning	<u>p.13</u>
Chapter 2.	A conventional view of Bentham on public reasoning	<u>p.62</u>
Chapter 3.	A participatory view of Bentham	<u>p.92</u>
Chapter 4.	Bentham reconsidered: the hybrid view	<u>p.127</u>
Chapter 5.	Implications of the hybrid view	<u>p.170</u>

Introduction

This thesis considers two separate areas of debate in Bentham scholarship and political theory to discern what the latter might learn from the former. In doing so, it considers both interpretive and normative questions. The former concern whether different interpretations accurately explain Bentham's views. The latter query whether those accounts tell us more about how citizens should reason together before voting.

In contemporary democratic theory, the concept of "public reason" has been developed largely by proponents of deliberative democracy. These thinkers commonly believe that talking between voters holds a special value because it allows citizens to engage with one another in a particular way (Habermas, 1997). This mode of participation or "deliberation" is often designed to provoke individuals to move beyond their "mere" self-interest, and instead consider what is best for the community before they vote. Rawls (1997), for example, believes that this type of talking between citizens should be subject to certain constraints. Citizens should appeal only to shared or "public" reasons when justifying arguments on fundamentally important matters.

This basic idea has many different formulations, with different implications for how we view public reasoning within democracy. In Chapter 1, I explore an existing distinction from the public reasoning literature to contrast "deliberative" and "aggregative" approaches. The deliberative approach to public reasoning appeals to some shared sense of how justifiable or generalizable arguments and reasons are in public discussion. These theories commonly ask citizens to focus on something beyond self-interest. They often suggest that deliberating can transform—rather than merely correct or clarify—preferences. Often, this type of deliberation is held to have a particular status because it reflects, creates or constitutes important democratic, even moral, norms. Deliberation is also often expected to create a particular type of agreement, or consensus, between citizens.

These accounts are often contrasted with aggregative views. These may encourage citizens to focus on either their own self-interest, or wider concerns. They insist that individuals' basic preferences are fixed, and do not transform during discussion. Usually, aggregationists grant pre-voting discussion no particular moral, or democratic, status. They usually see consensus as unnecessary. In their view, good decisions can be reached by counting votes rather than the products of discussion.

One way to understand the differences between aggregative and deliberative theories is to examine how they respond to particular challenges. Two of these are outlined in Chapter 1. The first challenge is an accusation often made against aggregative theory by proponents of deliberative democracy. It is to provide a plausible explanation of why discussion before voting should be considered more than a “mere preliminary” to choosing at the ballot box (Barber, 2003:198). Deliberative theories tend to provide a convincing response to this challenge. They explain in detail both the nature of “valuable” discussion before voting, and the democratic benefits it is expected to create. In contrast, aggregative views often struggle to answer this challenge. They fundamentally assume that discussion cannot “transform” preferences in the way that deliberative democrats suppose. This curtails the extent to which aggregative accounts may explain the role of public reasoning.

The second challenge is sometimes levelled against deliberative democracy by aggregative theorists. It is to explain how discussion before voting can be protected against manipulation by elites. In this thesis, an elite is considered an individual, or group, which has influence over discussion that is not merited by its expertise, or number. Manipulation may happen when deliberating citizens are induced to hold opinions that they would not have held given the opportunity to consider relevant information fully themselves (Zaller in Shapiro, 1998:514). Aggregative theories are in a relatively strong position to answer this challenge. Since they give public reasoning a less ambitious (transformative) role, aggregative conceptions of pre-voting discussion are less likely to produce both deliberative benefits, and manipulated harm for citizens. In contrast, the more ambitious role given to pre-voting discussion by deliberative theorists potentially creates more opportunities for manipulated outcomes. These two challenges therefore illustrate strengths and weaknesses of different approaches to public reasoning.

In this context, Bentham’s political theory is particularly interesting. It shows the potential for a middle way between aggregative and deliberative democracy. Bentham is often considered to be an aggregative democrat with little to say about the value of public reasoning. I believe that re-examination of his views suggests that he had far more to say about the value of pre-voting discussion than often assumed. Admittedly, the contemporary relevance of these Benthamite views will depend on the viability of some key assumptions. These include the possibility of reaching the public interest through an aggregation of self-interested votes. The validity of those assumptions—which relate to matters other than public reasoning itself—will not be in the scope of this thesis.

How exactly might Bentham fit into this debate? Our understanding of his views on publicity and public opinion has matured in recent years. New evidence and sophisticated analysis have emerged in the shape of re-edited texts (Bentham, 1983b) and acute contributions to the secondary literature (Rosen, 1983; ed. Zhai, & Quinn, 2014). Despite advances¹ in both fields, however, there has not been a concerted effort to understand how Bentham's views might fit into this contemporary debate. On the one hand, Bentham scholars principally analyse texts in terms of their place in the history of political thought. These contributions rarely consider the terms and questions considered relevant to debates in normative theory. On the other hand, many contemporary normative theorists overlook Bentham: they often assume that he was an aggregative theorist with little to contribute. This thesis addresses this by showing that public reasoning can be understood in relation to principles that emerge from Bentham's work on publicity and aptitude. I term this "public aptitude"; a quality among the public which is fostered by a particular type of reasoning between and among citizens. This reasoning is part of a "hybrid" interpretation of Bentham, which I draw in contrast to "conventional" and "participatory" views.

Chapter 2 focuses on a "simple" or "conventional" view of Bentham as an archetypally aggregative theorist. That interpretation suggests that there is little, if any, significant role for public reasoning in Benthamite democracy. This is because the public interest is reflected through an aggregation of self-interested votes rather than public-spirited discussion. That conclusion follows from some other assumptions of Bentham's work. Three of these are particularly important. First, the belief that Bentham defines utility as the satisfaction of "raw" preferences suggests that discussion cannot hold value in light of its ability to "reform" those preferences. Second, Bentham's quantitative utilitarianism suggests that public reasoning cannot be valuable in light of its ability to stimulate qualitative improvements in citizens' preferences. Third, Bentham's aggregative definition of the public interest—which Postema (1986) refers to as his "individualism"—means that aggregation without deliberation is, in any case, sufficient to reach decisions that reflect the public interest. Clearly, this suggests that the conventional view could leave little room for a plausible account of the value of pre-voting discussion.

¹ There has been a great deal of fruitful Bentham scholarship in recent years, often sparked on by the publication of re-edited editions of Bentham's works. Amongst other things, this evolving literature has produced a comprehensive clarification of Bentham's views on democracy (Rosen, 1983), and the story of how and why he became a democrat (Schofield, 2006); analysis of how he believed legislators should use measures beyond legislation to promote the greatest happiness of the greatest number (Brunon-Ernst, 2012); and explanation of how he believed officials should develop the skills needed to create good legislation (Bozzo-Rey, 2010). Similarly, the distributive implications of Bentham's utilitarianism have been reappraised as scholars have explored the implications of the "subordinate ends" he suggested to legislators (Quinn, 2008). Often, this scholarship has corrected misapprehensions about Bentham's views and, in doing so, presented Benthamite responses to normative questions when some observers had assumed there was none. The foundational assumptions behind Bentham's ethics and political theory have also not been immune to reinterpretation, with varying degrees of success. Lyons (1973), for instance, suggested that Bentham intended the principle of utility to apply in two different ways to private ethics and public government. To take just one example, many contemporary theorists had assumed that Bentham offered little by way of an account of rights until Paul Kelly (1990) and others had corrected that misapprehension by analysing the importance of security to Bentham's political theory.

In contrast, participatory readings—outlined in Chapter 3—suggest that public reasoning has a more important role to play in Bentham’s democratic theory. I draw participatory approaches from a number of scholars. Some of these—including Niesen and Crimmins—imply that we should reject the assumption that Bentham wished citizens to consider and vote according to their own self-interest. I believe that this would be a mistake. A self-interested view of pre-voting discussion sits more comfortably with Bentham’s analysis of language and epistemology, and with his argument for the secret ballot. Relatedly, I do not think that Bentham should be read as a deliberative democrat; contrary to arguments proposed by Ben-Dor. There is little to suggest that he espoused the kind of public-spirited preference transformation and full consensus championed by Habermas and his contemporaries. As one delegate at a recent conference said of Ben-Dor’s reinterpretation: “it’s beautiful...but it’s not Bentham”. In relation to public reasoning, at least, I agree with that verdict. Other participatory views do reflect Bentham’s focus on self-interest. Often, however, they do not directly address central questions in the debate on public reasoning.

In general, and for different reasons, these more “participatory” views of Bentham do not provide grounds for a reading of Bentham which answers both of the challenges outlined in Chapter 1. Some fail to provide plausible grounds for an interpretation of Bentham’s views (Crimmins, 2011; Ben-Dor, 2000; Niesen, 2011). Other accounts do provide a convincing examination of Bentham’s views, but do not consider questions that are directly relevant to the contemporary normative debate. Participatory views thus do not provide the resources we need to understand Bentham’s views and respond to the two normative challenges presented.

In Chapter 4, I aim to outline a reading of Bentham which is consistent with the trends in his writing. This begins with a summary of Bentham’s logic for democracy, which was not founded on an inalienable public right to democratic participation. Rather, Bentham became a democrat because he came to believe that democracy was necessary to counter the “immediate cause of all mischief of mis-rule... [which was]that the men acting as representatives of the people have a private and sinister interest...producing a constant sacrifice of the interest of the people” (Bentham, 1818a:24).

Put simply, particular interests were “sinister” when the public interest was sacrificed in their pursuit (Bentham, 1821a:80). Officials would only serve the public interest if they had the right kind of aptitude. Bentham believed that this “official aptitude” had three aspects: moral, intellectual, and active. Intellectual aptitude meant having the capacity to analyse and understand measures that would promote

the public interest (Schofield, 2006: 274). Active aptitude was the physical activity of performing one's duty, rather than simply understanding it in the abstract (1989:87).

Moral aptitude was most important to the role of public reasoning. It was a negative quality: the absence of a motivation—generally found in man—to act in one's own interest at the expense of all other interests. It was not innate to the character of virtuous men. Rather, moral aptitude was to be created and encouraged by the environment officials found themselves in (Bentham, 1989:14-5). Low levels of moral aptitude were a feature of situations where officials could pursue sinister interests with little chance of suffering as a result. In those situations, the divergence between rulers' particular interests and the public interest exposed them to great (sinister) temptation. In contrast, official—particularly moral—aptitude was encouraged when officials were made dependent on the public they served (Bentham, 1818a:24). This was essential to democracy, and central to public reasoning.

That was where public opinion, and public reasoning came to the fore. Rulers were to be dependent on the public through elections and constitutional reform; but also through the complementary operation of public opinion: formalised through what Bentham called the "Public Opinion Tribunal" (POT). The POT was a fictitious entity that Bentham used in his constitutional writing to explain the role of public opinion. As Rosen states, "Bentham's conception of the POT, on which reform depends, is not altogether clear" (Rosen, 1983:39). Nonetheless, the POT had some distinct features. It could speak for the public and serve the public interest. In particular, it performed a "censorial" function over the conduct of public officials. It rewarded them for good conduct, with approval; and punished them for bad conduct, with disapproval. In doing so, the POT created incentives for officials to serve the public interest rather than sinister interests. By definition, this increased their moral aptitude, on which the goodness of government itself depended (Bentham, 1989:4).

The POT could only cultivate official aptitude, however, when it was generated by the right kind of participation from members of the public. In Chapter 4, I refer to this as "public aptitude", and explore some of its features by drawing on a variety of different strands of Bentham's thought. The features of public aptitude indicate the kind of public reasoning that Bentham valued. It is critical of—rather than acquiescent to—authority. It focuses on citizens' own self-interest rather than the public interest. It examines language to test its relation to pleasures and pains, and expects to reach decisions through majority voting rather than deliberative consensus. Public reasoning of this sort both requires *and* cultivates the reasoning capacities that members of the public need to hold their officials to account. The

aim of this form of public reasoning is to furnish individuals with a “correct” conception of their own interests, and to scrutinise the conduct of officials in line with that conception: both through voting and the judgements of public opinion (Bentham, 1989:68). This requires scrutiny of language to militate against the delusions and fallacies that Bentham believed caused citizens to misunderstand their own interests. This view of public reasoning asks citizens to be critical of claims made by putative authority. It focuses on the self-interest of citizens, and analyses and reflects on the language used in discourse. “Public aptitude” engenders official aptitude. It is the hallmark of what Rosen calls “educated critical opinion” (Rosen, 2003:243). This view suggests that public participation in a form of pre-voting discussion is vital, but that discussion is different from the Habermasian or Rawlsian ideals. Participation directed toward “public aptitude” is focused on clarifying individuals’ perceptions of their particular interests rather than transforming them; it holds no intrinsic moral or democratic status independent of the benefits it confers on decision-making; and happens with no expectation of full consensus between participants.

In Chapter 5, I examine the contemporary relevance—and limitations—of the hybrid view of Bentham by considering some potential objections. At an interpretive level, the “hybrid” account corrects some misunderstandings about Bentham’s own views. It also has some implications for contemporary theory. It illustrates just how much common ground there is between deliberative and aggregative democratic approaches. That conclusion should cause us to question some of the critiques of respective theories levelled by writers from these competing approaches.

Principles of participation which foster public aptitude are not entirely “aggregative” (i.e. reliant solely on the aggregation of votes). Rather, they inhabit territory between aggregative and deliberative contemporary theories. Like some aggregative theories, the hybrid view is focused on the self-interest of participants rather than the public interest, and reflects existing preferences rather than transforming them. Like some deliberative theories, the hybrid view provides an account of how citizens should interact with one another and reflect before voting. This view therefore illustrates that there is far more common ground between aggregative and deliberative democratic theory than proponents of each view often suppose. Aggregative democrats often deride deliberationists as naïve believers in a set of ideal discursive conditions which may never obtain (Posner, 2003). Similarly, deliberationists can claim that their perspective is essential to understanding all but the thinnest conceptions of democratic procedure (Barber, 2003:198; Fishkin, 2008). Many of those critiques are oversimplifications. We should not approach public reasoning by assuming that all deliberative theories must believe in the full consensus espoused by Habermas. Nor should we assume that aggregative theories necessarily assume the self-interest of participants (List & Goodin, 2001); what Habermas himself called the “egocentric viewpoint” (Habermas,

1989:45). Rejecting that binary distinction allows us to acknowledge views like the hybrid Benthamite account.

This final chapter considers separately Bentham's views, and wider Benthamite principles which can take inspiration from him. Bentham's views could meet a number of objections from contemporary theorists. In Chapter 5, I consider four of these: the free rider problem, the danger of sinister interests within the Public Opinion Tribunal, questions over Bentham's optimism about public opinion; and a challenge of how to deal with what Dworkin (1978) called "external preferences". On many occasions, the nuances of Bentham's views provide convincing responses to these objections. Consider the "free rider" problem. Some critics suggest that an account of public reasoning premised on the self-interest of participants would allow some participants to gain the benefits of public reasoning without the costs of participating. However, Bentham's understanding of self-interest was expansive enough to mitigate that concern. It is likely that participation in public reasoning could create benefits for citizens which could only be experienced first-hand; not through free-riding. Other challenges are more difficult to counter. For example, Bentham did not give any comprehensive explanation of how the press—on which the POT depended—could itself be protected from corruption by sinister interests. This is an important, though surmountable, shortcoming.

Chapter 5 also considers some critiques of the hybrid Benthamite view which follows broad principles outlined by Bentham. Three particular objections are considered, regarding: power structures, the aggregative mechanism, and potential for a tyranny of the majority. There is a convincing Benthamite response to some of these. For instance, it has been suggested that aggregative views which reflect rather than transforming individuals' preferences reinforce existing power structures (Gutmann & Thompson, 2004). However, the kind of inclusive and critical dialogue that Bentham wanted to promote would encourage citizens to question existing power structures, so should counter this concern.

Further, the way that aggregative and deliberative principles are reconciled in the hybrid Benthamite approach is itself distinctive. Contemporary deliberative views often rely on contested notions of the public interest, or reasonable agreement. The hybrid Benthamite view, in contrast, merely encourages citizens to clarify and understand their own self-interests before voting. I believe that this has some advantages. Given the common liberal assumption that individuals should be considered the best available judge of their own interests—the "best judge principle" (Goodin, 1995:127)—it may be more difficult to manipulate citizens engaging in discussion about their own self-interest than those deliberating on wider

ideas. It should be more difficult to suggest that citizens have misunderstood their own preferences than to imply they have misconstrued the public interest, for example. Consequently, the hybrid Benthamite view is less susceptible to creating false consensus than many—particularly Habermasian—deliberative views.

The Benthamite view is also less susceptible to important critiques levelled against Rawlsian accounts of public reasoning. For instance, Phillips (1995) and others have contended that Rawlsian public reason is exclusionary. Definitions of “reasonable persons” and “public reasons” may in fact reflect existing orthodoxies and power structures. In turn, this means that discussion may not be open to the breadth of input from members of the community that it should be. It is therefore an advantage of the Benthamite account that it relies on no such notions. Rather, the POT was open to any individual with an interest in the matter being discussed. Further, it allowed a wide variety of modes of communication and participation.

This “hybrid” interpretation thus clarifies our understanding of Bentham in relation to contemporary writing on public reasoning, and strengthens the view that aggregative and deliberative democracy should be seen as compatible, rather than opposing types of approach. This view has been championed by contemporary theorists like Dryzek and List (2003).

Chapter 1. Two challenges for an account of public reasoning

In the coming chapters, we will explore some alternative accounts of how citizens should engage and communicate with one another before voting on important questions. I will refer to this general question as the debate on “public reasoning”. It concerns the way that members of the public, rather than politicians or designated experts, should reason and communicate with one another before they vote.

This chapter outlines two challenges for an account of public reasoning: effective discussion, and preventing manipulation by elites. These two challenges are not a comprehensive list of everything that we might demand from a theory. They are however of particular interest. This is because they illustrate some key aspects of how the contrast between aggregative and deliberative political theory is often set up by contemporary theorists. Implicitly and explicitly, it is common to find deliberative theories which suggest that mere aggregation ignores essential features of democratic procedure. Similarly, aggregative theorists sometimes allege that deliberation is associated with a set of paternalistic assumptions, which can be manipulated. On this view, some deliberative theory carries with it the insinuation that citizens reach the “wrong” conclusions in the real world because of their deliberative failings; or those of their environment (Posner, 2004). As that distinction is often defined, I believe that fundamental assumptions made by some aggregative theories reduce their ability to explain the value of effective discussion. Conversely, deliberative assumptions allow a convincing explanation of the value of effective discussion but also create opportunities for the manipulation of outcomes by elites in doing so. Freeman provides an excellent summary of why aggregative views may struggle to explain pre-voting discussion:

“individuals vote their private preferences and group interests . . . in effect they act like economic agents removed to a different forum. What point could there be in public discussion of their self-seeking and competing group purposes with others who have opposing interests?” (Freeman 2000, 373)

At one end of this spectrum, a purely aggregative view would understand democracy purely as a way of counting the preferences that individuals already have, rather than changing them through discussion. Clearly, a view of this sort leaves little work to be done by discussion between citizens before voting. Self-interested preferences, associated with some aggregative views, are thus often associated with counting votes rather than persuasion through deliberation. The fact of someone else’s preference is rarely considered a *reason* to change my own, particularly when theirs is a self-interested preference. It is often thought that self-interested preferences lack persuasive force in the same way that someone else’s preference for one type of ice-cream over another lacks the ability to persuade me to change my own preferences for different flavours of ice-cream.

At the opposite extreme, an account which makes very ambitious claims about the transformation of preferences that can be achieved through deliberation before voting may create more opportunities for outcomes to be manipulated through discussion. To explain this divergence, I will examine some potential views of public reasoning from different approaches to democratic theory. In particular, some of the differences between theories commonly thought of as ‘aggregative’ and ‘deliberative’ illustrate the tensions between providing a plausible account of discussion and mitigating opportunities for manipulated discursive outcomes. Of course both aggregative and deliberative approaches are diverse and complex; it is important to acknowledge that complexity. Nonetheless, the debate between these two approaches is common in the contemporary literature on pre-voting discussion. In particular, the accusation made against aggregative theories by deliberationists should cause us to consider whether an aggregative response is possible. Similarly, the accusation—made by Posner—that deliberative theories imply a kind of paternalism that may lend itself to elite manipulation should prompt investigation of deliberative replies.

A contrast between aggregative and deliberative democratic theories

Contemporary democratic theories all endorse the idea that citizens should talk to one another, and vote. In explaining these ideas, however, different approaches endorse different configurations of these two concepts and emphasise different aspects to different degrees. In broad terms, theories are often thought of as falling into two broad approaches: deliberative and aggregative. Simply put, deliberative theories primarily emphasise the importance of *talking* (under certain important conditions) between citizens before voting; aggregative theories instead emphasise the importance of *counting* votes to reflect individuals’ interests. This distinction is, of course, somewhat stylized because democratic theories are usually nuanced rather than exclusively aggregative or exclusively deliberative. Nonetheless, there are some important differences between these types of approach, and the use of this distinction is strikingly common in the contemporary literature.

Contemporary theorists often contrast aggregative and deliberative approaches by claiming that they wish citizens to vote according to different objectives. Freeman suggests that this is perhaps the clearest dividing line between the two approaches:

“...the relevant distinction...concerns the object about which citizens deliberate and vote and the kinds of reasons that they take into account in coming to their collective decision. Whereas an aggregative view counsels voting one’s informed preferences regarding one’s own good or partial group interests...deliberative democracy counsels voting one’s deliberated judgments (or informed preferences) for the common good” Freeman (2000: 337).

The object of voting is, however, just one of a number of dimensions over which accounts of public reasoning differ. More fundamentally, deliberation is often “intended to change the preferences on the basis of which people decide how to” vote (Przeworski, 1998: 140). In many cases, its aim is to enable citizens to approach consensus on what serves the public interest. In contrast, aggregative theories hold that laws should *reflect* individuals’ preferences, rather than being the product of a prior process of deliberation that changes them. Sometimes, but not always, this reflects Freeman’s distinction above: self-interested preferences are not held to have the same persuasive power as reasons presented during deliberation. Deliberative democracy holds the view that aggregation alone provides an inadequate or incomplete account of democracy. As Dewey said, the “counting of heads compels prior recourse to methods of discussion, consultation and persuasion” (ft. 2 in Knight and Johnson, 1994). Counting alone misses out the benefits of learning through persuasion between peers.

In many ways, some aggregative and deliberative democrats approach public reasoning from fundamentally different perspectives; the implications of this are far reaching. Beneath these broad differences, however, aggregationists and deliberative theorists disagree on more specific questions. Below, I illustrate this distinction by explaining four of these questions. Given the breadth and diversity of deliberative theory, I also distinguish between “stronger” and “weaker” deliberative views to acknowledge that responses to each of these questions vary significantly within each approach. As we proceed to investigate the details of different accounts, we will discover more varieties of both approaches, and perhaps more common ground between them. Broadly speaking, however, the aggregative view suggests that it is essential for public decisions to be “positively responsive to the interests [preferences] of each person” (Cohen, 1999: 186), usually construed as their preferences as expressed at the ballot box. Particular theories across the boundaries explained above vary in respect to the focus, transformative power, status, and output of public reasoning. Some responses to these questions from different theories are outlined below to illustrate.

First, consider the focus of public reasoning: what are citizens asked to reason and vote according to? Deliberative democracy is usually defined in opposition to expressions of self-interest, power relations and bargaining between participants. It typically demands that democratic processes “are oriented around discussing [the] common good rather than competing for the promotion of the private good of each” (Young, 1996:120-1). Many deliberative theories thus demand that citizens should vote according to what they believe will best serve the “public interest”, which may be defined in a number of ways:

“Usually, ‘the public interest’ is taken to describe a good that is common to an entire society, rather than to some particular section of it” (O’Flynn, 2010:300)

In order to cast their vote in this way, citizens need to deliberate together before voting in order to understand the common interest and what it requires. These approaches therefore ask citizens to learn about the public interest. This may be defined in a number of different ways, but is always more than “mere” self-interest. Deliberative theories thus often hold that “instead of reasoning from the point of view of the private utility maximiser, through public deliberation, citizens transform their preferences according to public-minded ends” (Young, 1996: 120-1). This is reflected by the fact that “public spiritedness” is commonly an important condition of deliberation. We reach an understanding of the public interest through talking together, with something of a shared perspective, rather than simply the counting of votes.

Different deliberative accounts expect different degrees of commitment to public orientation from participants. Some deliberative theorists explicitly acknowledge that citizens may be motivated to deliberate by their own interests (Manin, 1987). Others hold that a public focus from citizens is part of the definition of properly conducted deliberation, meaning that partial or self-interested claims should not make it into the deliberative forum. This public focus is a common part of the type of public justification of arguments, in front of one’s peers, that many contemporary democrats believe is essential to good deliberation (Goodin, 1995: 143). The benefits of deliberation are held to depend on a variety of conditions, of which a move beyond the mere self-interest of participants is one. Habermas—and the deliberative tradition he inspired—distinguished his own approach from aggregative theory by contrasting self-interests and objectives beyond them.

Habermas argued that when the ideal terms of deliberation obtain, participants move from strategic action, aiming merely at “success”, to communicative action, which is the cooperative search for mutual understanding. The latter aims at an agreement that is justifiable by reasons in light of the generalisable interests of all relevant parties, expressed through their agreement. One of the conditions for this kind of interaction is discursive equality: each participant is given an equal opportunity to introduce or question any assertion and to express attitudes, desires, and needs (Habermas, 1990:89, 1996: 305-306, 308, 2001: 34). The aim of this kind of deliberation is the generalisable interests of individuals, reached through discursive consensus rather than the aggregation of individual interests through voting. It is a precondition of Habermasian deliberation that participants must overcome their “egocentric viewpoint” (Habermas,

1989: 45). I will refer to Habermasian deliberation as a “stronger” deliberative view, in part, because it firmly requires a public-spirited—rather than self-interested— participation from citizens.

Of course, there are important variations between different conceptions of deliberative democracy. Some deliberative views place much less stringent requirements on participants to reflect generalisable interests. I will mark these out as “weaker deliberative” views. Take Gutmann and Thompson, for example, who argue that one of the purposes of deliberation is to “encourage public spirited perspectives on public issues... by encouraging participants to take a broader perspective on questions of common interest” (Gutmann & Thompson, 2004:10-11). These authors value deliberation in part for its public orientation, but they also acknowledge its ability to express and clarify self-interested and other claims. For them, deliberation is valuable both because it can produce decisions which better reflect the public interest, but also because it acknowledges something more basic: a duty of reciprocity between citizens. Although intended to encourage a public-orientation amongst participants, this kind of deliberation also leaves room for citizens to express self-interested claims so long as the reasons presented to justify them are public and accessible to others (Gutmann & Thompson, 2004: 4-5). “Weaker” deliberative views leave more space for the inclusion, consideration and expression of self-interest before voting. In this vein, Mansbridge has also made similar, and convincing, modifications to the deliberative project in a series of influential recent works. In 2010, she joined other renowned deliberative democrats to argue that a new approach was required towards the role of self-interest in deliberation. Rather than seeing self-interested claims as inimical to properly functioning deliberation, deliberative democrats should agree that self-interest, “suitably constrained, ought to be part of the deliberation that eventuates in a democratic decision”. Expressions of constrained self-interest during deliberation should thus feature prominently in a reformulation of “the regulative standard to which real deliberations should aspire” (Mansbridge et al, 2010: 64). “Weaker deliberative” views of this sort still intend deliberation to produce decisions which reflect more than the mere self-interest of those involved. But they tend to stipulate appropriate conditions for the expression of self-interested claims during deliberation rather than encouraging participants to overcome an egocentric viewpoint.

At one end of the deliberative spectrum, we thus have stronger theories. These have been modified by “weaker” views, more accepting of an appropriately constrained role for self-interest during deliberation. We therefore see a diversity of deliberative views in respect to orientation. Partly for this reason, a distinction between aggregative and theories, like Freeman’s, which is based on orientation alone cannot tell the whole story. Traditionally, the benefits of deliberation have been associated with over-coming mere self-interest, particularly by those following the “stronger” Habermasian view. “Weaker” deliberative democrats like Mansbridge, and Gutmann & Thompson have, however, amended that view.

Indeed, aggregative theories do not necessarily require self-interested participation from citizens. Consider the (aggregative) social choice tradition. A fundamental assumption, and the fourth of Arrow's conditions, of this approach "expresses fully the idea that social choices are determined by individual desires [citizen sovereignty]" (Arrow, 1963:29). Those desires may be collective or personal, depending on the details of a particular theory. These details vary, so there is a great diversity of aggregative approaches. I shall attempt to acknowledge this by considering two broad categories. "Rational choice" aggregative theories assume that self-interested preferences should be aggregated through a social decision mechanism. In contrast, "social choice" aggregative theories may assume either that preferences are self-interested, or reflect some wider considerations. This distinction also requires a caveat. Some rational choice theorists (Downs, 1957) endorse an ordinal view of preferences. This means that preferences can be considered simply individuals' orderings of possible states of affairs. Views of this kind may be ambivalent about the basis on which individuals rank possible states of affairs.

What aggregative views do have in common, however, is that they do not begin with an exogenous account of the public interest, and assess individual preferences by reference to it. For Arrow, doing this would require us "to assume, with traditional social philosophy of the Platonist realist variety...an objective social good" (Arrow, 1963:22). Instead, we should trust individuals' preferences as a reflection of the interests in question. On a deliberative account, of course, valuable preferences must be the result of a transformation during deliberation under the right conditions, with the right information available. They must result from the right kind of procedure, or have the right kind of justification in underlying reasons. These kinds of conditions are traditionally absent from aggregative views.

Among aggregative theories, the "egocentric viewpoint" criticised by Habermas is often received with far less, if any, opposition. Consequently, aggregative views are often also less committed to demands for public-spirited deliberation. Riker (1982), for example, models democratic choices in terms of the actions of self-interested individual utility maximisers. In doing so, he never suggests that the individuals involved in collective decisions should move beyond their own interests and towards the public interest. Similarly, rational choice theory starts from an assumption that "individual behaviour is motivated by self-interest, utility maximisation, or, simply put, goal fulfilment" (Pettracca, 1991: 289).² Using these assumptions, it then seeks to understand how individuals' pursuit of their own goals plays out in different strategic situations. Two assumptions are central to this kind of analysis. First, methodological individualism implies that outcomes can be understood in terms of the preferences, motivations and

² Again, a caveat applies here regarding rational choice theorists who espouse an ordinal conception of utility.

actions of individuals rather than groups or communities. Second, individuals are assumed to be rational in pursuit of their goals:

“We assume that every individual, though rational, is also selfish... Throughout our model, we assume that every agent acts in accordance with this view of human nature. Thus, whenever we speak of rational behaviour, we always mean rational behaviour directed primarily towards selfish ends” (Downs, 1957: 27, in Petracca 1991: 294-5)

Despite the diversity of aggregative and deliberative views, a basic contrast here is clear. “Weaker” deliberationists like Gutmann & Thompson are more permissive to self-interested arguments and reasoning than Habermasian counterparts. Even they, however, nonetheless stipulate that arguments presented must assume a “moral point of view”, so “presuppose a disinterested perspective” (2004:72) to be suitable for the public forum. This condition is similar to the “stronger deliberative” Habermasian view in that it asks deliberators to move beyond self-interest. Unlike Habermas, however, they do not expect that even ideally conducted deliberation should produce full consensus. They are more accepting of lingering disagreement after deliberation; as well as self-interested claims before it. Gutmann & Thompson’s “weaker deliberative” view thus accepts that there will be persistent fundamental disagreements in society owing to epistemic factors similar to Rawlsian “burdens of judgment” (2005: 54). We will return to this point below.

The shared assumption behind both “stronger” and “weaker” deliberative approaches is “epistemic moralised” deliberation: that discussion is valuable partly because it encourages citizens to consider the questions impartially, rather than reflecting mere self-interest. The importance of public-spiritedness in deliberative theories is reflected by the kinds of conditions they stipulate for ideal deliberation. In turn, these reflect what is necessary for a full transition from “mere” interests, to deliberative “reasons”. The presentation of these reasons has the capacity to persuade fellow citizens and change their preferences rather than merely informing or clarifying them. Public orientation is thus integral to epistemic moralised discourse, and the potential consensus it is expected to achieve.

The use of a self-interested orientation in rational choice analysis has led to a perception amongst some commentators that aggregative theories necessarily understand citizens’ democratic behavior in self-interested terms. In reality, aggregative democracy is not quite so simple. Whilst some rational choice models may assume self-interest, aggregative models for understanding democratic decisions need not do so. For example, Hylland acknowledges that, according to social choice theory, “preferences can be based

on egoistic concentration on private consumption, on altruism, on ideal principles of ethics, or on any combination of these factors” (Hylland in Elster, 1986:53). More recently, Dryzek and List (2003: 2) have emphasized this point. They reaffirm that there is a clear distinction between conducting a logical and normative analysis of aggregation mechanisms (social choice) and any behavioural assumptions about individual beliefs or preferences. One may model the aggregation of preferences without necessarily making particular assumptions about the individual behaviour or motivation which produces them. Social choice is a group of models for understanding how the democratic inputs provided by individuals, which might be preferences, votes, or something else, are aggregated to produce collective decisions.

It follows from these distinctions that discussion before voting under rational choice models (with the caveat that some are ordinal) will be geared towards informing voters’ self-interested preferences rather than the public interest. In contrast, aggregative theories that depart from the assumption of self-interest need not conceive of pre-voting discussion in such self-interested terms. The orientation of broadly aggregative theories lends itself to information sharing to inform how existing preferences cause citizens to vote. The orientation assumed by many deliberative theories, in contrast, lends itself to persuasion, using reasons, in a particular way. Orientation therefore implies something about whether preferences can, or should be transformed before voting. This takes us to the second dimension in which public reasoning accounts vary: transformative power.

The differences between aggregative and deliberative theories are particularly clear-cut when it comes to views put forward about the transformative power of deliberation before voting. We have already seen that some deliberative theories require a more public orientation from individuals than aggregative counterparts do. This also implies the possibility of expressed preferences changing during deliberation in response to the reasons and arguments presented. On most deliberative views, this change, or transformation, is essential because the exchange of reasons in justification of arguments presented before voting is valuable in part because it can provoke citizens to reconsider their prior views and judgments (Gutmann & Thompson, 2004: 13). Many contemporary deliberative accounts place preference change at the centre of democratic theory. For Knight & Johnson, for example, deliberation is marked out from other types of communication precisely by its capacity to transform “the substance of participants’ preferences” (1994: 282). The idea of preference change is so fundamental to contemporary conceptions of deliberation that a democratic view could scarcely be called “deliberative” without it. Similarly, Dryzek states that deliberation “as a social process is distinguished from other kinds of communication in that deliberators are amenable to changing their judgements, preferences, and views during the course of their interactions, which involve persuasion rather than coercion, manipulation, or deception.” (Dryzek, 2002: 3)

In part, preference change flows from a public-orientation demanded from citizens. It follows that since individuals are required to vote on the public interest—or at least something beyond informed self-interest—they benefit from learning more about it through deliberation before voting. “Epistemic moralised” deliberation is required: to have due regard to the interests of my nation, I must learn from fellow citizens and move beyond my merely self-interested “raw” preferences. In some accounts, this entails that citizens have a duty to appeal only to certain types of reasons when deliberating together on important questions. The appropriate conditions for deliberation are thus understood to limit the content of how citizens may justify arguments in their attempts to persuade others while deliberating. Theories of this type rely on an assumption that there is a “set of public reasons that ought in principle to be shared by all” (Bellamy, 2009: 103). Rawls suggests that citizens have a duty to refer only to “public reasons” when debating “constitutional essentials and matters of basic justice” (2005:214). These reasons are “public” in the sense that reasonable persons can affirm them from within their own “reasonable comprehensive doctrines” (2005:58). The Rawlsian duty of public reason thus clearly prescribes some conditions for pre-voting discussion. Constraints of this sort can be justified either in reference to the reciprocal duties that citizens owe one another, or in terms of the improvements they produce in the quality of decision making by a community. Gutmann & Thompson, for example, appeal to both intrinsic and instrumental reasons in defence of their formulation of deliberation (2004: 79).

Public orientation and preference transformation go hand-in-hand. For many deliberative theorists, part of what makes deliberation valuable is that norms associated with it promote a “concern...for the public interest” among citizens (Sunstein, 1993:246). Conversely, a departure from the first of these claims accompanies a move away from the second in many aggregative views. Preference change through deliberation is thus essential to deliberative democracy, but inimical to traditional aggregative democratic approaches. We have already briefly touched on social choice theory as a family of models to understand how individual preferences or votes should be combined to produce collective decisions. Accounts of this kind analyse ways of aggregating individual interests, or preferences, into social outcomes rather than explaining how they should transform before voting (Knight & Johnson, 1994: 279). Indeed, modelling the aggregation of preferences would be extremely difficult if those preferences were considered something of a “moving target” rather than fixed inputs.

The lesser emphasis placed on aggregation by deliberative theorists means they positively wish this “target” to move. This reflects a reorientation, to some degree at least, of the individual away from her

own mere self-interest, and towards the community interest, Like Sunstein, Fishkin—a leading deliberative democrat—cites this as a key benefit of deliberation:

“when citizens discuss public problems together they come to place greater value upon the interests of the broader community” (Fishkin, 2009:103).

Of course, preference transformation also implies that preferences can be “improved” in some sense. The transformation of preferences during deliberation thus reflects the deliberative theorist’s willingness to step back from and evaluate individuals’ preferences. The claim that a post-deliberative preference $A > B$ is “superior” to pre-deliberative preference $A < B$ cannot be explained simply by reference to preferences themselves: that would be circular. Instead, deliberation causes an individual to hold a preference that is superior, or preferable in some way. This superiority can be explained either in (substantive) terms of the public interest, or by the discursive terms that produced it, because epistemic moralised deliberation was conducted in the “right way”. So, here again, the public interest plays a pivotal role in many accounts of deliberation: it provides one measure by which the value of preferences can be assessed. A deliberative democrat may claim that discussion has worked because participants have arrived at a preference that better serves the public interest; for example, $A > B$. Theoretically, at least, individual votes can thus be assessed as good or bad reflections of the public interest. The value of deliberation can be explained by its ability to improve preferences and promote convergence on measures that benefit all, rather than just some individuals or groups. An “epistemic moralised” conception is thus often thought necessary to explaining discussion’s importance to democracy, and preference change.

In contrast, aggregative theorists model preferences by assuming that “...agents are... endowed with preferences that are...given and not subject to change in the course of the political process” (Elster, 1986:105). Without some fixed or defined set of preferences to model, it would be difficult to explain and understand different ways in which they can be aggregated to produce collective decisions. Democracy is thus considered more of a “marketplace”, where citizens express their preferences through votes; like consumers purchasing products. These “consumers” may research the products on offer, but will not fundamentally change their preferences.

Views on orientation and transformation tend to complement one another. Accounts that demand voting on the public interest demand more learning by voters, and greater potential for learning, through preference change. The converse is true of accounts that ask citizens to vote on self-interest. If, as some deliberative views contend, the public interest requires equal regard for the interests of all members of the

community, then it is plausible to assume that individual voters can learn from one another. Unlike aggregative citizens, deliberators will search for more than mere information on how different policies or priorities affect their existing preferences. Those preferences may or may not be self-interested; depending on the type of aggregative or deliberative account espoused. Deliberators will learn from other citizens' experiences, changing irrational views in light of their contributions to debate. The act of presenting reasons, and justifying expressed preferences is thus essential to most deliberative accounts, but peripheral to aggregative views. Instead, aggregationists conceive of pre-voting talk as a way of informing existing preferences. Talk between citizens has an immense, transformative power for deliberative democrats. It can change their judgements, values, and world view. For many aggregative theorists it is, at best, a source of information for agents already endowed with their preferences. The transformation of preferences is made possible by the presentation of reasons that persuade others to change their minds. In contrast, aggregation may model the exchange of information between citizens, but not transformations in their fundamental underlying preferences. This leads us to the third way in which public reasoning accounts vary: status.

As we have seen, most views of public reasoning claim that there are some benefits from discussion before voting. Aggregative theories generally identify some instrumental benefits that are created when individuals are given better information before they express their preferences. Most deliberative democrats also cite instrumental gains. These may include the further benefits of preference change as citizens reflect on their own views in light of the new arguments and evidence presented to them. For differing reasons, therefore, and in slightly different ways, ascribing an instrumental importance to this kind of discussion is a common feature of most democratic theory, whether theorists understand democracy in broadly aggregative or deliberative terms. Whilst a belief in the instrumental value of discourse is common to most democratic views, deliberative theory goes further than this. It often ascribes an important intrinsic value to the right kind of deliberation.³

Individual aggregative and deliberative theories vary significantly on whether discussion has a further "moralised" value beyond its ability to improve outcomes. It is clear that aggregative theories take a functional approach. They usually understand the value of discussion exclusively in terms of the improvements it creates to decisions through increased information. Epistemic democrats like Estlund, for example, argue that participation in democratic deliberation is valuable primarily because of its ability to create better decisions (Estlund, 2008:23). For them, the central question is the instrumental efficacy of particular ways of making collective decisions rather than the character of the discursive procedures used

³ Rawls (2005) provides an intrinsic view: public reason is what reasonable persons owe one another in debate. For a contrasting instrumental view see Arnson (2004). Gutmann and Thompson (2004) combine intrinsic and instrumental arguments for deliberation.

(Coleman, 1989:196). Similarly, Ober, for example, suggests that the value of discussion resides in the part it plays in “the organization of knowledge (accurate information and true beliefs) rather than simply preferences or interests” (2011:1).

In contrast, some deliberative views ascribe different, non-instrumental, value to public reasoning. This can take a number of different forms. In many theories, these types of value may overlap, coincide or complement one another. Specifically, deliberation can be said to reflect, create, or constitute democratic values. Consider the first of these: the reflection of democratic values. Rawls’ (1997) theory of public reason has come to define much contemporary debate on this question. He suggests that citizens hold a duty towards one another to appeal only to reasons that are justifiable to all reasonable persons. On this view, the right kind of deliberation between citizens has intrinsic value: it reflects what reasonable persons owe one another. What Rawls calls the “natural duty of respect” is reflected by there being a shared basis for determining the principles that it will uphold. In line with this, citizens respect one another as individuals when they take account of one another’s claims and interests and act on reasons which they are “prepared to explain to them in the light of mutually acceptable principles” (2003: 373; see Rawls, 1997:156). Mutual respect is, therefore, a good that can be achieved only through the principles of justice that meet the reciprocity condition being publicised and understood by those individuals subject them. This argument goes far beyond merely asserting that well conducted deliberation produces better results; it ascribes a further, “moralised” value for the right kind of deliberation between citizens. Rawlsian public reason is intimately related to a wider, and very ambitious, normative project in political theory.

The idea that reasoning together on democratic questions reflects duties that citizens owe one another is not unique to Rawls. As we have seen, Gutmann & Thompson suggest that deliberation fulfills a similar principle of reciprocity. That “core principle of democracy” implies that “citizens owe one another justifications for the mutually binding laws and public policies they collectively enact” (Gutmann & Thompson, 2004: 98). Deliberation thus acknowledges important duties between citizens. Taking collective decisions without appropriate deliberation would be to neglect what citizens are owed by one another and their community. As well as reflecting democratic values, public reasoning may also be held to create them.

On this view, a deliberative process of reason-giving, as well acknowledging duties, is also required for laws to be legitimate and just (Ibid:101). Properly conducted deliberation therefore grounds two of the

most significant values to be found in a democracy; it both reflects duties between citizens and creates new ones by generating laws that are binding. This idea, that appropriate deliberation is required for democratic legitimacy is common to many deliberative theories. Goodin has thus gone so far as to generalise that deliberative theories “suppose that outcomes will be democratically legitimate only so far as they emerge through [an] external-collective processes of deliberation involving a free and equal exchange” (Goodin, 2000: 82). Similarly D’Agostino—like countless other contemporary theorists—considers a variety of different models of public reasoning to examine whether they can provide an account of how, and whether political regimes are justified to their citizens (D’Agostino, 1996:57). Alongside this notion of justification and legitimacy, other deliberative democrats claim that deliberation has the capacity to cultivate and improve an “excellence of character” amongst those involved; a quality amongst the citizenry which “permits a democracy to flourish” (Gutmann & Thompson 2004:79). I do not wish to suggest that every deliberative theory does, or must, see political legitimacy in terms of particular type of public reasoning before voting; many do not. It is clear, however, that this kind of moralised claim is common to many deliberative views and goes far beyond the more functional value attributed to pre-voting discussion by aggregative theorists.

In addition to creating or reflecting a particular kind of democratic value, public reasoning is also sometimes held to constitute a democratic value itself. Habermas, for example, argues that the value of deliberation is only fully realised when certain ideal conditions obtain. These are premised on the demanding view that ideal deliberation takes place “Only in an egalitarian public of citizens that has emerged from the confines of class...and exploitation” (1997:308).⁴ This implies that the results of deliberation enjoy a particular legitimacy or value, but also that this kind of discussion between citizens itself actually constitutes what is valuable about democratic governance. For Habermas, appropriate discourse between citizens generates legitimate outcomes because norms owe their legitimacy to a kind of recognition that is only found in rationally motivated agreement (1999:940). These norms could not be generated or defined without the right kind of participation from citizens. By stipulating discursive conditions of this sort, deliberationists like Habermas buttress their claim that deliberation serves an important moral function. Decisions reached under ideal deliberative conditions are thus thought to enjoy a special legitimacy or value, by definition. This “status” is often tied up with the type of agreement expected from deliberation. This brings us to the final dimension of public reasoning accounts.

Views of public reasoning also vary in the degree to which they expect citizens to reach consensus following discussion. Broadly, there are two general approaches to this question: consensus and

⁴ The assessment of a difficult claim like this is hardly likely to be uncontentious in any real world democracy.

convergence. Some views orientate discussion around the public interest (O’Flynn, 2010:300) or generalisable interests whilst others are directed towards the self-interest of participants. Alongside these variations, theories also vary in their view on the reasons that individual citizens can, or should, have for agreeing to principles for collective decisions. To illustrate, let us consider thinkers who examine deliberation in reference to wider normative questions in political theory. Consensus and convergence can be understood in relation to these normative questions—see above—or in isolation from them. What matters for our immediate purpose is the type of reasons that citizens are expected to have for agreement, rather than the normative implications of their doing so.

Consensus views generally wish and expect deliberators to reach shared conclusions for the same, shared reasons. As we have seen above, Habermas, for example, argues that “insofar as norms [created by discourse] express generalizable interests, they are based on a rational consensus” (Habermas, ft. 71 in D’Agostino, 1996: 49). Aspects of Rawlsian public reason also suggest a consensus view of deliberation because, on constitutional essentials and matters of justice at least, it aims to secure agreement between citizens based on shared, public reasons. Rawls defines this kind of shared public reasoning in contrast with alternatives through which citizens could reach agreement for different reasons, for example merely “as the result of political compromise” (Rawls, 2005: 218).

In contrast, convergence views of public reasoning expect agreement, when it does emerge, to be reached by different citizens for different reasons. Ackerman, for instance, argues “justification is to be found in the web of talk that converges upon it from every direction” (Ackerman, quoted in D’Agostino, 1996: 49). Similarly, many commentators assume that a utilitarian view of pre-voting discussion would also allow participants to agree for different reasons. Generally, utilitarianism suggests that actions and processes are right insofar as they promote happiness; often construed as the satisfaction of preferences. After all, it seems plausible that a utilitarian perspective should be interested in the fact that citizens reach agreement, and the effects of this agreement, rather than the character of their reasons for doing so. Of course, we shall return to details of a possible utilitarian response to public reasoning in subsequent chapters.

So accounts of public reasoning vary in regard to the reasons they expect individuals to have for reaching agreement through pre-voting discussion. They also vary in the degree to which they expect agreement to emerge when deliberation is done in the right way. In some cases, public-spirited deliberation is expected to cause citizens to reach agreement as more people realise what “truly” reflects the public interest. This view of deliberation combines the dimensions we have considered. It suggests that reasoning together provokes public-spiritedness (orientation) and preference change (transformative power) because it makes unacceptably partial or biased preferences give way to “those that can form the basis of a public

consensus” (Gaus, 2003:148). Historically, an expectation of consensus of this sort has been expressed by great Enlightenment political philosophers.⁵ It reflects a view that there are reasons which are discoverable to all through rational discourse. Views of this sort tend to also be perfectionist in nature; they spell “out a set of controversial metaphysical and ethical doctrines concerning the nature of value and the good life... then recommend[s] political principles built upon these” (Nussbaum, 2011: 3). The shared, discoverable reasons, available to all, are thus those which reflect these principles.

Many contemporary theorists reject the idea that discussion should be understood in these perfectionist terms. This rejection is perhaps most clear in Rawls’ work on public reason which makes a conscious and explicit effort to depart from an historic approach, which conceives of “people in a standard, or canonical fashion so that they might all accept the same kind of reasons” (Rawls, 2003:368). In other cases, the rejection of perfectionism by deliberative theories takes place through an insistence that deliberation should be understood in terms of its procedural rather than substantive value.

Despite these significant differences, contemporary deliberative democracy does sometimes share with Enlightenment liberalism an expectation that properly conducted discussion before voting should yield agreement. Often, these views imply that deliberation will cause citizens’ judgments to align on particular answers that best represent the public interest, however defined. As Gaus (2003:144) notes, some deliberative theories hold that we should reach “full consensus” on key questions through deliberation, and that this should take place under the right discursive conditions. The term “consensus” comes from the Latin verb “consentire”, which means “agreed”. It implies that a collective decision is reached through agreement, rather than closed off by through voting.⁶ In contemporary use, it has also come to take on wider meanings. For example, Rawlsian “overlapping consensus” entails agreement for the same reasons (1997). Following the themes mentioned above, I will refer to those deliberative views which expect deliberation to produce consensus on important decisions “stronger deliberative” accounts. They make more extensive claims on behalf of deliberation, including the expectation of full consensus between participants, as well as giving deliberation a greater normative role and, often, permitting less of a role for self-interest in pre-voting discussion.

“Stronger” deliberation before voting, therefore, plays an important role. Part of this is to allow citizens to approach consensus, which itself generates, reflects, or constitutes important democratic norms in many theories. Since we all deliberate together regarding the same question—what is best for society?—this kind of consensus may seem plausible. I may enter a process of deliberation with a selfish preference for

⁵ Gaus (2003) provides an interesting analysis of how convergence conceptions of public reasoning have developed over this longer period, and how they contrast with contemporary equivalents.

⁶ <http://www.oxforddictionaries.com/definition/english/consensus>

building a new opera house in my home city; based on the fact that I personally enjoy opera more than football. I may finish deliberation understanding that a new football stadium better serves the public interest because more people will enjoy it, or it produces greater economic benefits for the community.

"Stronger deliberative" accounts thus broadly follow the work of Habermas in respect to consensus as well as orientation. In contrast, "weaker deliberative" views (Gutmann & Thompson, 2004) are more ready to accept that even well conducted deliberation may not yield full consensus. Of course, I am conscious that the different dimensions we have outlined are differently configured and emphasised in every individual account of public reasoning. A deliberative theorist may thus be "stronger" in respect to orientation, but "weaker" in regard to consensus. Nonetheless, while acknowledging this complexity we need to make some generalisations in order to identify important trends in the literature.

In that vein, let us turn to consider aggregative views, and how they regard consensus. Those aggregative theories which accept the behavioural assumptions of rational choice theory may ask voters to vote according to what they believe is good for them as individuals.⁷ Since the "public interest" is taken to be merely the aggregate of individuals' interests, voting in this way better reflects the public interest than any other mechanism. The aggregation of individuals' interests, expressed through voting, thus takes us to the public interest, without requiring voters to consider anything more than what is good for them. Public interest decisions are reached by counting votes, not talking. Logically, this makes agreement before voting unnecessary. An "epistemic prudential" conception of pre-voting discussion is thus implied: individuals should learn about their own interest before voting, rather than the "public interest" or interests of all. This is illustrated by a paradigm case of aggregative democratic theory: social choice. Here, various ways of aggregating individuals' preferences—taken to represent their interests—are modeled. Unlike deliberative theories, the emphasis is on how individual preferences can be aggregated to produce a collective decision rather than a process of deliberation that provokes citizens to agree or before they vote, on a shared understanding of what serves the public interest (List, 2006). There is little consideration of consensus in this view because aggregation resolves disagreement. For example, some writers have focused on the "miracle of aggregation", which means that good decisions can be reached through weight of numbers; even if many electors have mistaken views before voting (Page & Shapiro, 1993).

⁷ It is possible for a theory to focus on the aggregation of "public spirited" preferences, but for clarity I focus on the idea of aggregating self-interested preferences. This follows the caveat that aggregation may refer to ordinal preferences.

To summarise, I have tried to identify some trends found across a wide variety of writing on public reasoning while also acknowledging the degree to which individual accounts depart from these. Broadly, aggregative and deliberative views differ in how they respond to four important questions: focus, transformative power, status, and output. Each account is different, and has its own configuration of positions on these dimensions. Consequently, a degree of judgement is inevitably necessary to determine which category best characterises a given account. Different judgments could be made. Though imperfect, these categories allow us to focus on some important trends and contrasts in contemporary writing on public reasoning. I summarise these below, in Table 1. Broadly, “rational choice” aggregative theories assume a self-interested orientation from citizens. They reflect preferences rather than transforming them; do not ascribe a special status to discussion; and do not expect consensus before voting. “social choice” aggregative theorists vary in respect to orientation. They more readily model preferences that are not self-interested because they do not necessarily make assumptions about the values citizens use to order states of affairs to produce preferences.

On the deliberative side, “weaker” deliberative theorists are relatively permissive of self-interested arguments during deliberation and voting. They see the clarification of citizens’ existing positions as a key part of deliberation, as well as the transformation of preferences. They may or may not expect full consensus, or believe that deliberation has a particular and special normative value. Finally, “stronger” deliberative views do expect consensus before voting; and for it to be realised through transformation of preferences during a kind of deliberation that moves well beyond citizens’ mere self-interest. Often, they ascribe to that kind of deliberation a particular normative value.

While imperfect, this taxonomy allows us to look across some of the terrain on which the contemporary debate takes place, while acknowledging that differences across categories can often be those of emphasis rather than absolute opposition. In the next section, we shall move on to explore some normative implications of these distinctions.

Table 1. Contemporary accounts of public reasoning

	<u>“Rational choice” aggregative</u>	<u>“Social choice” aggregative</u>	<u>“Weak deliberative”</u>	<u>“Strong deliberative”</u>
<u>(a) Focus</u> (what are citizens asked to vote according to?)	Self-interest	Either self-interest or public interest	Self-interest or public interest	Public interest
<u>(b) Transformative power</u> (does discussion before voting transform or reflect preferences?)	Reflect	Reflect.	Reflect and/or transform.	Transform.
<u>(c) Status</u> (Does discussion play a moralised role?)	No.	No.	Often moralised	Moralised
<u>(d) Output</u> (Is consensus expected?)	No.	No	Only sometimes.	Yes
<u>Examples</u>	Riker, rational choice	List, social choice	Gutmann & Thompson, Fishkin, List	Habermas, Dryzek.

Two diverging challenges: discussion and avoiding manipulation

Let us consider how other examples of public reasoning accounts could fit into this structure. Using the categories above, we can see some important variations. In turn, these variations have significant implications for how well particular theories can respond to a variety of normative questions. I want to focus on two of these normative challenges in particular: providing a plausible understanding of the value of pre-voting discussion; and limiting opportunities for the outcomes of discourse to be manipulated.

These questions are at the centre of the contemporary debate. So often, deliberative theorists dismiss aggregation because it underplays the role, status, or value of pre-voting discussion. Conversely, aggregationists can be dismissive of deliberation. They sometimes suspect that the notions used to explain the deliberative ideal can be a veil for theorists' desire to achieve particular outcomes. Underlying this criticism is the suspicion that accounts of deliberation may mask an underlying paternalism about how citizens' preferences should be "improved".

Having surveyed a variety of contemporary public reasoning accounts, I believe that accounts which: focus more on the public interest; espouse transformative deliberation; give discussion a moralised role; and expect it to achieve consensus are better placed to provide a convincing account of what constitutes effective discussion before voting. This is largely because they give that discussion a more important normative role in democracy. The theorist who bases justice, legitimacy, or other important values on a foundation of democratic deliberation should have a convincing response when asked: "why is public reasoning important?" This holds even if the basis for her views on those critical values is subsequently rejected. In contrast, theories which deny those premises by directing discussion towards: self-interest; information rather than preference transformation; a functional rather than moralised role; and have no expectation of achieving agreement through discussion have fewer responses to the same question. If aggregation achieves good outcomes, why devote too much effort to talking beforehand?

Deliberative theories have a more defined view of what discussion "should" be. They often describe what it should produce (consensus), and how it should work. Sometimes, this more defined view is also linked to specific outcomes. For example, Gutmann & Thompson argue that a "deliberative deficit" helps to explain society's failure to reach conclusions they advocate on matters like healthcare and welfare (2004: 12). I believe that arguments like this leave deliberative theorists open to a charge of elitism and paternalism. Where theorists establish "ideal" conditions of deliberation, and link these to substantive conclusions they themselves endorse, this seems to imply that other would agree: *if only* they deliberated "properly". I believe that views of this sort leave citizens more susceptible to manipulation because they establish—sometimes quite prescriptive—expectations around the way that citizens should engage with one another. In contrast, aggregative theorists are generally more willing to simply trust the preferences that individuals express as a reflection of relevant interests. Some of the normative claims made by theories to explain the importance of deliberation also create opportunities for outcomes to be

manipulated. In general terms, accounts are less susceptible to allowing discussion to be manipulated when they: espouse self-interested discussion before voting; see the value of that discussion predominantly in non-transformative terms; do not believe it holds a special value; and do not expect consensus. A less ambitious account of pre-voting discussion may be flawed in other ways, but it has the advantage of offering less scope for manipulation of preferences during discussion. In contrast, accounts create more potential opportunities for outcomes to be manipulated when they give deliberation an important public-interest-focused role, through preference transformation achieved on moralised terms, which should reach full consensus.

In short, stronger deliberative theories are good at accounting for the value of discussion before voting but can be vulnerable to manipulation. Aggregative theories are often poor at accounting for the value of discussion but provide fewer opportunities for manipulation. These two challenges therefore provide an interesting insight into the range of available theories: neither the stronger deliberative nor traditional “rational choice” aggregative can adequately answer both. Exploring accounts of discussion and manipulation may therefore teach us more about the distinction between aggregative and deliberative theory which is so commonly sketched out by contributors to the literature on public reasoning. Let us focus on the first of our challenges: effective discussion. Why might aggregation be considered to provide an inadequate account of the value of pre-voting discussion?

Our first challenge is the principal deficiency identified in aggregative democracy by its deliberative critics. An accusation commonly leveled at aggregative theories is that they fail to provide a plausible account of the importance of discussion before voting. Barber makes this criticism when claiming that:

“Talk to these thinkers is little more than a... preliminary to...choosing...to limit democracy to selection among preferences and to think of efficient decision-making as its sole measure is to ignore all but the thinnest features of democracy” (Barber, 2003:198).

The concern here is that aggregation demotes the task of pre-voting discussion to that of informing pre-existing, fixed preferences rather than transforming them. By doing so, Barber suggests, aggregative theorists ill-equip themselves to explain the importance of citizens’ interaction before voting. The “stronger deliberative” view responds to this concern in detail. It stipulates conditions for the just formation of preferences amongst individuals. These conditions themselves take on great normative importance in a Habermasian democracy, for example. They are held to produce full consensus in ideal conditions. “Weaker” deliberative views (Gutmann & Thompson, 1996; 2004) also stipulate equivalent conditions; but with fewer claims about the subsequent agreement, and norms, they will create.

In contrast, some aggregative theories are believed to suggest that the aggregation of pre-deliberative preferences is sufficient to produce appropriate democratic decisions: no particular discursive conditions are necessary for them to be of value. At the root of this criticism is the fact that many aggregative theories suggest that the public interest is reached through aggregation of self-interested votes rather than enunciation of the public interest through deliberation. Further, some aggregationists like Riker (1982) and Downs (1957) incorporate from rational choice theory the assumption that individuals will act to further their self-interest rather than wider interests in the community when participating in pre-voting discussion.

Thus, aggregative democrats are often read as failing to address the importance of discussion. Counting (votes) is critical to good decisions; not talking. An emphasis on self-interested voting seems to leave little space for citizens to transform their preferences, and learn about the public interest, through discussion, in the way that deliberative democrats demand. Many social choice accounts do not assume the self-interest of citizens at the ballot box. Nonetheless, they are wedded to the idea that preferences are fixed through the democratic process, and that discourse must be functional and informational, rather than moralised, as a result.

Consequently, neither self-interest nor non-self-interest aggregationists can claim, as Elster does, that deliberation in public causes arguments to move from bargaining using interests towards the exchange of reasons:

“ the effect of an audience is to replace the language of interest by the language of reason and to replace impartial motives by passionate ones” (Elster, 1998: 111)

This seems like a fundamental difference. The absence of “public reasons”, and persuasion using them, makes it difficult for aggregationists to explain the value of discussion. Indeed, aggregative theories in general do not want citizens to move beyond interests, but instead to *reflect* them when voting. This idea is prominent in Habermas’ own characterisation of the interest-based, what he calls “liberal view”, theories that he seeks to move beyond. In “Beyond Facts and Norms” he captures this concept by explaining that the “liberal view” treats individuals as independent variables in the democratic equation, and conceives the “democratic process...exclusively in the form of compromises among interests” (Habermas, 1997: 296). Under ideal Habermasian deliberation, “participants question and transcend whatever their initial preferences may have been” (Habermas, 1997: 449). Exchanging arguments with others, under the right conditions, can thus move a citizen “beyond” her initial preferences. Without these

ideal conditions of deliberation, it is alleged that aggregative democracy fails to provide a plausible account of pre-voting discussion. It lacks an account of appropriate *persuasion* using *reasons* rather than merely information-sharing to inform existing preferences. Of course, the idea of persuasion using reasons relates to other deliberative assumptions like the capacity of pre-voting discussion to transform (not just inform) preferences before citizens vote; and to create consensus between participants.

Under a self-interested view of aggregative democracy, an “aggregative citizen” will reason, discussing her vote with others to seek out the information necessary to understand how her preferences translate into voting intentions. She will ask how proposed policies are likely to affect her job, community, healthcare, and so on. She will not, however, revise her fundamental preferences—whether they are assumed to be self-interested, public-interested, or motivated by something else—in light of the reasons presented to her by fellow citizens. This is precluded by a foundational assumption of much aggregative theory: the independence of voters.

To understand that assumption, consider a thinker who is often thought to be the father of aggregative democracy: Condorcet (1785). He modeled outcomes for a group of jurors facing a binary choice between two simple alternatives. Assuming that there is a single “correct” answer, he argued that the majority of members of a jury is more likely to be correct than any individual juror. This follows, assuming that each has an equal and independent—better than random but worse than perfect—chance of being correct. He reached this conclusion by multiplying out the probability that each individual juror has of reaching a correct decision. Doing so, however, requires that the chance of any individual juror being correct is independent of the chance of any others being correct (Estlund, 1994:131). This assumption explicitly precludes one juror, or voter, spreading a correct opinion to others through persuasion or deliberation. If we assume that each voter, or juror, has a better than half chance of being correct, it follows that increasing their number also makes a correct collective decision more likely:

“... 250 voters at competence of 0.51 have a group competence of 0.62, while a group of 10,000 at the same competence have a group competence of 0.98” (Estlund, 1994: 231)

For Condorcet, and his followers, voting thus increases the chance of a “correct” decision through multiplication rather than deliberation. The assumption of voters’ independence rules out preference change through persuasion within a group of voters. Aggregative democrats rely on weight of numbers rather than persuasion to produce good decisions. To that end, an aggregative citizen listens to others to unearth information that informs views on how her preference(s) will be satisfied (Fearon, 1998: 45-47).

Aggregative theories with a self-interest orientation suggest that, in gathering information on all of these matters, however, she will remain focused on her own interests; and select among available options to serve them as best she can. After collecting information, her vote should reflect her best judgment of what should serve her own interests, given all of the information gathered. In other words, the task of pre-voting discussion is to inform existing (self-interested) preferences, not to transform them. It is the way that our votes combine at the ballot box that increases the likelihood of good decisions rather than modes of talking before that. Following Condorcet, the individual voter may try to improve his competence from 0.51 to 0.52; but not to persuade his peers to change their minds.

Aggregative voters should thus focus on interests rather than deliberative reasons. The interests they concentrate on may either their own, or wider than this, depending on assumptions made by a particular aggregative theory. Thus, it is usually assumed that a type of instrumental reasoning, associated with aggregative approaches, is incompatible with preference change because it involves the collection of information rather than the transformation of preferences.⁸ For the deliberative theorist, this is—at best—an impoverished view of pre-voting discussion. For Barber and others, pre-voting discussion should do more than merely better inform the self-interested preferences of individuals; a theory that fails to take account of this neglects part of the value of the democratic process. This accusation is often associated with a sense that mere aggregation allows “raw” or unrefined preferences to feed directly into democratic decisions (Fishkin, 2009). This is a common accusation, to which we shall return in Chapter 2.

We may understand “raw preferences” simply as pre-deliberative: those which have not been transformed as a result of a process of deliberation. Deliberative theories thus have a ready response to the challenge of presenting a plausible account of what constitutes effective discussion between citizens before voting. What makes discussion vital is that it allows citizens to move beyond their raw preferences; to learn about the public interest from one another, and to acknowledge shared reasons about what is in the interests of the community. In doing so, they will sometimes be expected to reach consensus.

Aggregative theorists, however, do not usually stipulate conditions for the formation of valuable preferences. Instead, following Arrow’s condition, the expressed or “raw” preferences of individuals are held to determine social choices. The aggregative democrat cannot step outside those preferences to assess how “valuable” or “correct” they are. Rather, preferences are taken to represent individuals’ interests, and these, aggregated via voting, serve the public interest. On this view, the focus of democratic theory must

⁸ Of course, if we assume that “preference change” must be defined as transformation along deliberative lines, by reference to an “epistemic moralised” standard, this will be true by definition.

be on translating individual preferences into social decisions; not assessing preferences by reference to an exogenous account of what collective decisions should achieve. Consequently, this suggests a form of prudential reasoning applies either in reference to self-interest (on “rational choice” aggregative views) or wider interests (on some “social choice” aggregative views). An aggregative democratic theory thus emphasises counting (votes) and subsequently risks having too little to say about the value of talking. Without a moralised conception to call upon, aggregationists are unable to conceive of the value of talking in terms of preference change, or the presentation of reasons, and persuasion using arguments, that are justified to one’s peers.

Relatedly, in the absence of a deliberative account of the conditions under which preferences should be formed and revised, aggregative theories allegedly have too little to say about the context in which preferences should be formed (Christiano, 2002:32). Without an exogenous, public interest or procedural, measure of the “value” of preferences, aggregationists struggle to repel this critique, other than to simply assert the importance of individual preferences. Reaching the public interest through aggregation, rather than enlightening deliberation risks leaving us without a plausible account of why pre-voting discussion should matter, or take place at all.

How might aggregative theories respond? “Rational choice” aggregative theories demand that individuals vote merely according to their own interests, which are taken to be reflected by expressed preferences. For these views, the range of potential responses is limited. Using Arrow’s fourth condition (citizen sovereignty) implies that they cannot evaluate preferences by an imposed external standard. Clearly, the public interest would also be an inappropriate measure by which to assess individuals’ preferences if our democratic theory suggests that individuals should vote on the basis of self-interest. Some “social choice” aggregative theories are, of course, not committed to the assumption of self-interest. They could posit that voters gain a better understanding of the public interest through discussion. Nonetheless, they reject the idea of preference change so may explain the value of this discussion only in the functional terms of information sharing.

Consequently, aggregative views of pre-voting discussion have been thought impoverished. Without public-interest (or discursive) criteria for the assessment of preferences, the aggregative theorist seems obliged to accept the dictates of whatever expressed preferences individuals happen to form, no matter how unpalatable they appear. It is in this context that deliberative democrats often make passing reference to Bentham. Theorists like Gutmann & Thompson thus assert that:

“In the classical utilitarianism of Bentham, any end that anyone pursues counts as a kind of utility” (1996:169)

Without a plausible account of discussion, aggregative theories are bound to blindly aggregate and follow preferences: they cannot account for how discussion should improve them.

So, while deliberative democrats can explain that pre-voting discussion is vital because it enables the improvement of preferences towards consensus on the public interest, aggregative democrats can offer no equivalent response since:

“...they retain the “fundamental premise that there exists no procedure-independent fact of the matter as to what the best or right social outcome is [outside individuals’ preferences]” (List & Goodin, 2001:5)

The implications of this premise are extensive, and have been expounded in numerous ways. An example of an aggregative theory and social reform is illustrative. Despite his progressive work on women’s rights, J.S. Mill’s (1970) democratic theory has received criticism from feminists who accuse him of a reluctance to criticise “raw” preferences (Annas, 1977). Without an independent standard by which to evaluate expressed preferences; or stipulated conditions for their just formation, Annas claims that Mill is committed merely to satisfying the preferences that individuals happen to hold at present, rather than improving them.

This has pronounced effects for Mill’s writing on the emancipation of women. His commitment to a form of “utilitarianism” thus puts unjust, patriarchal preferences on a par with others because it appeals merely:

“to the satisfaction of desires that people actually have, not those they would have in some ideal condition” (Annas, 1977:181)

Annas accepts that there are other conflicting aspects to Mill’s thought. She insists, however, that the aggregative—or “utilitarian” as she terms it—impulse in his writing relies on this premise. This implies that a Millian utilitarian is bound to accept the expressed preferences of subjugated women to continue to be denied equal property rights as a reason to deny them this freedom. Without some moralised

conception of discussion, aggregative theorists are unable to explain why the “legitimate” desires of 18th Century women to hold property should be satisfied; but their “illegitimate” desires not to have this right should not be.

Note the obvious answer that deliberative democrats have to this challenge. The aggregationist is supposedly committed to counting the “illegitimate” preferences on a par with “legitimate” preferences because he rejects moralised deliberation. In contrast, deliberative theorists may deem subjugated women’s preferences for continuing subjugation illegitimate: either on moralised grounds that they do not represent relevant interests, or because they were not formed under the right discursive conditions. Indeed, in historic conditions where women’s education was skewed towards reinforcing patriarchal norms, Annas herself suggests that existing preferences are illegitimate for this reason:

“...women would collectively complain about the position of men if their education were not aimed at getting them to think of themselves as dependants with subservience to men as their natural goal.” (Annas, 1977:182)

Arguments presented by Sunstein illuminate other aspects of this challenge (Sunstein, 1993:245). He raises the possibility that individuals’ preferences and beliefs may “adapt” to an unjust status quo; particularly in reference to environmental policy. If citizens’ preferences do adapt to unjust situations, then it is difficult to understand why expressed preferences deserve the respect afforded to them by aggregationists. The aggregative democratic theory of Bentham is thought, by contemporary critics like Richardson, to suffer from the same problem. It appears to acknowledge the preferences of all people as holding equal value, even in apparently unjust situations (Richardson, 2003:43). Without a response to the challenge of explaining the value of discussion, aggregative theories seem powerless to control the implications of individuals’ expressed preferences. Aggregative democracy requires such a response in order to challenge the contention that it must satisfy *whatever* preferences individuals have.

It is not easy, however, to see how aggregative theories could provide such a response, without accepting some quasi-deliberative claims. These include: public orientation; transformative power; moralised status; or a consensus-generating view of pre-voting discussion. Unlike Fishkin (2009), they cannot argue that deliberation about the public interest, in concert with other citizens, causes us to move away from a self-interested preference for a new opera house, to a more public-spirited preference for a new football stadium. Unlike Habermas (1997), aggregationists cannot say that deliberation under the right (ideal)

conditions prompts us to better understand the public interest. Rather, aggregative theory relies on weight of numbers at the ballot box to achieve that goal, not persuasive (deliberative) reasons.

The aggregationist cannot assume, as Gutmann and Thompson (1996) do, that pre-voting discussion moves individuals beyond mere self-interest; helping them to understand the public interest through reason-exchange among equals. Another alternative is the “intrinsic” view that deliberation is, in itself, valuable. Deliberation could be seen as an acknowledgement of what reasonable persons owe one another; of a “value of fair cooperation among citizens on an equal footing of mutual respect” (Cohen, 1994:1505). This kind of claim may form part of the “weaker deliberative” argument,⁹ but this too is unavailable to aggregationists. The argument that deliberation embodies a duty of reciprocity between citizens (Gutmann & Thompson, 1996:52) is clearly exogenous to citizens’ preferences. It cannot be imported into an aggregative thesis without implying an epistemic moralised conception of deliberation; thus transgressing citizen sovereignty. More fundamentally, the very idea of persuasion using shared reasons while deliberating, is inimical to the aggregative assumption that voters are *independent* of one another before they vote.

Given these constraints, an aggregative theory must provide the following in order to answer the challenge of effective discussion:

- a. some epistemic prudential, not epistemic moralised, criteria for the assessment of preferences which can explain how one preference held by an individual may be “superior” to another preference held
- b. an explanation of how discussion amongst citizens before voting can “improve” expressed preferences held *without* relying on deliberative assumptions.

To recap; this challenge reflect the idea that discussion takes a back seat, or perhaps even becomes unnecessary altogether, once the main normative work within a democratic theory is being done by the counting of votes rather than discussion between citizens. Answering this concern could assuage some of the fears held by deliberative democrats about the aggregative tradition. The question remains: can an account which denies the importance of public spirited “deliberation”, and persuasion using shared reasons, nonetheless explain why discussion is important? What role remains for discussion in a theory

⁹ The different dimensions of public reasoning accounts examined sometimes pull in different directions. I have considered deliberative theorists who reject or amend full consensus, but endorse the intrinsic value of participation as “weaker” deliberationists. In doing so, I accept that could be considered otherwise in light of their views on other dimensions.

which supposedly endorses “mere aggregation, through majority vote, of existing, unreflective, presumably selfish preferences” (Posner, 2003:131)?

Preventing manipulation

We have examined the principal accusation levelled against aggregative theory by its deliberative critics. Now let us turn to a principal criticism of deliberative theory, put forward by aggregationists. A second challenge for accounts of public reasoning is to explain how discussion before voting can be protected against domination by elites, particularly through their manipulation of the outcomes it produces. This is related to a particular concern about the deliberative project.

Deliberative theorists like Gutmann & Thompson have specifically implied that citizens should reach different substantive conclusions if they deliberate better (2004:12). Thus, Posner (2004) suggests that providing an account of ideal deliberative conditions implies the expectation that others may agree with my own substantive views; if only they deliberate properly. There is a concerning potential for paternalism, and elitism is some strands of deliberative democracy. Providing more substantive explanations of the “improvement” in preferences explains the value of pre-voting discussion. It also, however, potentially establishes the expectation that ordinary voters can be thought to have simply failed in their deliberative duties, and reached inadequate conclusions as a result. The aggregative theorist’s suspicion is that perhaps “...what motivates many deliberative democrats is... a desire to change specific political outcomes, which they believe they could do through argument, if only anyone could be persuaded to listen, because they are masters of argumentation” (Posner, 2004).

This concern is related to the idea that deliberative norms might become manipulative if they are coupled with a belief that citizens have reached the “wrong” conclusions. A number of interpretations of “manipulation” have been put forward. These generally hold that a deliberating citizen is manipulated when placed in "a situation in which elites induce citizens to hold opinions that they would not hold if aware of the best available information and analysis" (Zaller in Shapiro, 1998:514). Similarly, Klemp defines manipulation in terms of two necessary and sufficient conditions:

Agent B is said to have been manipulated where “agent A uses hidden or irrational force” to affect B’s choices and “agent A acts intentionally” (Klemp, 2011: 60)

This definition has several elements. Klemp’s view of manipulation involves two agents, and requires that one agent influences an outcome through the use of force over another. The manipulated party is thus somehow *made* to play his part in that outcome because his choices are *affected* rather than freely chosen, and independently determined. It is because B’s choices are affected by A; rather than consciously and

independently determined in an informed way, that the means of manipulation is hidden. It does not engage B's capacity to make an informed choice. The force exerted may alternatively be irrational, because it does not reflect the appropriate consideration of relevant reasons that B would ideally give to the question at hand. Political scientists have identified and studied a number of different ways in which deliberators may be manipulated (Druckman & Nelson, 2003). These include the use of framing effects to make a particular outcomes more likely. While considering whether an extremist political group should be allowed to protest, for example, describing the question in terms of free speech, or in terms of public safety may make opposite results more likely (Druckman & Nelson, 2003:731). Critics have identified this as a danger of deliberation, since "elites shape citizens' views on matters of public concern by framing them in persuasive ways" (Stokes, 1998: 124-6).

It is important to be specific about the kind of manipulation we are concerned about here. There is an extensive literature on agenda manipulation: how the order of choosing between options affects how preferences translate into decisions.¹⁰ This is not, however, our focus. Agenda manipulation is a danger for all democratic accounts which require voting, so tells us little about different conceptions of pre-voting discussion. Rather, our focus is on the process of participation before voting; how preferences might be manipulated during this; and how well accounts of public reasoning conceptualise, and combat, that danger.

There are two broad ways of explaining why manipulated outcomes are undesirable. First, we may object that manipulation does not respect the autonomy and dignity of the individual being manipulated. Perhaps there is something intrinsically wrong about not allowing individuals to make choices in an informed way which allows them the opportunity to make up their own mind rather than, perhaps inadvertently, mirroring the views of others. Effectively controlling the decisions made by another person does not respect their autonomy to make that decision for themselves. It could also be argued that manipulation entails failing to treat the subject of manipulation with the respect they deserve (Sunstein, 2015:7). These objections suppose a particular type of ethical view: they object to the character of manipulation, and the fact that it fails to afford individuals what they have a right to, rather than its consequences. The deontological objection to manipulation is that it fails to observe the right of deliberators to take a decision themselves.

¹⁰ See, for example, (Dietrich & List, 2007).

Alternatively, manipulation could be wrong because it leads to worse outcomes than available alternatives: it fails to serve the good of those concerned. This “welfarist” objection plausibly supposes that the outcomes of manipulated agreement or action will not serve the interests of individuals involved as well as their own informed and independent decision-making could. This relies on something approaching the “best judge principle.”¹¹ It rests on the assumption that, under the right conditions, individuals themselves are better at assessing what will serve their interests, than potential manipulators can be. Under the right conditions, therefore, her choices further her welfare more effectively than theirs would do.

The welfarist objection might not condemn every instance of manipulation in every circumstance. It is conceivable that some sufficiently wise and informed manipulator might be justified in some unusual situation in producing a particular outcome.¹² It is likely, however, to condemn manipulation of most kinds in most contexts. The welfarist view is evident in many contemporary instances of regulation where practices which resemble manipulation are permitted in light of the benefits they are believed to confer. For instance, behavioural economists like Laibson and Choi (Choi et al, 2003) have shown that consumers tend to follow default pathways of action when presented with complex choices. Here, the default is defined as a course of action to which the individual will remain committed unless she incurs some cost by actively choosing to opt out (Ibid:2). The costs involved may simply be the minor decision costs of opt-out. In part, this is explained by a tendency to procrastinate when faced with long term decisions. In cases like this, the choices made by consumers, even when they are given the time and information needed to assess their options, often do not reflect their best interests: “people may decline to change from the status quo even if the costs of change are low and the benefits substantial” (Sunstein, 2011: 1351). Behavioural insights like these have been summarised by the Thaler and Sunstein in “Nudge: Improving decisions about health, wealth, and happiness” (2009)

Research of this sort has led to regulatory changes in a number of countries to re-orientate defaults for consumers to provide better outcomes. In the UK, for example, legislation passed in 2008 and implemented from 2012 was used to switch the default option for UK workers from not saving for retirement to saving into a plan provided by their employer.¹³ Whilst these workers are offered the option to opt out, very few choose to do so (HMG, 2013). Debatably, the right to opt out could render this an

¹¹ For an explanation of the implications of this principle, see (Goodin, 1995:127)

¹² As we will see in subsequent chapters, Bentham interestingly presented arguments of this kind himself in his early career. Whilst he never departed from the view that the rightful end of political decisions was the public utility, the early Bentham sought that end by offering evidence, advice and information to a benevolent dictator, who could effectively manipulate his subjects through indirect legislation; the ability to shape their incentives and perceptions. As Schofield (2006) and others have illustrated, it was only later in his career, from around 1803, that Bentham conclusively became a democrat and argued that voting by citizens rather than simply decision making for them was essential to public utility.

¹³ For a brief overview, see The Behavioural Insights Team, 2014.

active choice for some consumers. In reality, the effect of consumer inertia rather than active, deliberative decision making has been that over 90% of those automatically enrolled since October 2012 have continued to save rather than opting out of doing so. That policy received widespread support on essentially welfarist grounds: there was strong evidence that regulating to shift this default would benefit consumers in the long-term by increasing participation in retirement saving.¹⁴ Exceptions like this, where apparent manipulation is accepted on clearly welfarist grounds are, however, rare. Both the welfarist and intrinsic perspectives provide strong reasons to object to voters being manipulated during pre-voting discussion. It is important that an account of public reasoning responds to this danger, and mitigates the opportunities for manipulation to take place during pre-voting discussion.

Responses to the danger of manipulation often follow the four dimensions outlined above: orientation; transformative power; status and expectation of consensus. Deliberative democrats usually claim that manipulative influence is precluded by the conditions which define properly-conducted deliberation. James Fishkin has extended that argument by arguing that we cannot even understand what manipulation is, much less counter its threat, without some conception of deliberation. In “When the People Speak (2009), Fishkin provides a wide-ranging, and compelling account the benefits of deliberation between citizens before they vote on questions which are important to their community. The book contains detailed analysis of how deliberation in carefully designed public forums can lead to better decision-making.

For Fishkin, part of what makes decisions from these forums better than the outcome of what he calls “raw public opinion” is that it produces a more developed and mature response to the facts, which cannot so easily be swayed or manipulated. In contrast, “[r]aw public opinion is vulnerable to manipulation because it is volatile, based on low information levels, susceptible to misinformation, strategically incomplete information, and priming. In the setting of a deliberative microcosm, a scientific sample of ordinary citizens should become thoughtfully empowered rather than manipulated.” (Fishkin, 2009: 125). This research suggests, therefore, that it is a central quality of well conducted deliberation that participants move beyond their initial views which can be easily changed by suggestion and the framing of information and towards a stable and reliable set of views which cannot be changed so easily to reflect the interests of others. Deliberation immunises against manipulation. To do so, however, it must fulfil some important conditions. The quality of deliberation, according to Fishkin, depends on five important characteristics which essentially distinguish deliberation from much ordinary conversation: information, substantive balance, diversity, conscientiousness, and equal consideration (2009:34).

¹⁴ For an overview, see: (DWP, 2010).

The first of these conditions requires that participants should be given access to “reasonably accurate information that they believe to be relevant” (Ibid). Substantive balance requires that arguments presented from one side or perspective should be answered by considerations offered by those who hold other perspectives. Diversity implies that discussions should be broadly representative: that “major positions in the public are represented by participants”. Participants will fulfil the need for conscientiousness to the extent that they sincerely weigh the merits of arguments presented. Finally, equal consideration ensures that the arguments presented by different participants are considered on their merits; irrespective of who has presented them.

Together, these conditions are meant to imply that a deliberating group reaches decisions in the right way. For the group, they provide the equivalent of an individual’s informed consent before a medical procedure. They assure that a significant level of understanding, reflection and intent was present for decisions on policies, for which “we all have to live with the results” (Ibid). High quality deliberation is thus likely to secure the benefits of better decision-making and reflect a more balanced and informed set of judgements in response to relevant facts (Fishkin in ed. Cheminant, 2011:34). Conscientious deliberators meet to listen to opposing arguments and assess the reasons used to justify them. They do not simply proceed with fixed preferences and reason instrumentally about the best ways available to satisfy them. This is often said to be a key dividing line between deliberative and non-deliberative democratic views. Phillips, for instance, claims that the “common core” of deliberative theories is a sense that political engagement can “change initial statements of preferences and interest”. The alternative, she suggests, is to conceive of “political activity primarily in terms of instrumental rationality” and “government as engaged in aggregation” (Phillips, 1995:149).

Conversely, the absence of these conditions diminishes the quality of deliberation and is likely to erode the quality of the judgements it produces. For instance, an imbalance in the information presented is likely to skew resulting decisions. Good quality deliberation thus, by definition, does not allow participants to be significantly misinformed or imbalanced in their collective decision making. It also has a further important characteristic, beyond this. Quality deliberation, as Fishkin describes it, has the ability to immunise participants against manipulation. It moves them forward in their own thinking on a particular question to a point where they are less liable to be unduly influenced by imbalanced or manipulative arguments which do not reflect the public interest. As a consequence, “Imbalance in the argument pool... has little effect on those who become more informed” (2009: 127) in this way.

Part of the aim of this deliberation is thus to protect participants against being misled by an imbalanced set of information in future. Imbalance in the pool of arguments presented may happen by accident. It

could also be the result of deliberate, perhaps self-interested, intervention. It is only where intentional intervention of this kind takes place that Klemp's (2011:60) definition of manipulation will be satisfied. Like Fishkin, Sunstein also describes manipulation in contrast with deliberation. For Sunstein, the "problem of manipulation arises when choosers justly complain that because of the actions of a manipulator, they have not, in a sense, had a fair chance to make a decision on their own" (Sunstein, 2015:6). There is therefore a distinct contrast between deliberative choices, made in an informed and reflective way, and manipulated ones which have been skewed by influences which do not appeal to participants' capacity for rational and informed reflection.¹⁵

Whether on intrinsic or welfarist grounds, the danger of manipulated outcomes has some likely implications for a view of how citizens should reason before they vote. Manipulated reasoning before voting will fail to respect the autonomy or dignity of voters on the former view, and will produce worse decisions on the latter. In this context, we have seen that Fishkin argues good quality deliberation both rules out manipulative influences by definition and immunises participants against manipulative influences when they take important decisions. He also goes on to make a stronger claim; that some form of deliberative democratic view is not only potent at countering manipulation but essential to understanding what it is.

Thus deliberation reflects the appropriately informed and considered views of all participants, reflected through the fairness of the conditions in which they engage with one another and the sincere, committed spirit in which they do so. In contrast, manipulation denies citizens the ability to reach decisions in that way. Deliberative democrats therefore provide a starting point for understanding manipulation. They begin to outline the conditions or circumstances whose absence defines, or at least facilitates, the manipulator's success. These deliberative conditions are vital because they are the background against which voters form preferences which are then counted at the ballot box. A deliberative account therefore provides a description of appropriate conditions for preference formation. It is in distinction to this notion of appropriate discourse before voting that a definition of, and antidote to, manipulation is found. That definition usually focuses on the information, consideration, and legitimate influence, that citizens should be subject to before they vote.

¹⁵ Unlike manipulation, coercion is transparent to all concerned. A coerced agent may face a threat associated with noncompliance, but he will be aware of this threat. To illustrate this difference, imagine a strong opponent of abortion considering how to vote in a forthcoming U.S. election. If I wanted that person to vote Democrat, I could prevent him finding out that many Democratic candidates are pro-choice. Concealing information in this way is, Klemp contends, a paradigm case of manipulation; it takes effect without the victim being aware: "When I get you to vote Democratic by depriving you of information, I take away your choice as to whether to comply with my wishes...I simply change your actions invisibly, without your knowledge, choice, or consent" (Klemp, 2011: 62).

An example might illustrate why manipulation is such a concern. Consider occasions in history when voters have received misleading information. For example, the Zinoviev Letter, was published four days before the UK's 1924 General Election.¹⁶ It is widely thought to have given British voters a misleading view of the degree of likely integration between UK Communist Party, the Labour Party, and the wider international communist movement. Estimates vary as to how significant its subsequent effects were on the electoral fortunes of the nation's first Labour government. There is, however, little doubt that some voters were given misleading information by the press—in 1924—when the letter was reported as genuine. Later claims by newspaper proprietor Lord Rothermere to have won a hundred seats in that election¹⁷ corroborate suspicions the story was deliberately manipulative rather than merely accidentally misleading. Fishkin's view implies that in order to define or understand what is wrong about instances of manipulation like these, one needs an account of the appropriate conditions for preference formation. Without a view on what the right conditions for voters to make their decision in 1924 were, we are unable to say what was wrong about the provision of misleading information, manipulation or imbalance in the pool of arguments expressed through the media. If we want to object to cases like these—Fishkin implies—an account of deliberation is required to understand what goes wrong when voters are manipulated. The acknowledgement that there is something wrong about individuals being misled or manipulated before they make important decisions thus inevitably leads to a deliberative view of decision making. For Fishkin, this is because:

“Once the issue of preference formation is highlighted, then there are grounds for demanding that people get good information, have access to arguments on competing sides, and have the chance to weigh the merits of those arguments—in short that they deliberate to some degree” (Fishkin, 2009: 37).

Conversely, this argument has implications for non-deliberative views. These accounts, since they do not attend to the importance of the conditions for preference formation, must instead focus on the ways in which preferences are aggregated. This would be problematic if it meant that a view of public reasoning was, for example, content to simply aggregate the preferences expressed at the end of a set of misleading discussions based on incorrect information and manipulation. Elsewhere, Fishkin equates these alternative approaches with what he calls “competitive democracy”, a view of voting which is more concerned with the process of counting votes and competition between governing elites, than reflecting the will of “we the people” through decisions. (Fishkin, 2009:85).

¹⁶ See Andrew, (1977).

¹⁷ See Taylor (1972:223-41)

Fishkin's research suggests that some view of the appropriate conditions of pre-voting discussion is required to understand the threat of manipulated outcomes. I do not believe, however, that anything in Fishkin's argument dictates that an account of this kind must be "stronger" in the Habermasian deliberative sense or even "deliberative" at all. The research does demand, however, that an account of public reasoning must have some account of appropriate conditions of pre-voting discussion. In some way, it must answer the challenge of explaining what constitutes "effective" discussion. It does not, however, dictate that any particular conception of the orientation; transformative power; status; or consensus-generating ability of that discussion is essential. Below, I explore how different contemporary accounts explain and seek to negate the threat of manipulated outcomes.

First, in respect to orientation, we have seen that many deliberative theories demand a public orientation from citizens when they engage with one another. Asking citizens to move beyond their own interests and consider those of the whole community, however, can itself present opportunities for manipulation. In particular, there is a danger that what can appear to be in the interest of "the public" often turns out to be in the interest of its dominant members. The distinction between public and private interests can become blurred (O'Flynn, 2010:312). This blurring is problematic if we rely on pre-voting discussion to reflect the public interest. Clearly, conflating public and private interests could directly de-rail that enterprise. If this is the case, deliberators may be manipulated to reflect the interests of dominant persons and groups. This is, therefore, a particular danger for stronger deliberative views which most emphasise public-orientated deliberation. Whilst the boundaries between different interests may be blurred under any type of public reasoning, these conflation are likely to have less effect under accounts where citizens are not asked to deliberate on the public interest at all. Citizens' mistaking particular interests for the public interest is likely to be a less serious problem for public reasoning accounts which ask citizens to vote according to their self-interest, for two reasons. First, there may be fewer—potentially misleading—references to the public interest in discourse before voting, because these would not be directly relevant to the orientation of debate. Second, where voters do nonetheless conflate the public interest with particular interests, this should not itself bear quite so directly on how they vote because they are not—in any case—voting for the whole community.¹⁸

¹⁸ The potential for deliberation to fail in this way is widely acknowledged by contemporary theorists. For instance, the use of symbols in political discourse can be used to manipulate opinion. It is understood by many political scientists to evoke "rather unthinking, reflexive, affective responses...rather than...calculations of probable costs and benefits" (Sears, 2001:17). Symbols might be used in all sorts of ways in political discourse, whether they refer to public interest or self-interest. In studies of this sort, symbols are defined as focusing on "some enduring evaluative predisposition", created at an early stage in a person's psychological development (Sears, 2001:16). Defined in this way, their featuring in political discourse is likely to produce manipulated results.

This danger would be of relatively little concern if the effect of symbols in political discourse could be reversed or mitigated through deliberation. There have been many studies to demonstrate that well-conducted deliberation can have that effect¹⁸. There is, however, evidence

In contrast, many, particularly “stronger”, deliberative theories attempt to correct these potential sources of manipulation through a form of deliberation aimed towards the public interest. This poses an additional problem. Contemporary theorists of political language have found evidence that abstract language commonly used in debates like these can lead citizens away from a measured and rational assessment of costs and benefits. Edelman, for instance, concluded that abstraction—for example towards notions associated with the public interest—is frequently used to justify policies without accurate reference to their real costs and benefits:

“The revenue service deprives people of money, almost always involuntarily; the military draft imposes involuntary servitude... Usually the rationale for such restraints is an ambiguous abstraction: national security, the public welfare, law and order. We do not experience or name these ambiguous and abstract objectives as any different from goals that consist of concrete benefits, such as traffic control and disease control. Linguistic ambiguity spreads the potent rationale of these latter types of benefits to justify far more severe constraints and deprivations (including death in war) in policy areas in which benefits are non-demonstrable and doubtless often nonexistent.” (Edelman, 1974: 301).

Abstract language is therefore another potential means of manipulation. In this respect, theories which are directed at self-interest hold an advantage because they are less reliant on a form of discussion between citizens which can be manipulated; both by conflation of public and private interests, and by ambiguity of language. An account which makes less ambitious claims on behalf of pre-voting discussion is less likely to expose participants to potential manipulation through conflation of public and private claims, or through abstract language associated with generalisable interests, rather than particular ones.

This issue is thus particularly pertinent for those deliberative theories which see public reasoning in moralised terms. This gives pre-voting discussion a more significant role in democratic theory and helps to explain why discussion before voting might be important. It also, however, creates a new set of demands on citizens. Consider the effects of conceiving of deliberation in moralised terms as generating important normative values like legitimacy and justice. If, as seems likely, moralised deliberation of this sort also involves reference to abstract notions, there will be more opportunities for manipulation of those

that early acquired dispositions persist through an individual’s life, and continue to shape their judgments (Sears, 2011:20). This particular mode of manipulation could be a danger for all theories of public reasoning, whether aggregative or deliberative. “rational choice” aggregative theories might attempt to weed out this effect through a process of discussion directed towards voters’ self-interest.

terms in their use. Talisse (2009: 54-58) has linked these kinds of damaging discursive norms with the epistemic criteria stipulated by contemporary (particularly Rawlsian) public reasoning accounts. In these views, the moralised role of discussion creates particular duties, and expectations, on those participating. These, in turn, have the capacity to limit discussion and allow its outcomes to be manipulated. Notions of reasonable persons and acceptable reasons have the potential to be manipulated to diminish the input of some citizens. It is thus an advantage for accounts of public reasoning if they can explain the value of discussion without recourse to these.

It is partly for these reasons that some (“weaker”) deliberative democrats are less stringent in the public orientation which they believe properly conducted deliberation demands of citizens.¹⁹ Accounts of this sort are likely to be less susceptible to allowing the manipulation of outcomes in this way. Similarly, “social choice” aggregative views are likely to be susceptible to manipulation in this way only to the degree that they demand a public-interested perspective from discussants. Of course, this varies significantly between specific accounts. For similar reasons, “rational choice” aggregative theories are extremely unlikely to allow manipulation through citizens being encouraged to conflate the interests of dominant participants with those of the public at large. Since these views do not require a departure from the voter’s own self-interest, the question (or misunderstanding) of what constitutes the public interest is less likely to arise at all.

The same principle applies to other dimensions of accounts of public reasoning. A transformative conception of pre-voting discussion also carries with it some danger of manipulated outcomes. Of course, preference change has both positive and negative implications. It is possible that deliberation will cause individuals’ preferences to align on measures which serve the public interest, as intended, and improve the quality of decision making. It is also, however, possible that expressed preferences could worsen. They may change in line with errant, or mistaken, views of the public interest in a way which reflects manipulated discourse. Transformative theories are therefore subject to both the potential gains of well-conducted deliberation, and the potential problems of manipulated discussion. In contrast, non-transformative theories are immune to both. According to a transformative account, the potential for manipulation is a necessary risk of conducting the kind of deliberation. This is essential to producing public-interested decisions. Without it, citizens’ pre-deliberative views would be simply aggregated to produce decisions that do not serve the whole community. Imagine, however, that it is possible to reach the same result (decisions that serve the public interest) without recourse to a transformative conception of public deliberation. This would make the same, positive, result achievable without the risk of preferences

¹⁹ See, for example, Mansbridge (2006), Young (1990)

“worsening” because of transformations that have been manipulated by dominant voices. A non-transformative account, that can reach the public interest, for example through aggregation, would therefore be at an advantage over transformative counterparts in respect to preventing manipulated outcomes.

In theory, at least, a successful aggregative account offers many of the benefits attributed to successful deliberation, but without some of the potential pitfalls. It would provide fewer opportunities for manipulation, *if* a reliable aggregation of self-interested preferences could be shown to take place. The less ambitious role attributed to discussion by aggregative theorists has both advantages and disadvantages. My claim here is that successful aggregation of votes can do the normative job often attributed to deliberation. This means that there is less need to risk potential manipulation in search of completion of that job during pre-voting discussion. Of course, it is also possible that a pre-deliberative preference could be the product of some prior manipulation; or that no such aggregative mechanism exists for combining self-interested preferences. Those questions are, however, outside of the scope of our current focus: the normative merits of different accounts of public reasoning, rather than wider democratic theory. To consider public reasoning in its own right, we need to assume that citizens approach discussion with preferences which reflect their own interests. The relevant question for accounts of public reasoning is: “what should happen next?”

Aggregative views are less reliant on preference change. They are also less reliant on certain discursive norms. This is a potential advantage in respect to manipulation, because those norms may be used to manipulate citizens’ deliberative participation. The discursive ideal put forward by “stronger” deliberative democrats relies heavily on there being the right kind of cultural background against which discussion can productively take place. In non-ideal conditions, the potential for manipulated discursive outcomes could become a significant problem. Contemporary theorists often claim that justifying decisions to the public has an important “civilizing” effect on governors. It is important to remember that this too, like the other benefits associated with deliberation, is contingent on certain social norms against which discussion takes place. To understand this, it is useful to take a look at how the views of one very prominent democratic theorist, Jon Elster, have evolved over time. In 1998, Elster famously argued that the presence of a public audience to which officials feel the need to justify their arguments in terms of the public interest imposes constraints on the actions and arguments that those officials subsequently pursue. Self-interested officials may make “public interest” arguments to justify views which are actually based on mere self-interest or prejudice. By appealing to the public interest in this way, however, they then become subject to two constraints: “imperfection” and “consistency” (1998:104). The former means that the official will need to

ensure that his ongoing “impartial” arguments are not seen to too closely mirror his own interests²⁰ since this would undermine their credibility:

“If the impartial justification corresponds *perfectly* to the speaker’s interest, the disguise may be too transparent to work” (1998: 102)

Consequently, he will have an incentive to ensure that his subsequent “impartially” justified arguments are an imperfect reflection of his own interests. Similarly, justifying an initial argument on grounds of the public interest also constrains the positions that an official can take: self-interestedly, she will wish to appear consistent. An official who argues for a policy that aligns with his self-interest will be seen as opportunistic if he then deviates from it when it stops serving his needs (1998:104).

Together, these constraints “civilise” the claims made by officials during deliberation. Elster’s argument interestingly relies on the self-interest of the officials rather than their public-spiritedness. It provides an explanation of how a deliberative setting can encourage officials to make proposals that serve the public interest, even if their true motivation is self-interested. Unlike idealised accounts, his argument does not rely on the realisation of ideal discursive conditions. Like other deliberative theses, Elster’s argument is, however, reliant on the dominance of a discursive norm which favours justification in terms of public interest, and provides incentives for participants to appear impartial in its pursuit. This is more likely to be associated with epistemic moralised than epistemic prudential reasoning.²¹ The former more often espouses discussion in pursuit of the public interest, which implies the potential for citizens to persuade one another, using reasons which reflect that end. It is likely that many “rational choice” aggregative audiences will be more swayed by arguments that are relevant to their own interests, rather than “impartial” appeal to the public interest. “Rational choice” aggregative citizens reasoning together will not need to appeal to ideas like the “public interest” as much as deliberative citizens because their task is prudential. The aggregative citizen will listen to those claims that she believes bear directly on her self-interest, however. The exact details of the kind of public reasoning that an aggregative account encourages will determine how relevant the consistency and imperfection constraints are to aggregative officials.

²⁰ Note that the reference to his *perceived* interests does not appear in Elster’s account. This will be important in subsequent chapters when we contrast the role of hypocrisy under a Benthamite account.

²¹ This is not necessarily only because officials will never appeal to what is “good for society” under an epistemic prudential system, but because the culture of such a system is less likely to esteem public spiritedness over self-interest quite so much as an epistemic moralised system

More recently, however, Elster has also conceded that this “civilising force of hypocrisy” is itself contingent on public esteem for public-spiritedness over self-interested motivations, and that this may not apply in many cultures (2011), perhaps even including the United States. Without a public disdain for self-interested arguments, officials will not feel the full force of the imperfection and consistency constraints because being “exposed” as self-interested would not discredit them. This underscores how reliant Elster’s argument for the benefits of public justification is on an epistemic moralised culture which esteems impartiality and appeal to public interest arguments. Without that culture, the hypocrisy of officials cannot civilise their behaviour. This implies that deliberative accounts which assert the benefits of public justification as a response to the potential self-interest of elites must also show why their—sometimes implicit—assumption of public esteem for impartiality holds. We can also expect them to show how their argument works in conditions where this assumption does not apply. Conversely, an aggregative case for preventing manipulation of discursive outcomes does not rely on public respect for impartiality, so will be at an advantage over deliberative counterparts. In general, aggregative views are less reliant on these cultural factors.

This is particularly interesting in light of Elster’s admission that many cultures do not esteem public spirited motivation above “partial” claims.²² The importance of discursive norms is also illustrated by an interesting cultural example from Gambetta. In his essay “*Claro!*”, he notes the effects that a particular type of discursive culture can have on the efficacy of debate. Under such a culture, norms dictate that an individual’s failure to win an argument on one particular question discredits her, and causes her to be less respected when other questions are discussed amongst citizens. A “Claro!” culture takes hold when wider assumptions about knowledge are prevalent: expertise is seen as “indexical” (Gambetta, 1998: 24-5) or general rather than subject-specific; individuals are seen, in general, as learned or not, rather than expert in some areas and ignorant in others.

These norms shape the incentives that societies create for individuals. In particular, they provide a strong incentive for individuals to conceal their ignorance of a particular question. “Claristas” pretend that they already know what fellow citizens were going to say: that their point was clear, or obvious (“claro”). Individuals debating together change their approach to public discourse because they have much, in social standing and influence on future debates, to lose from “losing” any particular debate. This promotes discussion on the basis of reputation rather than reasons. “Winning” or “losing” debates will therefore come to reflect the forcefulness of those taking part rather than the merits of their arguments. “Claristas” do not learn together about the public interest; or other relevant reasons for a collective decision. Instead,

²² Elster’s thesis seems to discount the possibility that a deliberative audience could be so manipulated that they cannot discern between “public” claims and self-interested claims from officials. This is a danger that the Benthamite account takes more seriously.

they scrap for status; only taking on discursive battles that they can be confident of winning. Assumptions about knowledge thus change incentives. These, in turn, alter the nature of debate. A type of “discursive machismo” takes hold, and prevents the cooperative learning about the public interest on which deliberative democracy relies: “the distinction between arguments based on pride and arguments based on reason may be blurred” (1998:33).

One of the great advantages claimed for deliberation above aggregation is that it harnesses the benefits of persuasion using justified reasons, rather than merely aggregating relevant interests. Consequently, a “Claro!” culture has extensive implications. I believe that these illustrate some wider potential problems in deliberative theory. Claristas leave others’ questionable claims unchallenged when they are not confident that they can win a “fight” to disprove them. A “Claro!” culture unduly dismisses deliberators who have a sound understanding of one particular question (e.g. transport) if they have previously “lost” unrelated debates (foreign affairs). Consequently, the community will lose the benefits of their (transport) expertise in relevant areas. Debate will cease to revolve around justified reasons, and instead turn on the status of those involved. Gambetta’s argument emphasises the great impact that these kinds of norms can have on the efficacy of deliberation, and the potential for discussions to reflect the will of dominant voices rather than true agreement based on relevant reasons. This is a particular concern for theories which understand the capacity of public reasoning to reach agreement on the public interest, or something beyond mere self-interest, as essential to effective decision-making. Deliberative failings like these are critical if we rely—as some deliberative theorists do—on deliberation to reach the public interest.

Aggregative views of manipulation

In contrast, some aggregative approaches explain the interaction of citizens and elites without assuming the background of a society that esteems public-spiritedness. These aggregative views model the danger of elite control in a different way. Assume that citizens and officials are self-interested actors and wish their preference regarding a given policy decision to be realised. In this context, the discussion between citizens and elites can be modeled as the exchange of information that can inform parties' actions. In this situation, aggregative voters need to learn more about policy options in order to inform their "technical beliefs"; those about the causal relations between policies and outcomes. These dictate how citizens' preferences can be realised in practice (Austen-Smith, 1992:47). Pre-voting talk can also convey important information about the intentions of other participants; citizens will factor these into their beliefs. The exchange of such information will play an important role if information and expertise are unequally distributed across a society. For citizens, it may be the only way to garner essential information before deciding how to vote. Indeed, aggregative theorists have suggested the fact that individuals exchange this kind of information before voting is itself evidence that expertise and information is unequally distributed (Przeworski, 1998:145).

Prudential reasoners will thus be reliant on receiving information from elites before they vote, if those elites have access to information they need. Without this kind information, citizens will be unable to understand how their preference for lower mortgage rates or lower investment volatility should translate into a political preference for one policy or another. My preference for a lower mortgage rate for my home will thus translate into a policy preference only via a number of technical beliefs. These concern how domestic government spending, regulation, political priorities and many other factors influence the rate I pay. These beliefs will, therefore, be critical in determining how preferences produce votes on particular issues. Indeed, the same preference could lead to two opposite voting outcomes, depending on the technical beliefs of the individual who holds it. There is a problem with citizens relying on information from elites in this kind of way. If, as seems likely, interests—as well as information—are unequally distributed across society, then elites will have a strong incentive to misinform citizens. This is a significant problem: technical beliefs, informed by statements by elites, can cause an individual's preference to produce wildly different conclusions about specific policies. The unequal distribution of interests in society thus gives elites the opportunity to serve their own interests by misinforming voters. This is an important reason why rational choice theorists have been sceptical about the ability of talk to change preferences. Since talk is cheap, it is a convenient way to mislead. Citizens will thus be placed in a situation where they require information--through discussion—but are also aware that those providing information have an incentive to deceive them. Rational listeners will discount talk in many situations.

Pre-voting talk, then, *should* lack credibility unless it is in the speaker's interest to tell the truth, or it is not in her interest to lie (Austen-Smith, 1992:47).

While aggregative accounts are potentially less susceptible to manipulation of abstract terms used in deliberation, misinformation also poses a manipulative threat.²³ The scale of this threat will vary according to how strong discussants' incentives to lie are, relative to the cost of talking, thus "when lying is easiest, it is least effective because rational listeners take no notice of relevant assertions" (Austen-Smith, 1992:49). Listeners will discount statements in this way if they are acting rationally. We might ask what role misinformation plays when listeners are inhibited from doing so by pervasive norms. In this sense, aggregative accounts may also be reliant on discursive norms in a way which is analogous to deliberative counterparts.

Elite manipulation can thus be a threat via misinformation²⁴ and via unequal access to deliberation. It can, however, also take place through rhetoric, and the (mis)use of language. Habermas alludes to this problem when citing Becker's view that the "normatively laden but vague terms of political debate have...an emotional significance...not a cognitive function" (Habermas, 1997:294). This is something that rational choice approaches sometimes discount. Rational choice theorists (Austen-Smith, 1992) analyse speech in strategic terms, and explain its importance solely in terms of the information conveyed rather than the style or rhetoric used.²⁵ If deliberation evokes emotions rather than prompting citizens to learn, it fails in its primary purpose.

Rhetoric is usually accounted for in deliberative theories via the conditions stipulated for deliberation. Members of an elite who attempt to deceive citizens by acting like Gambetta's *Claristas* will fail to live up to an "epistemic moralised" Habermasian ideal of communicative action. In seeking to promote their own interests, they will not engage in the kind of cooperative mutual search for truth that is required. Similarly, "weaker deliberative" conditions for deliberation will also commonly seek to rule out the abuse of rhetoric

²³ The more that preferences among the discussing parties vary, the greater the incentives become for elites to deceive citizens, and the greater the problem of elite control is likely to be (Fearon, 1998:47). So, the aggregative view understands the threat of elite control in terms of the exchange of information primarily influencing technical beliefs rather than perverting deliberation.

²⁴ On the aggregative view, the truthful sharing of information through discussion aids preference satisfaction through voting. Once again, however, this exposes a strain between the challenges of PD and NEC. By providing an explanation of the importance of information sharing to aggregative democracy, aggregationists do show that pre-voting discussion plays a central role, so answer that aspect of the PD challenge. An aggregative response to manipulation thus will evade concerns about the propensity of deliberation to be controlled by elites. In doing so, however, it must also explain how aggregative citizens can counter the problem of elites providing misleading information as part of pre-voting discussion. We return to this point in chapters 4 and 5.

²⁵ A notable exception is Riker's "The Strategy of Rhetoric" (1996) which does seek to explain the value of rhetoric. It does so primarily by trying to explain the advantages associated with negative campaigning, however, rather than attempting to combat the problem of elite manipulation through rhetoric.

by stipulating discursive norms which preclude it. This will be considered an appeal to emotion or preconception, rather than a deliberative appeal to reason. Analogously, when the effect of rhetoric is to promote false beliefs that stymie individuals' preferences being satisfied, aggregative theories can clearly model this. The effects of rhetoric can be wider: influencing preferences, as well as technical beliefs. Stokes has argued that rhetoric during deliberation can influence fundamental beliefs about what is valuable, citizens' sense of identity, and other key factors in their voting decisions. We should consider the capacity of discussion to “influence citizens at a deeper level: to mold [sic] their very sense of who they are and what their capabilities are” (1998:124). It would be odd to care about the manipulation of voting intentions during discussion, but ignore potential to manipulate underlying preferences themselves. Aggregative theories could be accused of doing so by assuming that preferences are fixed through the democratic process.

An aggregative response to manipulation should thus address the concern that elite control of pre-voting discussion can influence citizens' political decisions and actions in a variety of ways, beyond misinformation. Elites may skew our causal beliefs, but this does not make our more fundamental values, or even sense of identity, immune from improper influence. By conceiving of discursive manipulation primarily as misinformation regarding causal beliefs, aggregative theories have often neglected this point. I believe that aggregative theories should address the important question of whether and how preferences can be manipulated before discussion. That question is not—itsself—however within the scope of this work. Our focus is instead on how citizens should reason together before voting rather than the complete process they should engage in prior to this.

Another “weaker deliberative” thesis is presented by Rawls. In acknowledging the “burdens of judgment”, he certainly does not expect ideally conducted deliberation to result in consensus as “stronger” Habermasians do. This helps explain why pre-voting discussion is of normative importance. Unfortunately, however, it also weakens the Rawlsian response to discursive manipulation. When criteria are established for encouraging citizens to preclude certain types of reasons from debate, so it becomes more likely that those criteria could be exploited by members of political and other elites to promote their favoured outcomes. Rawlsian public reasoning could be seen in this light.

There is therefore an aggregative suspicion that deliberation is associated with some form of paternalism about the conclusions that citizens “should” reach; if they deliberate properly. Even weaker deliberative views can be more exposed to discursive manipulation, therefore, because the norms they believe are served by well-conducted deliberation, although not productive of full consensus, have the potential to exclude legitimate input from citizens. Rawlsian “public reasons” thus risk introducing elite bias to

deliberation, as the history of conflation between perceptions of public and private spheres suggests.²⁶ The prescription of reasons of this sort makes it possible for some legitimate underplayed during, or excluded from, deliberation.

Advocates of non-deliberative accounts, like Posner (2004) have, therefore, implied that contemporary deliberative democracy could amount to a “power grab” by those with the intellectual and social resources needed to win public debates. Without endorsing such an accusation, we might still worry that the conditions required to participate in meaningful deliberation can exclude some in society and favour others. Since voting provides each individual with a clear way of influencing political outcomes, and relies less on discursive conditions, aggregative accounts are less susceptible to this concern. This speaks to a central strength of aggregative theory. Its seeming ambivalence about the substance of citizens’ preferences means that it is less likely to create norms which allow citizens to be told that their preferences are simply “wrong”, or have been reached in an inadequate way.

Thus, aggregative, and some “weaker deliberative” theories hold an advantage in this respect because they do not demand that citizens reach consensus through deliberation at all. At one extreme of this spectrum, a “rational choice” aggregative view could ask citizens to vote on solely according to their own interests; it would be no surprise if this caused individuals to disagree. Aggregation of self-interested preferences allows for more divergence because the public interest is reached through counting, not talking. Since this makes convergence unnecessary, aggregationists need not avail themselves of epistemic moralised discursive norms to promote convergence. As we have seen, aggregationists may also reject the “weaker deliberative” theorist’s more modest discursive norms which are designed to exclude “mere” self-interest from debate. In this respect, aggregative citizens are less at risk of a system that favours elite interests by pushing them towards consensus via discursive norms. Non-consensus-forming theories of public reasoning thus reduce potential for elite control through those norms. To achieve this, they sacrifice the ability to claim the gains that theorists like Elster attribute to justification in reference to the public interest produces.

²⁶ They also risk creating a variety of normative problems, and undermining the epistemic progress which discussions can make. See Talisse (2009: 42-77) so an overview

Conclusion

We have looked at a variety of contemporary accounts of public reasoning and examined how they vary across four key dimensions: orientation, transformative power, status, and expectation of consensus. We have also examined two normative challenges for any account of public reasoning: providing an account what constitutes effective discussion, and preventing elite manipulation of outcomes. Broadly, public reasoning accounts are better placed to explain what constitutes effective discussion when they have four characteristics. These views are most often held by “stronger” deliberative theorists who follow the Habermasian tradition:

- public orientation (information and learning required);
- preference transformation;
- moralised deliberation;
- expectation of full consensus.

For the reasons discussed, however, these four characteristics are also potential sources of manipulated discursive outcomes. Thus, while “stronger” deliberative theories provide a good answer to our first challenge, they provide a weak answer to the second. Conversely, a “rational choice” aggregative approach provides opposite strengths and weaknesses. It provides a poor account of the value of public reasoning before voting, but a good account of preventing manipulated discursive outcomes.

Between these two extremes, individual accounts vary in relation to each of the four dimensions noted. Some demand public orientation from deliberators; rely on the transformation of preferences; associate important democratic norms with moralised deliberation; and expect full agreement between participants to result. Nonetheless, it is now possible to revisit the trends identified above, in Table 1. Doing so should allow us to see how two normative challenges identified apply to a variety of contemporary views. Table 2, below, summarises these trends. In short, aggregative theories are often fundamentally opposed to the premises which allow deliberative theorists to paint a compelling picture of why pre-voting discussion is so important.

To summarise:

Table 2. Contemporary account of public reasoning, revisited

	<u>“Rational choice” aggregative</u>	<u>“Social choice” aggregative</u>	<u>“Weak deliberative”</u>	<u>“Strong deliberative”</u>
<u>(a)Focus</u> (What are citizens asked to vote according to?)	Self-interest	Either self-interest or public interest	Self-interest or public interest	Public interest
<u>(b)Transformative power</u> (Does discussion before voting transform or reflect preferences?)	Reflect	Reflect.	Reflect and/or transform.	Transform.
<u>(c)Status</u> (Does discussion play a moralised role?)	No.	No.	Often moralised	Moralised
<u>(d)Output</u> (Is consensus expected?)	No.	No	Only sometimes.	Yes
<u>Examples</u>	Riker, rational choice	List, social choice theory	Gutmann & Thompson, Fishkin.	Habermas, Dryzek.

I believe this presents a key, and difficult, challenge for aggregative theories of public reasoning. Once an aggregative account attempts to mimic deliberative theories by defining conditions for the improvement of individuals' preferences, it risks departing from the assumption of the "citizen sovereignty". Doing so also makes accounts subject to the accusations of paternalism and elitism expressed against some deliberative democrats. Preference transformation may provide a neat explanation of the importance of pre-voting discussion. The idea of transformative deliberation is, however, in fundamental tension with aggregative theorists' view that individuals' preferences should be aggregated rather than assessed against some prior—moralised—notion of the "public interest". Most intractably, aggregative theorists often assume the independence of voters, and this precludes altogether their influencing one another through deliberation. Fundamentally, aggregative theorists will instinctively resist calls to move away from preferences as a definition of value: to reject ideas of a "good [defined] independently of anyone ever thinking it to be good" (Goodin, 1995: 119). There is merit in this rejection. It means that aggregative views are less likely to allow citizens to be told that their preferences are simply "wrong", or have been formed in the wrong way. It also makes it more difficult, however, for aggregationists to explain why public reasoning should be an important component of democracy.

Despite these limitations, or perhaps because of them, aggregative theories are well-equipped to answer or second challenge. The less ambitious role they ascribe to pre-voting discussion offers fewer opportunities for manipulated discursive outcomes. Given this, the key question is: "can aggregative theories provide a better account of effective discussion *without* sacrificing their advantages in respect to preventing manipulation?". To begin answering that question, we will examine one interpretation of a prominent writer in the development of aggregative political theory: Jeremy Bentham.

Chapter 2. A conventional view of Bentham on public reasoning

"The common view of Bentham as a paternalistic utilitarian who wants to employ a corps of civil servants to measure utility and then govern to maximise it is clearly at odds with a free, indeterminate public opinion in a mass democracy"

(Cutler, 1999: 324)

Introduction

A relatively simple view of Bentham on public reasoning is often held; it concludes that there is little, if any, role for pre-voting discussion in democracy. While seldom expressed in explicit and positive terms, I believe that this simple or “conventional” view is evident in passing references made to Bentham by deliberative democrats, and in the relative inattention paid to the conditions of public reasoning in some Bentham scholarship. This view has clear implications. It means that Bentham’s democratic theory lacks the normative resources deployed by contemporary deliberative theorists to explain the value of pre-voting discussion. Contemporary “stronger” deliberative democrats follow Habermas (1997). They may cite four ideas to explain the value of pre-voting discussion: public orientation; preference transformation; moralised deliberation; and expectation of full consensus. In contrast, the simple or “conventional” view of Bentham cannot cite these concepts; doing so would contradict its fundamental premises. The conventional view is archetypally aggregative. It provides few normative resources to explain the value of pre-voting discussion. Conversely it also, therefore, provides few expectations of deliberation and—potentially—fewer opportunities for discursive manipulation, as a result.

In recent years, much comprehensive historical analysis has focused on how Bentham’s political views evolved. Particular attention has been given to his transition to democracy (Schofield, 2006); his plan for parliamentary and legal reform, and other related important aspects of his work. It is striking, however, that the story of how Bentham’s democratic views developed is seldom told in terms of how those views provide, or fail to, an account of public reasoning. Often, the focus is instead on the types of democratic institutions that Bentham wanted to build, and franchise reform required to deliver them. Below, I attempt to address this by looking at how Bentham’s relevant views are often understood. In later chapters, I focus on Bentham’s views on public reasoning, and how these can be understood to imply that a particular type of character, termed “public aptitude”, is required from citizens.

Before presenting that understanding of Bentham, we should first consider some alternatives. The “conventional view” is the first of these. This chapter thus attempts to do three things: to define this

“conventional” view of Bentham; to locate that view in relation to contemporary accounts by using the four dimensions identified in Chapter 1 (orientation, transformative power, status, expectation of agreement); and to examine whether this view meets the challenges of offering a plausible explanation of the value of discussion, and preventing manipulation of discursive outcomes.

What is the “conventional view” of Bentham?

What I shall refer to as a “conventional view” of Bentham on public reasoning is grounded in some well-known aspects of his ethics. At its simplest, this reading of Bentham suggests that the public interest is reflected in democratic decisions through the aggregation of self-interested votes alone. There is, therefore, little—if any—significant role for pre-voting discussion to play. Counting votes does the normative work; not (deliberative) talking between citizens. Pateman, for example, contrasts contemporary theorists who believe that democratic participation is valuable for a number of reasons with Bentham, for whom “The participation of the people has a very narrow function; it ensures good government... through the sanction of loss of office” (Pateman, 1970: 19-20). The implication here is that this sanction operates through the ballot box, alone; not discussion.

The conventional view is often contrasted with democratic theories which emphasise democratic participation from citizens, including through public reasoning, to a greater degree. Gutmann’s work (1980) is an interesting reference point here. Along with Thompson, she has been at the forefront of contemporary debates on deliberation; including how normative theory can be translated into better decision-making in the real world. These two authors have, through a series of complementary analyses, set out what they believe to be the significant benefits of increased deliberative participation in politics from citizens. The foundations for these more developed deliberative views (Gutmann & Thompson, 1996; 2004) were, however, set in earlier work. In “Liberal Equality” (1980), Gutmann outlines how a number of historical arguments have changed our understanding of democratic participation. Prominent among these examples is Bentham. Other philosophers conceptualise participation in different ways. But Bentham is used as an example of a theorist who is concerned with the outcomes of political decision-making and remains neutral between different ways—bureaucratic or democratic—of arriving at those outcomes (Gutmann, 1980:26). A crucial distinction is thus drawn between the variety of contemporary theories which place a high value on participation in a deliberative process, and Bentham who supposedly did not:

“In Bentham’s own theory, democratic participation has no value independent of the limits upon political tyranny that it is expected to effect. However, were participation seen as (in part) a

pleasure in itself, or as a means of increasing an individual's future capacity for pleasures, the realm of democracy might be increased...Mill's revision of classical utilitarian doctrine serves to remind us of this" (Gutmann, 1980:26)

This is the essence of the conventional view of Bentham. It suggests that "bureaucratic determination" (Ibid) is vital to good political decisions, and that voting is a necessary check on that exercise of power. Beyond these two concepts, there is apparently little space left for an active process of public reasoning in which citizens participate to promote the public interest. I believe that the tendency to see Bentham in these terms is a consequence of three, related but different, aspects of Bentham's thought:

- i) The belief that Bentham defines value as "raw" preference satisfaction.
- ii) Bentham's purely quantitative utilitarianism.
- iii) Bentham's aggregative definition of the public interest; what Postema (1986) refers to as his "individualism".

The first of these aspects of Bentham's thought suggests that there could be little role for pre-voting discussion in reforming or improving individual preferences. The second implies that there can be little significant role for public reasoning in "educating" individuals about the superiority of some (types of) preferences above others. Finally, the third aspect suggests that aggregating self-interested votes through the ballot box produces decisions which reflect the public interest. This means that public reasoning directed toward the public interest would, in any case, be redundant.

Below I examine each of these claims in turn and explore why they have provoked the conclusion that there is little, if any, role for discussion in Benthamite democracy. Finally, I outline how these three claims have interacted by examining some examples of the conventional reading of Bentham.

i) Bentham defines utility as "raw" preference satisfaction

For good reason, Bentham is best known as a utilitarian: he argued that we should evaluate "every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question" (Bentham, 1970a: 12). He stipulated that this principle had universal scope: it applied "not only of every action of a private individual, but of every measure of government" (1970i: 12). In the case of government, Bentham held, decisions should serve the public

interest, and be assessed by their tendency to do so. In contemporary thought, utilitarianism is usually taken to imply that utility is defined as preference satisfaction. So, on this basis, many writers appear to assume that Bentham believed that the right action is that which causes the most preference satisfaction.

This underlines the contemporary assumption, which we shall investigate below, that Benthamite utilitarianism calls for the satisfaction of whatever “raw” preferences happen to exist in a society. If utility is preference satisfaction, and Bentham assesses political decisions on utilitarian grounds, then it follows that individuals should vote according to their own preferences, and public decisions should reflect the aggregate of what these individual votes dictate. It is presumably on these grounds that Bentham argued that legislators should gather information about citizens’ existing (not ideal, fully rational) preferences before making laws. (1970:56). Bentham’s acceptance of the value of existing preferences provides a stark contrast with the views espoused by some deliberative democrats. For instance, it is difficult to see how the conventional reading of Bentham could account for the kind of “improvement” in preferences, envisaged by contemporary deliberative democrats like Benhabib (Benhabib, 1996: 73). Of course, different writers mean different things when they refer to "raw" preferences. Three particular conceptions of this idea are particularly relevant. Preferences might be “raw” in the sense that they incorporate some form of referential error. Second, a raw preference might may be personal, rather than social or communal. Finally, raw preferences may be considered those which were produced without adequate deliberation.

The referential aspect of raw preferences is exemplified by a critique of Bentham put forward by a respected contemporary political theorist, Kymlicka. This implies that while an individual may have a stated preference for “X over Y”, that preference will in fact be raw if she has misunderstood what X, or Y is; if her preference refers to the wrong things. Imagine I have two glasses of liquid in front of me and am asked to choose between them. The glasses are close enough to see but not to pick up or smell.

On the left is glass X, a tumbler filled with clear liquid which looks like water. On the right is a wine glass filled with a red liquid which looks like red wine. If asked to choose between these two options by an observer, I might respond “I prefer glass X over glass Y”. I may do so while thinking to myself “this is the right choice because I prefer water to red wine”. That stated preference will be accurate if the two glasses do in fact contain water and red wine. Imagine, however, that I then move to pick up glass X and, just as I take it my mouth, ready to drink notice the strong, and alarming smell of paraffin. In this case, my stated preference for one glass over the other was based on referential error. It did not accurately reflect a true state of affairs. In this sense, my preference could be considered “raw”.

This conception of raw preferences is implied by Kymlicka. In his *Contemporary Political Philosophy: An Introduction* (2002:31ff), he criticises a utilitarianism which equates value with the satisfaction of preferences. Interestingly, he cites Bentham as a key example to make this argument:

“Satisfying our preferences does not always contribute to our well-being. Suppose that we are ordering food for lunch, but some of us want pizza, while others want Chinese food. If the way to satisfy the most preferences is to order pizza, then this sort of utilitarianism tells us to order it. But what if, unbeknownst to us, the pizza we ordered is poisoned, or just rancid? Ordering it now would not promote our welfare. When we lack adequate information, or have made mistakes in calculating the costs and benefits of a particular action, then what is good for us can be different from the preferences we currently have. . . Utilitarianism of the preference-satisfaction variety says that something is made valuable by the fact that lots of people desire it. . . [but] having the preference does not make it valuable—on the contrary, its being valuable is a good reason for preferring it. And if it is not valuable, then satisfying my mistaken preference for it will not contribute to my well-being. My utility is increased, then, not by satisfying whatever preferences I have, but by satisfying those preferences which are not based on mistaken beliefs” (Kymlicka, 2002: 31ff)

Like the choice between glass X, which actually contains paraffin, and glass Y which holds red wine, this counterexample relies on the assumption that utilitarians like Bentham are bound to satisfy preferences which are “raw”: they refer inaccurately to relevant states of affairs. The example of poisoned pizza is straightforwardly based on individual in question lacking some critical information: that one of their culinary options is dangerous. If he had been better informed, the individual in question would have held a “better” preference; based on a fuller and more accurate description of the options open to him. Consequently, any preference utilitarian view which requires “adequately informed” rather than “raw” preferences would be unaffected by this critique. The fact that this example is used against Bentham illustrates that Kymlicka implicitly ascribes to Bentham a “raw preference satisfaction” definition of utility.²⁷ In Chapters 3 and 4, we shall explore whether Bentham really did believe that "utility is increased...by satisfying whatever preferences I have" (Kymlicka, 2002:31ff).

²⁷ Similarly, Kymlicka cites an individual’s preference for reading poetry which he finds ‘disturbing rather than pleasurable’ (2002: 13) as a counterexample, on the same grounds.

If understood in these referential terms, accepting raw preferences for aggregation entails that public reasoning cannot be used to correct even the most basic referential errors in citizens' expressed preferences. For instance, a theorist who accepts the value of (referentially) raw preferences would be content to aggregate citizens' expressed preferences for an increase in Value Added Tax (VAT) even if—in expressing that preference—they think that VAT is the same as income tax, or that it only applies to the wealthy. Rather than correcting this misunderstanding in what an expressed preference refers to, a raw preference theory would merely aggregate views and determine decisions. It would add together the—mistaken—preferences that citizens express without attempting to improve them through dialogue. Clearly, a view of this sort would leave very little scope for public reasoning for voting to play any meaningful role: it does not value its capacity to correct even basic errors. This claim entails not just that discussion is not needed to transform preferences towards the public interest, but also that it can play no role in reforming expressed preferences to better reflect self-interest. This stronger claim might sound like a “straw man”, or an extreme criticism. It can, however, be found in the work of some highly sophisticated contemporary thinkers. It is usually expressed in the distinctions drawn between the benefits of active deliberation and passive voting without deliberation (Gutmann & Thompson, 1996:170; Richardson, 2003:43; Chambers, 2004).

There is a second sense in which preferences may be “raw”. They may be personal rather than social; they may reflect the interests or concerns of an individual rather than those of the whole community, or a wider group. Relatively frequently through the literature on deliberative democracy, proponents of deliberation explain their position by distinguishing it from a view of democracy which merely adds together votes as expressions of personal interests rather than refining or developing them through the exchange of reasons and arguments between fellow citizens. Cohen, for example, argues that properly conducted democratic politics “involves deliberation focused on the common good” (1989:19), and that those “interests, aims and ideals that comprise the common good are those that survive deliberation” (1989:25). As O’Flynn (2010) has pointed out, there is certainly a potential tautology in this. Nonetheless, there is a sense in contemporary theory that deliberation is essential to voters’ holding views which reflect the public interest rather than their own narrow interests. In turn, those preferences which have not gone through refinement through deliberation are often considered “raw”. As a consequence of being pre-deliberative, they are considered unready to be counted at the ballot box, or accepted by all in a public forum.

Slightly different implications follow this second conception of “raw” preferences. A utilitarian who is content to aggregate this type of raw preference will see no need for discussion that transforms preferences toward the public interest. This clearly precludes a Habermasian form of deliberation; aimed at consensus on the public interest, achieved through preference change (Benhabib, 1996; Habermas,

1997). These theorists tend to define raw preferences in distinction to preferences which have been transformed through deliberation. This implies that a raw preference utilitarianism will entail a more self-interest based view of discussion and voting. If Bentham is a raw preference utilitarian in these terms, his work is more likely to sit in the “rational choice” aggregative category identified above.

Finally, raw preferences can be understood as those which have not resulted from well-conducted, appropriate deliberation. Like many deliberative theorists, Cohen (1998) combines the claim that raw preferences are those which do not reflect wider interests, and the view that they are also those which have not been improved through public deliberation. The latter implies the former. These two definitions are, however, distinct from one another. Consider, for example, a “weaker deliberative” view which does not demand public orientation from citizens. An account like this may acknowledge that a preference has been through an appropriate and full process of deliberation yet may still reflect personal interests rather than the public interest. This is a point that some deliberative theorists, like Mansbridge have been determined to emphasise (Mansbridge et al, 2010). Goodin, for instance, argues that preferences can be thought of as “raw” not because they are self-interested, but because they are in some sense inconsistent (1995:IV). On this view, "raw" preferences might be those which have not undergone this process of examination to become more consistent. This examination could take place through a variety of different means. These could range from an individual privately reflecting on her views, to full public deliberation, or a retrospective analysis of expressed preferences.

In other works, Goodin implies something which is explicitly closer to a self-interest-based definition of raw preferences. He has argued, for example, that the benefits of deliberation can be achieved through citizens engaging in a form of "imaginative empathy" by considering the perspectives of fellow citizens. This allows the interests of fellow citizens to be considered, even when they are absent from discussions, thus unable to deliberate themselves (Goodin, 2000: 84).

A conception of raw preferences is implicit in Gaus' 1999 book, “Social Philosophy”. In that work, Gaus alleges that the divergence between expressed preferences and interests undermines a preference utilitarian view: “An alcoholic may want to keep on drinking, but it is not obvious that this is good for him” (Gaus, 1999:52).

Both Kymlicka's and Gaus' objections assume that a basic form of utilitarianism—often associated with Bentham—defines value as the satisfaction of existing “raw” preferences. By highlighting examples where raw preferences seem to obviously diverge from “real interests”, they suggest that this form of

utilitarian democratic theory cannot plausibly characterise citizens' interests. The accusation is that a basic raw preference view deems damaging and harmful courses of action valuable; even though they cannot—intuitively—be so (eating poisoned pizza, or providing an alcoholic with alcohol). Deliberation is often proposed as a remedy to this problem. It provides individuals with the perspective, information, time and capacity to properly consider relevant interests (Fishkin, 2009:125).

As we have seen, the notion of raw preferences appears in three slightly different guises. Kymlicka's critique appeals to the idea that raw preferences can contain some form of referential error and still be maximised according to Bentham's view. Gaus' criticism is slightly different. Instead, it appeals to the reader's intuition that the alcoholic is making poor decisions. His stated preference for continuing to drink does not reflect his real interests because, under better circumstances, with full information and a full awareness of the consequences of his different options, the drinker would not choose to harm himself in this way. This stated preference is pre-deliberative, in the sense that it has not resulted from an appropriate process of consideration and information. It follows that if we define value in terms of the satisfaction of this kind of raw preference that there would be little conceivable role for pre-voting discussion in transforming raw preferences into post-deliberative preferences which are a suitable basis for collective decisions.

The exact implications of a raw preference view depend on exactly how we define that concept, however. The conventional view is right to attribute a self-interest-based view of public reasoning. Indeed, Bentham repeatedly emphasises that citizens voting should do so on grounds of self-interest. It is partly for this reason that Bentham endorsed the secret ballot; on grounds that it would lead to each vote becoming a better reflection of the individual's own interest (James, 1981: 55). Bentham's belief in the value of self-interested voting is famously illustrated by the differences between his views and those of JS Mill (Schofield, 2006: 144-50; Lever, 2007; Park, 1931). There is a considerable amount of evidence, which we will return to in subsequent chapters, to suggest that Bentham held what some contemporary deliberative democrats would classify as a "raw preference" view. Indeed, a self-interested understanding of voting underlines not only Bentham's own arguments for the secret ballot, but also his utilitarian logic for voting. Nonetheless, it would be a mistake to describe Bentham as a "raw preference utilitarian" in the other two respects considered above. We shall explore evidence for rejection of these views in subsequent sections.

A "raw preference" has obvious implications for public reasoning. It would suggest that Bentham's work can conceptualise no significant role for pre-voting discussion. This claim may be allied to either a deliberative conception or non-deliberative conception of "raw" preferences. The former, implied by

deliberative democrats like Habermas and Gutmann & Thompson, entails that no transformation away from self-interested preferences can take place on Bentham's view. The stronger, latter claim, entails that no role for pre-voting discussion could be conceptualised by a Benthamite approach, even on prudential rather than deliberative grounds. This conclusion follows from the assumption that raw preferences are an inadequate reflection of individuals' interests; and that Bentham was content to define utility in terms of their satisfaction.

In some cases, like those above, the attribution of a raw preferences view to Bentham does not seem to be based on a significant sample of his work. In other cases, the view is advanced as part of a far more sophisticated and detailed understanding of Bentham as a writer. Some scholars, for example, cite Bentham's readiness to define the satisfaction of subjective preferences as being rooted in his epistemology:

“the subjectivist Bentham is led in the direction of recognizing as many different truths as there are minds, with truth for each of us consisting in the set of propositions in which we believe” (Quinn, 2012:20).²⁸

A number of commentators have recognised just how extensive Bentham's commitment to subjectivism was throughout his writing career. Harrison notes, for example, Bentham seemed less concerned with the metaphysical questions or scepticism and realism than with pragmatic decisions about the kinds of enquiry that are likely to improve utilitarian outcomes (Harrison, 1983:53-4).

The subjectivist view of Bentham is buttressed by much of his political writing. The Bentham who wrote a constitution for the Islamic state of Tripoli, for example, wanted to accommodate and satisfy existing preferences rather than attempting to “correct” or “reform” them (Bentham, 1990:12). It is for this reason that his constitution for the territory is prefaced by a survey of sociological facts about existing practices and beliefs. These included the size and geographical location of mosques (1990:12). At its simplest, therefore, the raw preferences view of Bentham is based on a passing reference to his work. When allied to an understanding of his subjectivism, and approach to punishment and constitutional writing, however, it takes on a more plausible interpretative guise. The Bentham who recognised the value of utility that criminals derived from crime seems an unlikely proponent of any view which demands that citizens should move beyond their "raw" preferences before voting.

²⁸ Quinn also goes on to identify the 'objectivist Bentham' in the same article: this quotation is provided to illustrate that the subjectivist reading reinforces a 'raw preference' view.

b) Bentham is a quantitative, rather than qualitative utilitarian

The belief that Bentham defined utility as the satisfaction of raw preferences is complemented by the assertion that he was a quantitative rather than qualitative utilitarian. This implies that, since:

- (i)- the satisfaction of no (type of) preference is intrinsically more valuable than another; and
- (ii)- the value of a process designed to “reform” preferences cannot be explained in terms of its ability to provoke individuals' transition from holding one (type of) preference to another; therefore...²⁹
- (iii)- the value of pre-voting discussion cannot be explained in these terms.

The grounds for this understanding of Bentham are well known. Mill famously proposed a qualitative conception of utility, thus could conceivably deny (ii). In contrast, however, Bentham was famously adamant that the value of any activity or preference could be judged only by reference to the quantity of utility it produces (Crisp, 1997: 23). Quantitative utilitarianism does not rule out public reasoning creating significant utilitarian benefits. It does limit the way in which those benefits can be explained, however. It means that Bentham's form of utilitarianism cannot readily be associated with a public reasoning which induces qualitative improvements in the preferences held by citizens. Quantitative utilitarianism is not a natural complement to the kind of preference change envisaged by Habermasian deliberative democrats, for instance.

A quantitative notion of utility thus complements the “raw preference” view, and reinforces the conventional reading of Bentham. If utility varies only in quantity rather quality then one explanation for the value of pre-voting discussion—that it induces qualitative improvement in preferences—is ruled out. Since individuals' experience is the highest evidence of their own interests, and these preferences cannot be exchanged for qualitatively “better” alternatives, a Benthamite government seems bound to simply aggregate existing preferences via voting. “Conventional” Bentham is the simplest kind of hedonist. He defined the good as pleasure and did not discriminate between “good” and “bad” preferences. This is perhaps what Bentham is most famous for: the idea that pleasure is all that matters, and pleasure is merely the satisfaction of one's preferences, however apparently misled they might appear to be.

c) Bentham is an individualist (defines the public good in terms of individual welfare)

²⁹ Note that (ii) does not preclude the possibility that the value of pre-voting discussion could be explained in terms of its ability to provoke a move to preferences whose satisfaction create a greater quantity of utility, but this would require an extra-preference definition of utility.

The first two foundations for the conventional view mean that the value of public reasoning cannot be explained in terms of reforming, or inducing qualitative improvements in, “raw” preferences. Let us now turn to Bentham’s individualism. This assumption suggests that the public interest is no more than simply an aggregate of individuals’ self-interest. This entails that the value of pre-voting discussion cannot be explained, in Bentham’s view, in reference to some wider, non-individualistic, conception of the public interest.

Bentham defined the public good solely in terms of individuals’ welfare; and for this reason endorsed the secret ballot (1983b). This assumption means that Bentham could not, and did not, endorse a non-prudential (or ‘moralised’) conception of pre-voting discussion which emphasises the transformation of preferences so that they better reflect the public interest rather than self-interest. The textual basis for this view is well known. Most famously, at the start of an “Introduction to the Principles of Morals and Legislation”, Bentham explicitly posits individualism:

“The interest of the community then is, what is it?—the sum of the interests of the several members who compose it” (Bentham, 1970a: 2).

In reference to pre-voting discussion, we may reason from this statement that, since:

- i) Bentham is an 'individualist'.
- ii) 'individualist' accounts hold that public interest decisions can be reached through aggregation of self-interested preferences; and
- iii) public-interested preferences are not necessary to this aggregation; therefore
- iv) individualist accounts, including Bentham's, cannot explain the importance of pre-voting discussion by reference to the transformation of preferences towards the public good.

I believe that these three complementary claims have, together, provided the basis for a contemporary assumption that there is little significant role for pre-voting discussion in Bentham’s democratic theory. The implication of this view is that deliberators following Bentham’s democratic principles cannot promote the public good by better informing (self or public-interested) preferences. This is because he accepts “raw” preferences in his definition of utility. The conventional view cannot provoke a move from qualitatively “worse” to “better” (self or public-interested) preferences, because Bentham held that utility varied only in quantity rather than quality. Finally, and in any case, the conventional view has no need to

move individuals' preferences towards the public interest before voting. Bentham's individualism—like Condorcet's (1785) theorem—entails that the public interest can be reached through counting votes, without talking. These three premises thus provoke the conclusion that little of value can be gleaned from Bentham's writing about the importance of public reasoning.

Of course, we must take care not to caricature the relevant literature. For example—like many Bentham scholars—neither Lieberman (2014) nor Hume (1978) suppose that Benthamite democracy relies exclusively on aggregation; and that the citizenry should play no role other than voting. They would both also question the grounds for attributing a “raw preferences” view to Bentham; on grounds we shall investigate in the next chapter. There is, however, in these and many other works, including those of Bahmuller (1981), Manning (1968), and Himmelfarb (1969), and Mack (1962: 255-61; 450), the repeated implication that voting and institutions are the critical means of serving the public interest in Bentham's political theory, rather than a deliberative engagement from citizens. The emphasis in this kind of secondary literature is consistently on *counting* rather than *talking*. Perhaps as a result of this emphasis, other contemporary theorists (Richardson, 2003) have also accepted the conventional view of utilitarian democracy as focusing on counting votes, and institutions rather than public reasoning. Along with this presumption, these theorists also import the conclusion that Bentham's democratic theory lacks the normative resources needed to explain the kind of discussion that citizens should take part in when forming their preferences before voting.

In more general contemporary debates, Bentham is often mentioned as promoting an individualistic view of wellbeing. He is, according to these writers, towards the more extreme version of a political worldview which discounts the importance of society, and only has regard for individuals. A good example of this kind of cursory interpretation of Bentham is found in Keen's recent popular book “Debunking Economics”, which portrays Bentham as the father of neo-conservatism, based on this individualism. In making this claim, Keen relies on the quotation above to ascribe egoism to Bentham: “Behind apparent altruism, behind apparent selfless behaviour, behind religious commitment, lies self-interested individualism” (Keen, 2001:26). Of course, the reality is more complex. A variety of different conclusions have thus been drawn from Bentham's individualistic definition of the public interest (above). These views reinforce the conventional reading. Like consumers in a marketplace, citizens in the society Bentham envisaged would register their self-interested preferences on the ballot paper. They were not to be public-spirited deliberators, in search of the mutual understanding of common values.

Where does conventional Bentham sit in relation to contemporary accounts?

Having explained some of the central aspects of the conventional view of Bentham, it is important to understand how this view fits into contemporary debates on the value of discussion between citizens before voting. In Chapter 1, I focused on four dimensions in which contemporary accounts tend to differ from one another: orientation, transformative power, status, and expectation of agreement. Where does “conventional Bentham” sit in relation to these four questions?

First, consider the orientation of discussion. Conventional Bentham suggested that individuals could generally be expected to act according to their own interests, and that public-orientated preferences were not necessary to public-interested decisions because the public interest could be reached through an aggregation of individuals’ self-interested preferences. The conventional view of Bentham therefore suggests that public-orientated deliberation is unnecessary. In this respect, it follows “rational choice” aggregative theories which rely on an aggregation of self-interested preferences rather than public-orientated deliberation. Like Condorcet’s theorem, the conventional view of Bentham seemingly reaches towards the public interest through multiplication rather than deliberation. Individuals are the best available judge of their own interests. Indeed, this is the basis of Bentham’s argument in IPML that we may trust citizens to vote. By reflecting self-interest at the ballot box, they play their part in creating collective decisions which serve the community as a whole, since the rightful end of these decisions should be nothing more than the aggregate interests of all individuals in the community:

“...if a man can well judge his own interests, then ‘why not in the measures to be observed by public men in the direction of those’ interests?” (Bentham, IPML: ft. 40 in Schwartzberg , 2007b` :574)

The self-interested view of voting, coupled with an individualist conception of the public interest thus mitigates the degree to which pre-voting discussion can play a useful role: we are left with a conventional view that Bentham was an aggregative rather than deliberative democrat. It follows from this that since preferences are supposedly fixed rather than subject to transformation during deliberation, discussion before voting can –at most- be understood as the sharing of information between citizens who seek to satisfy their own preferences when they vote.

Second, contemporary views differ in their understanding of whether public discussion before voting has the power to transform expressed preferences. Often, but not always, this transformation is understood to take place towards the public interest, and seen to weed out incomplete, partial, or merely self-interested preferences (Benhabib, 1996). Here, again, the conventional view of Bentham appears to sit alongside “rational choice” aggregative rather than contemporary deliberative views because transformation towards

the public interest is unnecessary to good outcomes. It might, however, be compatible with other forms of preference change which are considered integral to deliberation. “Weaker deliberative” theorists like Gutmann & Thompson (2004:1-15) propose that deliberation performs a valuable clarificatory role in regard to decisions. It helps participants to sort through and reconsider their preferences in light of new information. Whilst fundamentally different from the kind of public-interested transformation demanded by some “stronger deliberative” democrats, it is possible that the premises outlined for the conventional understanding of Bentham’s view on public reasoning do leave some space for clarificatory discussion of this sort, which primarily exchanges information and signals between participants rather than deliberative reasons. In this sense, the conventional view sits alongside some contemporary aggregative approaches (Austen-Smith, 1992); it conceives of pre-voting talk as functional rather than transformative.

In respect to the third of our categories—the status of discussion—this naturally leads the conventional view of Bentham to a functional rather than moralised account of pre-voting discussion. The conventional view of Bentham interrogates the value of public reasoning in terms of the benefits it can produce for society through the production of better political decisions; primarily in respect to the subsequent aggregation of votes at the ballot box. Whilst it might, therefore, endorse various democratic practices as useful to the public utility, it would do so not in an absolute way; but only contingent on their usefulness. Unlike contemporary “stronger deliberative” theorists, therefore, the conventional view of Bentham does not ground fundamental political values like justice or legitimacy on the act of participation in public deliberation. Bentham’s work as cited as part of the conventional understanding of his democratic theory makes no mention of a need to reflect fundamental values like “natural duty of respect” (Rawls, 2003:373). Similarly, conventional Bentham did not portray public reasoning as the part of a rights-based definition of the legitimate exercise of political power (D’Agostino, 1996:57); nor did he claim that the right type of discussion between voters constituted vital democratic norms, as Habermasian democrats do.

The conventional reading of Bentham usually has less to say about the fourth of our categories: expectation of consensus. The conventional response to the previous three categories we have examined may suggest that there should be few reasons to expect full agreement from public reasoning. After all, conventional voters will consider their own interests at the ballot box and may, or may not, engage with one another in something akin to deliberation; depending on how useful it happens to be to the public utility. The individualist conception of the public interest renders consensus redundant to achieving political decisions which serve the public interest. On this conventional reading, Bentham certainly did not expect full consensus to emerge from properly conducted deliberation as “stronger deliberative” theorists like Habermas (1997) and Dryzek (2002) do. This conception of public reasoning is not “destined” for a “stronger deliberative” full consensus of all participants for the same shared reasons. It is

less clear, however, how far Bentham, on the conventional view at least, expected agreement to emerge from well-informed, self-interested voters exchanging information before casting their ballots.

The implications of the conventional view are clear. These observations on Bentham prompt the conclusion that he was a “rational choice” aggregative democrat. He emphasised counting votes over talking; encouraged a self-interested focus from voters; rejected transformative dialogue before voting; attributed no moralised deliberative functions to discussion between citizens. Finally, he did not suggest that such discussion—properly-conducted—should produce full consensus between participants.

If we reject this strong version of the “raw preference” view, however, there may be prudential grounds for pre-voting discussion in a conventional view of Bentham’s work. Public reasoning could be justified on grounds that it somehow allows individual citizens to glean a “better” or more accurate set of preferences which reflect their 'real' interests. On one construal of this view, some prudential reasoning may be possible; on the other none is possible. On neither view is any “moralised” deliberation--orientated towards the public interest—necessary.

As an aggregative theory that emphasises the secret ballot, the conventional view of Bentham is thus somewhat similar to the principles of contemporary social choice theory. We should assume that satisfying preferences should be the goal of government. The role of democracy is thus to count these preferences, and satisfy them as well as possible. Deliberation should not encourage citizens to transform their preferences. Bentham is often seen in this light: as part of the social choice tradition. The conventional view thus seems to concur with the social choice theorist's assumption of "citizen sovereignty": the idea that social choices are determined by individual desires (Arrow, 1963:29). This corresponds with Bentham’s own suggestion that individuals' judgments regarding their interests should not be overridden on paternalistic grounds (1983b: 131) and complements his quantitative utilitarianism. Thus Bentham agreed that public decisions should be responsive to individuals' preferences rather than exogenous conceptions of the "public interest". It is for this reason that Schofield (2006) and others understand Bentham as fundamentally rejecting the view that the unenlightened preferences of the masses needed to be fundamentally reformed by members of a political and social elite.

Each of these claims supplements the conventional, non-deliberative picture. Individuals’ preferences were the best available guide to their interests; the community was nothing more than the sum of its

members; and government should seek to satisfy the preferences that individuals happen to have. Without any clear view of the conditions under which preferences should be formed, or how they should be assessed, conventional Bentham seems to lack any plausible explanation of why talking should be fundamental to democracy, not just counting (votes). As we have seen, aspects of this characterisation of Bentham have permeated not only cursory studies like Keen's but also the critiques of sophisticated deliberative democrats (Gutmann & Thompson, 2004) and contemporary political philosophers (Gaus, 1999). For the reasons stated, conventional Bentham is highly unlikely to meet the challenge of providing a plausible account of the importance of discussion before voting.

Aside from the question of pre-voting discussion, the conventional view has had wider implications. Bentham's individualist definition of the public good; his insistence on maximising the good; and on the commensurability of a simple definition of happiness between people have provoked a number of well-known objections in contemporary political theory and ethics. Foremost amongst these is Rawls' accusation that utilitarianism fails to take seriously enough the "separateness of persons" (McKerlie, 1989). It may be appropriate for an individual to sacrifice his own welfare at one point in life in order to gain greater welfare in the future; for example through "consumption smoothing". Many contemporary political theorists, however, believe that redistribution of welfare between people is of a different character; that we cannot justify sacrificing the interests of one individual to promote those of another. In many instances, maximising the total welfare without regard for its distribution may demand this. This concern has a number of further implications. We may worry that utilitarianism as a simple maximising theory pays insufficient attention to the interests of minorities. If the suffering of one or two individuals can bring utility to many then this kind of theory could urge us to inflict such suffering.

Millian utilitarians have an obvious response to this. They may argue that the "permanent interests of man as a progressive being" are not served by redistributions of this sort. Brink and others have provided perfectionist readings of Mill which provide a number of potential responses of this sort (Brink, 1992). In contrast, Bentham was bound by his assumption of quantitative utilitarianism so may give no such response.

A critique of the conventional view.

It is clear that the conventional view of Bentham does reflect some important ideas in the development of his political theory. It is in some ways, however, an incomplete view of Bentham's work on public reasoning and democratic theory. Since this reading essentially casts Bentham as a "rational choice"

aggregative theorist, the normative strengths and weaknesses attributable to it follow from the particular characteristics of this type of view. Along with other aggregative theorists of this kind, like Riker (1982), conventional Bentham studies the interaction of voters by assuming that their preferences are self-interested. He analyses different ways of using those preferences to produce collective decisions. Below, I examine briefly some of the normative merits of that view, as described. Having done so, we shall then consider whether Bentham was, in fact, “conventional” in the sense described.

First, let us consider conventional Bentham in reference to the orientation, transformative power, status, and expectation of consensus that are attributed to public reasoning. In each of these respects, this construal of Bentham follows a “rational choice” aggregative approach. Relative to rival contemporary views of public reasoning provided by deliberative and aggregative theorists, therefore, this reading of Bentham provides a weaker response to the challenge of explaining the value of pre-voting discussion, but a stronger response to that of preventing manipulated discursive outcomes. To explain why, it is worth briefly summarising why these conclusions follow from the normative characteristics attributed to conventional Bentham.

The orientation of public reasoning in question is tilted towards the self-interest of participants rather than the public interest, or other wider interests. There are two separate points to note here. First, the conventional view of Bentham focuses on voting and the logic of institutions necessary for a successful aggregation of votes. It therefore reflects the fact that Bentham spent a great deal of his writing career explaining and analysing the functioning of aggregative democratic processes and—compared with contemporary deliberationists—relatively little time explaining the details of how a process of public discussion should lead citizens towards understanding the public interest.

The second point to note is that quite apart from not appearing in a conventional construal of Bentham’s political theory, a view of public reasoning which was directed towards the public interest in this way would actually be contrary to some of its principles. While contemporary deliberationists write with enthusiasm about the benefits of discourse in the public realm, and the changes that it can induce both in citizens and the decisions they take part in, the conventional view of Bentham approaches public-orientated discussion of this sort with some suspicion. The conventional insistence that individuals should vote according to self-interest means that conditions for public-interested deliberation are not given significant emphasis. It also means that public-orientation of this sort is precluded. Conventional Bentham

is thus both “rational choice” aggregative in approach, and apparently hostile to public-interested discourse before voting.

It follows that the conventional reading of Bentham, like other aggregative views which rely on self-interested voting, is in a poor position to respond to the challenge of explaining the normative value of pre-voting discussion between citizens. It does not hold that public-interested collective decisions require successful public-orientated discourse in order to be realised. Consequently, there is no strong reason to believe that public reasoning should be essential to good decision-making; or that it plays any particularly important role.

As we have seen, the lack of a public-orientation at the ballot box also has likely consequences for “rational choice” aggregative views in response to the challenge of preventing manipulated discursive outcomes. If, as conventional Bentham suggests, the public interest can be reached through an aggregation of self-interested votes at the ballot box, it follows that this view, like others, offers the chance to produce good outcomes through collective decisions without the offering more opportunities for outcomes to be manipulated during deliberation before voting. Whilst public-focused deliberation may have a number of other benefits, it does also imply the potential for discussion to be manipulated to favour some interests over others. In particular, depending on how the term is understood, the demand for discourse to reflect only “public” reasons also carries with it the potential for public and private reasons to be conflated, leading to manipulated outcomes (O’Flynn, 2010). That conflation can only happen if discourse is expected to refer to, and clarify the public interest rather than self-interest.

Denying the value of a public-orientated deliberation before voting thus puts the conventional understanding of Bentham in a weaker position to explain the value of pre-voting discussion, but a stronger position to prevent manipulated discursive outcomes. In particular, the conventional view implies that individuals’ expressed preferences should be considered an reliable reflection of their interests. This is essentially the opposite of the approach implied by many deliberative democrats. While those theorists imply the need for an “improvement” in preferences, through an appropriate process of consideration and deliberation; the conventional view simply aggregates existing views. Consequently, the conventional view is unlikely to give cause for citizens to be told that they have the “wrong” preferences. While deliberative theorists have been criticised for implying that a failure to reach certain conclusions is the result of deliberative failings (Posner, 2004), conventional Bentham would be unlikely to level such an accusation at any preferences, or conclusions.

As a “rational choice” aggregative view, conventional Bentham also denies that public deliberation before voting has the power to transform preferences towards the public interest, or in related important ways. This reduces the potential value of discourse before voting, because it cannot radically change the underlying preferences that individuals hold. Consider, for instance, two competing goals: a desire to reduce carbon dioxide emissions, and for consumer energy prices dipping below £0.10 per Kilowatt-hour. Assume that this price level cannot be achieved without scrapping “green levies” on resale prices and therefore increasing the carbon dioxide emissions of created by electricity consumption. Assume that a national referendum has been scheduled to decide whether green levies should be scrapped.

A citizen has an underlying preference for lower energy prices over environmentally-sound energy generation. Many deliberative views imply that an individual like this may well be prompted to revisit and revise her preference as a result of deliberation. It is not merely that deliberation provides information about how lower energy costs and more sustainable generation can be achieved. It can, rather, fundamentally cause the voter to change her mind about which goal should be prioritised, and which should be sacrificed. In contrast, public reasoning according to the conventional view of Bentham is non-transformative. It does not claim the ability to cause the voter to amend her underlying preference for lower energy costs. Citizens may present information to one another, and consider relevant evidence, but the purpose of this input will be primarily to help citizens understand how their preferences relate to the question at hand; not to transform that underlying basis for making a decision on how to vote.

Assuming that transformations happen in a positive way, and deliberation acts as it should do, deliberative theories are better placed to explain the value of pre-voting discussion in light of the transformative conception of discourse they advocate. In comparison, the conventional view of Bentham does not appeal to this potential for significant, positive preference change. Thus, it cannot appeal to these kinds of benefits to explain the value of discussion before voting. So, citizens following this blueprint for voting cannot “worsen” their preferences through a process of deliberation which has gone astray. Nor can they “improve” their preferences through an enlightening and positive deliberative engagement. This makes public reasoning less potentially damaging, but also less important. Like other “rational choice” aggregative views, conventional Bentham may invoke a less paternalistic attitude toward preferences than deliberative counterparts. As a result, it may be less susceptible to potential manipulation of discursive outcomes. Similarly, however, denying the potential for preference change also makes it more difficult for the conventional view to provide a plausible explanation of why public reasoning before voting should be considered an important, or valuable component of democratic procedure. These concerns are also reflected by the status that conventional Bentham accords to public reasoning. There is nothing in the conventional view of Bentham to suggest that public reasoning should play a moralised role. It is not, for

example, associated with reflecting, creating, or constituting fundamental democratic norms like justice or legitimacy.

As a consequence, the conventional view of Bentham does not provide any obvious answers to the challenge of providing an explanation of the value of pre-voting discussion. Its emphasis on counting votes rather than talking to generate better public understanding of political decisions means that the value of discussion can be explained in functional rather than moralised terms. Further, the right kind of discursive interaction between citizens is not portrayed, on this reading, as the basis for the legitimate exercise of democratic power, or for the basis of wider principles of justice. Conventional Bentham thus provides a characterisation of democracy which is something very close to what Barber calls mere “selection among preferences...[with] efficient decision-making as its sole measure” (Barber, 2003:198).

For the reasons discussed, however, the lack of a concept of moralised debate on the conventional Benthamite view is an advantage in reference to our second challenge of preventing manipulated discursive outcomes. Insofar as a moralised conception of debate is thought to apply to different contemporary views, it is also believed to imply that further conditions or duties should apply to the type of deliberation that citizens engage in before voting. These conditions, in turn, create some opportunities for discussion to be manipulated to favour the interests of dominant figures rather than the whole of the community. In particular, they have been associated with the potential for elites to make paternalistic judgements about the validity of the preferences expressed by members of the community. As we shall see in subsequent chapters, attitudes like these may increase the danger of manipulated discursive outcomes—produced as a result of social pressure to reach a conclusion within prescribed times. Similarly, there is little in the conventional understanding of Bentham to suggest that deliberation should be expected to achieve full consensus between citizens on complex and fundamental questions. Unlike the “stronger deliberative” views of Habermas and Dryzek, conventional Bentham does not suggest that achieving this kind of complete agreement for shared reasons is reasonable expectation for well-ordered discussion between citizens before they vote.

Indeed, it would be strange if an aggregative view which asks individuals to vote according to their self-interest did expect citizens to fully agree in a Habermasian way. After all, the conventional view denies the public-orientation and transformative power of pre-voting discussion. As we have seen, the stronger deliberative view which expects deliberation to achieve full consensus offers a strong response to our first challenge of explaining the value of discussion, but also creates opportunities for discursive manipulation,

particularly through the danger of false consensus. The conventional view of Bentham denies the assumption of full consensus and, therefore, is less likely to subject discussion to the danger of false consensus. It is thus at an advantage in respect to responding to our second challenge of preventing manipulated discursive outcomes.

We have seen that the conventional reading of Bentham has both normative strengths and weaknesses relative to contemporary views on the role of public reasoning. Irrespective of these normative merits, I believe that the conventional view is, however, an incomplete reflection of the parts of Bentham's writing which are relevant to this question. Whilst Bentham did advocate the secret ballot and express an aggregative definition of the public interest, his account of individuals' motivation was more nuanced than often suggested. In particular, it would be misleading to ascribe a simple (referential) raw preference utilitarianism to Bentham; particularly in light of his writing on epistemology. Far from being happy to accept referential errors in expressed preferences as a basis for voting, Bentham was determined to acknowledge the potential for referential error. Indeed, the correction of errors was central to much of his writing.

Consider, for example, "A Table of the Springs of Action" in "Deontology together with a Table of the Springs of Action and The Article on Utilitarianism". This was Bentham's attempt to chart the sources of motivation for different types of action. Here, as we might expect, he argued that pleasure and pain lie at the foot of all motivations to act (Bentham, 1983c:98). Ultimately, he believed, members of society often made referential errors in the way they understood people's motivations. These errors could result either from the judgement of an observer being swayed by his interest in the matter in question, or his intellectual weakness. Judgment could be "influenced, and perverted, by interest", or an observer being subject to some form of "mere weakness- intellectual weakness" (Bentham, 1983c:111). Bentham thus explicitly acknowledged and addressed the capacity for error. This featured under the heading "Causes of misjudgement and misconduct" (Ibid). Here, as in many other places, Bentham does not sound like a thinker who could not distinguish between preferences based on referential error, and those based on sound judgment.

These works also demonstrate that Bentham was not the simple egoist that he is sometimes portrayed to be. Bentham certainly emphasised the importance of self-interest as a motivation throughout much of his career. Mack, for instance, points towards his writing in 1790 which notes "the universal, necessary, and undisputed, and not even to be lamented property in human nature, the predominance of the self-regarding

affections over the social” (Bentham in Mack, 1962:455). Passages like these became more explicit and common in the later Bentham’s work, as his political theory also developed in a democratic direction. In “First Principles Preparatory to Constitutional Code”, for instance, Bentham assumes something which sounds like a simple egoism as a starting point for understanding how public opinion should work:

“On every occasion, the conduct of every human being will be determined by his own interest... meaning according to his own conception of it, to the conception correct or incorrect entertained in relation to it by himself at the moment of action”
(Bentham, 1989:68)

Alongside this, however, in his study of motivation in IPML, Bentham acknowledged the pleasure that individuals take from the interest and welfare of others. He was, nonetheless, careful to take account of self-interest as a predominant, if not absolute motivation for human action throughout both his earlier and later writing on institutional design (Hume, 1978:7).

So, Bentham argued, pleasure and pain lay at the root of people’s interests and motivations. The utilitarian roots of interests and motivations were not, however, always perceived clearly. For a variety of reasons—which we will revisit in subsequent chapters—people could and did err in how they perceived their own interests, and those of others. Bentham’s recently re-edited writing on sex emphasises this point. Bentham believed, and readily acknowledged, that prejudice could pollute public opinion, often on a persistent basis. He argued, for example, that the exclusion of women from the franchise had been based on prejudice rather than a proper understanding of interests (Quinn in Zhai & Quinn ed. 2014:79).

In recent years, Schofield (2006:1-5) has radically clarified our understanding of Bentham’s views in this area. Where previous scholarship had referred to Bentham’s theory of fictions in general (Ogden, 1932), Schofield more precisely identified the relationship between real and fictitious entities, and later instances of misjudgement which resulted from their conflation. Often, Bentham believed, confusions in language caused people to misunderstand interests in society. Prime among these confusions was the mis-use of fictitious entities in common language which often implied that they were real and physical rather than merely a construction of language. Whilst real entities could be perceived by the senses, fictitious entities could not. They had to be carefully analysed to test whether they ultimately had meaning. Ultimately, Bentham believed that meaning could be derived only from reference to real i.e. physical entities (Schofield, 2006:8).

Indeed, Bentham believed that this form of referential error—between real and fictitious entities—was a serious problem. Once a name was attributed to a fictitious entity, and it featured in common discourse, that named entity would often be mistaken for a real entity:

“to speak of an object by its universally known name, is to ascribe existence to it. Out of this error, misconception, obscurity, ambiguity, confusion, doubt, disagreement, angry passions and even discord and hostility have to no inconsiderable amount had place” (Bentham, UC cii 447 (9 July 1826) [Bowring viii, 327-8] in Schofield, 2006:77)

Bentham’s political arguments in many areas, and their roots in his theory of language as clarified by Schofield, do not sound like those of a theorist unconcerned with referential error. Bentham’s arguments about language, and the distinction between real and fictitious entities, also support the idea that he was an individualist i.e. defined the public interest as the aggregate of the interests of relevant individuals. Clearly, there is textual evidence to support this view:

“The community is a fictitious body, composed of the individual persons who are considered as constituting as it were its members. The interest of the community then is, what?—the sum of the interests of the several members who compose it.” (Bentham, 1970: 12)

This definition is put to use in many aspects of Bentham’s later political theory. To take one example, in the “Constitutional Code” (CC), he considers the performance of the legislature in reference to “the national interest... [which is] nothing more than an aggregate of the several particular interests” (Bentham, 1983b:43). Here, and elsewhere, Bentham believed that a general and abstract notion was comprehensible only in terms of the physical interests of individuals involved. Bentham’s individualism makes much sense in the context of his views on language.

If, as Bentham asserts, the public interest is a fictitious entity, and individuals, along with the physical utility they experience, are real entities, then there is no other way in which the public interest can be understood. Bentham’s theory of language, and the categories he dictates mean that the public interest can be nothing other than individualistic. There is no scope in his approach for an alternative, abstract concept of the public interest, which exists apart from the concrete utility of individuals. The individualism Bentham asserts thus runs deep. It is etched into the foundations of his worldview; not just his politics.

Whilst the grounds for believing that Bentham was an individualist are strong, those for suggesting he was a raw preference utilitarian are not. The utilitarian painted by Kymlicka is content to accept

preferences containing referential errors. But Bentham himself consistently tried to seek out and correct mistakes like this. Perhaps most famously, Bentham applied this idea of error to the popular notion of rights (Bentham, 1998b); arguing that the widely held basis for individual rights was mistaken since it was not ultimately comprehensible in real (physical) terms. He did not respond by concluding that existing rights-practice was justified; simply because it was valuable according to existing preferences. Rather, he believed that rights talk was either comprehensible in terms of harms and benefits to individuals; or it reflected nothing more than the personal view of a speaker. In Bentham's view, uncritical talk of individuals' rights was harmful precisely because "most commonly when examined into, the only meaning which it is found to cover is the latter": personal prejudice rather than a reflection of the harms and benefits encountered by individuals (UC cliii. 133-8 in Schofield, 2006:107). Referential error was thus fundamental to Bentham's famous rejection of rights. Such discourse was meaningless because it did not refer to real entities.

This is just one example of a wider theme. Bentham wished to distinguish real entities from the properties ascribed to them. He believed that doing so would allow individuals to understand their physical environment better; and maximise their own welfare. There is some evidence that this, in turn, fed into Bentham's views on how public opinion would operate in a democracy. Whilst he believed that individuals should be amenable to arguments based on the harms and benefits of particular measures, this was no guarantee that they always would be. Just as Kymlicka's diner might make mistakes in assessing whether the pizza in front of him is his favoured lunch option, Bentham's citizens might also err in their appreciation of the utilitarian reasons to vote one way over another:

"To prove that an institution is agreeable to the principle of utility, is to prove...that the people *ought* to like it: but whether they *will* like it or no after all, is another question" (Principles of Penal Law", Part II: 'Rationale of Punishment', Bowring, i, 390-525, at 411, quoted in Quinn, 2014: 74).

This related idea of referential error in language was essential to Bentham's writing on logic. His "central concern in his writings on logic was to understand the way in which language might be used both accurately to describe, and inaccurately or erroneously to misdescribe, the physical world" (Schofield, 2006:13). With these writings on logic and language in mind, we can conclude with some confidence that Bentham did not espouse a referential raw preference view of welfare. Contrary to Kymlicka's implication, Bentham would not have been content to satisfy a diner's misplaced preference to eat pizza which, unbeknownst to them, has been poisoned.

With some notable exceptions (Schwartzberg, 2007b; Rosen, 1983), relatively little attention has historically been paid to the implications that these kinds of error had for Bentham's political thought. There may be a couple of reasons for this. First, it is only relatively recently, in 2006, that Schofield has established just how significant and clear the link between Bentham's logic, writing on language, and later political theory was. Seen in this context, Bentham's attack on abstract references to individual rights follows from his earlier insistence that fictitious entities must be analysed and understood in relation to physical terms. Only through that process could they be properly understood, rather than causing confusion and harm.

Second, many interpretations have rightly been influenced by contributions of H.L.A. Hart, who edited the "Introduction to the Principles of Morals and Legislation" in its most recent form. Whilst revelatory and fascinating in many other ways, Hart's analysis gives fallibility of judgement, and the potential for error, a rather cursory treatment. In his "Essays on Bentham" (1982), Hart outlines only briefly, then dismisses, the basis for Bentham's belief in the fallibility of judgement. Rather than suggesting that the capacity for error is an important idea in Bentham's thought, he instead suggests that it is wholly based on a "naively subjectivist" account of beliefs, and consequently, speech acts, that cannot adequately explain the "common sense" meaning of claims that are common to everyday discourse (Hart, 1982:11). In support of that conclusion, Hart cites Bentham's view that "That to which expression is given, that of which communication is made is always the man's opinion nor anything more" (Bowring, VIII, 321). A simple interpretation of this view would suggest that:

(S) the only logical meaning of any speech act "that p" is in fact "I believe that p", and nothing more.

Hart's explanation of Bentham's rejection of claims to infallibility was therefore that any such claims must necessarily be false. The speech acts of supposedly infallible speakers, like other speech acts are—by definition—incapable of relating to anything more than their speakers' mind states. This position is criticised as a "naive" basis for rejecting claims to the possession of infallible judgement regarding the external world (Hart, 1982:12). In a state of affairs where (S) holds, claims to infallibility are necessarily false; the speech acts of supposedly infallible rulers are by definition, incapable of relating to facts about the external world. However, (S) also entails more than this. There is some textual evidence for Hart's view. From the "elliptical" basis of speech acts, Bentham draws two principal conclusions: "avoid dogmativeness...still more avoid intolerance". A central tenet of fallibilism (his belief that human

judgment is always liable to error) thus emerges: “never cease to bear in mind how slippery and hollow the ground on which your opinion, and consequently the utmost value of any expression which you can give to it, rests” (Bowring viii 330).

Bentham did hold that “the falsity of all such claims to infallibility was a consequence of some simple truths about the character of human judgements”. This claim is not, however, necessarily reliant upon (S). Indeed, Hart quickly dismisses Bentham's fallibilism. He suggests that “...here I think his limitations as a philosopher begin to appear” (Hart, 1982:12-3). Hart goes on to assert that the weakness in this epistemic argument also damages Bentham's wider account of the nature of law; it “infects his conception of commands, prohibitions, and permissions” (1982:12). He quotes Bentham's insistence that “of no matter of fact external to, of no matter other than that which passes in a man's own mind can any immediate communication be made by language” (Bowring, viii 321). The term “immediate” is central to this quotation's meaning. Hart interprets this as textual evidence for the centrality of (S) as a premise for Bentham's rejection of claims to infallibility. An alternative explanation is possible. The quotation above reasserts that—in political terms—the possibility of error in the exercise of judgement should always be borne in mind. The falsity of claims to infallibility was thus explicable in terms of the possibility of error which was a recurrent theme of human judgement in Bentham's epistemology.

I believe that Hart is right to imply that this would be a confused basis on which to reject claims to infallibility. Instead of identifying specific sources of error in judgements—thus statements—about the external world, this premise would instead undermine the very basis for the possibility of a reliable relation between speech acts and external facts. The Hartian criticism is, however, flawed on two grounds. First, Bentham cannot be understood as working within the same analytic approach as many 20th century philosophers; particularly those who followed the “linguistic turn”. As a consequence, the counter argument that his theory does not accommodate everyday language is not conclusive. Work done by Schofield (2006:1-27) on unpublished manuscripts confirms that Bentham's project was not to work from the foundation of everyday language; discerning and clarifying the truth found in its use. Instead, he wanted to apply a radical new epistemology, using direct sensory perception of real entities as a starting point, and expose the systematic misunderstandings and biases ingrained in the application of language. For Bentham, language was not so much an authoritative epistemic starting point as an established custom; riddled with error and demanding criticism from an alternative perspective. Consequently, he emphasised the practical steps necessary to expose fallacies and the false beliefs they engendered. Doing this required an active and critically engaged response to language. This could only be achieved when claims to infallibility were rejected: fallibility of human judgement “ought never to be out of mind” (Bowring, viii:300n).

The “fact” that a thesis implies one cannot acknowledge as legitimate everyday references to apparently external objects is thus irrelevant. A central tenet of Bentham’s project was to confront and expose the illegitimacy of many such references. The root to correcting systematic errors of this sort lay outside, rather than inside the established customs they embodied.

Second, it is not quite as clear as Hart implies that Bentham’s understanding of the meaning of speech acts was entirely subjectivist. Rather—as Postema has argued at length—Bentham also believed that the value of speech acts could be assessed and understood in reference to their utilitarian value (Postema, 1983: 40-58). In turn, this relied on an interaction with a series of external facts about the external world: not merely an interlocutor’s own mind state in isolation. Bentham believed that the presentation of false beliefs still had potential to promote utility because any perception of their apparent falsity could only be probabilistic. Consequently, there was rarely any reason to exclude the expression of any specific types of argument in public discourse. Bentham thus links infallibility and the liberty of discussion, pre-empting a subsequent argument by Mill: whether this “theory of impeccability or infallibility emerged in a madhouse or in a den of thieves would be a question suitable for a debating club were it not that under this theory “all debating clubs, all meetings that can produce discussion on topics the most interesting to human happiness, are silenced” (Bentham, 1989: 276, quoted in Schwartzberg, 2007: 574).

So much for the idea of referentially “raw preferences”; what of pre-deliberative preferences? Is Bentham a “raw preference” utilitarian in this sense? The first thing to say, of course, is that the idea of preferences as pre-deliberative depends very much on which conception of deliberation is used to frame that idea. For instance, it should be clear from what we have noted about Habermasian deliberation that Bentham did not endorse a “stronger deliberative” conception of pre-voting discussion. There is little to suggest that Bentham endorsed a fully transformative, moralised form of deliberation, expected to reach full consensus. It follows from this that Bentham probably was a raw preference utilitarian if we accept this stronger notion of deliberation and define deliberation in “stronger” Habermasian terms.

Beyond this, things get more complicated. “Weaker deliberative” theorists emphasise the value of pre-voting discussion without necessarily holding that this should produce full consensus between participants, or that deliberation holds a particular moralised value in reference to the democratic norms it reflects, generates, or constitutes. For a weaker deliberative theorist, preferences might be raw if they are formed without full information, reflection, or consideration. Here, there is more common cause between

weaker deliberationists and many examples of Bentham's democratic writing. Bentham argued that people need adequate information and consideration to understand their interests properly. There are many examples of this in his writing. In the "Deontology", for example, he argued that individuals have been misled to believe that sacrifices of their own welfare are morally valuable. This caused them to misunderstand their own interests. His task was to liberate people from the influence of this "gloomy system of theology" (Bentham, 1983c:122); to clear away misleading use of terms like "virtue" and "vice". In reality, he believed, terms like these had no value aside from their relation to happiness, which was "all that is ultimately and for its own sake worth regard" (1983c: 122). Some of Bentham's constitutional writing reflects a belief that mistaken beliefs of this sort can be exposed and eradicated through rational argument. In "First Principles...", for instance, he focuses on the role of delusion in a political system, and the harm that it had done.

Whilst he did not want to speculate about all of the psychological and historical sources of such delusions, Bentham was clear—as we will see in Chapter 4—on its serious effects. Delusion in public discourse caused erroneous opinions about interests, which damaged citizens' true interests:

“Not one history out of a thousand—not one biography out of ten thousand—that tends not to keep the mind plunged into this error: and this is one of the ways in which the making men dupes has for its effect, and to a vast extent for its object, their being made slaves”
(Bentham, 1983c: 265)

The strength of language here is important. Deluded citizens had not merely been led astray, but actively enslaved by their rulers (Ibid: 264). Given the importance of delusion in Bentham's constitutional writing, it would be misleading to conclude that he believed that the satisfaction of informed, considered preferences was of the same value as those of uninformed, unconsidered preferences. It was precisely because reflection and information was necessary to proper assessment of interests that Bentham denounced "prejudice". He defined this as: an opinion "embraced without sufficient examination...a judgement which, being pronounced *before* evidence, is therefore pronounced without evidence" (UC cii. 540 (6 Aug. 1811) [Bowring, ii. 478] in Schofield, 2006: 265). There are, therefore, good reasons to believe that Bentham believed that individuals could significantly err in their understanding of interests, and that information and reflection, along with the clarification of language, was needed to help correct these errors.

Bentham argued that people need to reflect on this information before voting rather than merely vote according to an unconsidered initial view. The details of how this reflection and collecting of information should take place will be our focus in subsequent chapters. There are, however, some clear indicators that Bentham wished voters to inform themselves and reflect before voting, and for legislators to play a role in helping them to do so (Quinn, 2014:85).

Finally, if we consider raw preferences to be personal—rather than social or communal—the picture is somewhat clearer. Whilst Bentham wanted laws and political decisions to reflect the public interest, he also argued in the strongest terms that individuals should vote according to their own interests. He believed that secrecy was important to voting. It was “the only security for genuineness of suffrage” (Bowring, iii, 599). Protection from the effect of public opinion and, critically, the approval or disapproval of powerful and exalted patrons, allowed voters to express their real views at the ballot box rather than courting the approval of others. This argument complements the famous assertion in *IPML* that self-interest is a predominant motivation for much action (Bentham, 1970a: 97-8), as well as an appropriate evaluative standard for individual actions (1970a:12). Indeed, in Bentham’s political writing, the “general predominance of self-regarding affection in man’s nature over social affection” is sometimes simply assumed as “a matter of fact” (1989: 264-5).

To recap, the conventional view results from three assumptions about Bentham’s political theory: raw preference utilitarianism, a quantitative conception of utility, and an individualist definition of the public interest. These are sometimes accompanied by a crude view of Bentham’s work on motivation. In reality, this acknowledged both social and selfish motives for action, and the predominance of the former over the latter. Taken together, these suggest a modest and functional, rather than expansive and extensive, role for pre-voting discussion between citizens. As a result, the conventional view of Bentham is relatively weak in response to the challenge of explaining the normative value of pre-voting discussion. It grants discussion a functional rather than consensus-producing, moralised role. It is also relatively strong in response to the challenge of denying opportunities for manipulation of discursive outcomes, exactly because it denies discussion such a significant role in democratic theory. Additionally, it provides few “paternalistic” grounds on which the preferences of ordinary citizens may be criticised; relatively to some “stronger” deliberative views.

In interpretive terms, the conventional view is an incomplete reading of Bentham’s work. It is fair to say that Bentham emphasised self-interested action by citizens and voters; both at the ballot box and before this. He did not, however, endorse “raw preference” utilitarianism; understood in terms of referential error. Bentham was not a “stronger deliberative” democrat, so did not conceive of raw preferences as

those preceding this form of Habermasian deliberation. He did, however, conceptualise the importance of error in the formulation of preferences.

Having examined the strengths and weaknesses of the conventional view, we shall proceed to investigate an alternative. It is clear that the conventional view does not convincingly respond to the first of our normative challenges: effective discussion; though it does have some strengths in reference to preventing manipulation of discursive outcomes. In the next chapter, we shall discover if an alternative reading of Bentham can fare any better.

Chapter 3. A participatory view of Bentham

Introduction

Contemporary theories of public reasoning focus on the status and role of discussions between citizens, including before voting, and the obligations that apply to participants. In contrast, it is relatively rare in Bentham scholarship to examine explicitly Bentham's political writing in these terms. Instead, many commentators focus on the role publicity plays as a feature of the constitutional system Bentham espoused, and its relation to the architecture of government.

In this chapter, I examine some examples of Bentham scholarship which provide potential grounds for understanding Bentham's views on public reasoning. These are "participatory" in the sense that they suggest significant grounds for citizens' participation in discussion before voting. Unlike the conventional view, therefore, they attribute to Bentham something more than a simply aggregative democratic view, which holds that aggregation of votes alone, without significant discussion, is sufficient.

There is sometimes a gap between the terminology and concepts used by Bentham scholars and contemporary normative theorists. Most Bentham scholars do not explicitly use the normative concepts and terminology which characterises contemporary debate between deliberative and aggregative democracy. For example, they have not often systematically examined the obligations of participants in discussion, or the effect that discussion is believed to have on the character of citizens, and the nature of arguments presented. Nonetheless, concepts explored by Niesen, Ben-Dor, Crimmins, Lyons, and Schofield do provide many of the resources required to understand Bentham's views on pre-voting discussion. Each provides a valuable insight into how Bentham wanted democratic debate to function. In doing so, they challenge suggestions—outlined in Chapter 2—that Bentham wished voters to engage solely at the ballot box, and not before.

Some of these participatory views—for instance those based on work by Lyons and Niesen—imply that we should reject the assumption that Bentham wished citizens to deliberate, and vote according to, their own self-interest. I argue that this would be a mistake; for two reasons. First, Bentham's argument for the secret ballot places self-interested voting at the centre of his rationale for democracy. Second, a self-interested view of pre-voting discussion fits more comfortably with the context of Bentham's work on language and epistemology. Attempts to equate Bentham's views of discussion with those of Habermasian "stronger deliberative" theorists are likely to mislead: Bentham consistently advocated a

non-moralised form of discussion. This was based on an informed understanding of self-interest; not a “moralised” realisation of the public interest, as suggested by some deliberative democrats. Other participatory views, like those of Schofield, do reflect Bentham’s focus on self-interest. I believe that these views provide strong foundations for understanding Bentham’s views; but this scholarship tends to focus on the idea of publicity and democracy rather than the nature of citizens’ participation. There is, therefore, an important challenge ahead: to understand how the concepts identified by these scholars translate into a view of public reasoning, and how that view fits within the context of other competing views in the contemporary debate.

In general—and for different reasons—these more “participatory” views of Bentham do not provide grounds for a view which both explains the normative value of pre-voting discussion *and* negates opportunities for manipulated discursive outcomes. Niesen and Ben-Dor suggest that a type of participation was essential to the outcomes that Bentham wished to create through political decision-making. Schofield identifies the important place that publicity played in Bentham’s constitutional thinking; which could lend itself to a fuller picture of the type of public reasoning required. In contrast, the “dual standard” that Lyons attributes to Bentham could suggest that individuals should participate in public deliberation and have a duty to reflect considerations relevant to the public interest in doing so. But this view provokes further difficult challenges. Broadly, participatory views thus offer the option of either stretching Bentham’s views to overplay similarities with contemporary deliberationists; or analysing the concepts in Bentham’s thought without applying them to the terms used in contemporary debate.

Participatory views of Bentham

Let us begin by defining what we mean by “participatory” Bentham scholarship. There is a distinct contrast between the conventional view of Bentham, as an aggregative democrat for whom prior discussion is a relatively unimportant mere preliminary, and what Mansbridge et al. describe as the “classical ideal of public reason” under which “individuals “enter a deliberation with conflicting opinions about what is good for the polity, but after voicing and hearing the reasons for different the reasons for different options, converge on one option as the best, for the same reasons” (Mansbridge et al, 2010: 66). That ideal is a hallmark of what I have called “stronger deliberation”. That view holds that deliberation should have some key characteristics. These include: focussing on the public interest, transforming preferences, constituting or creating important democratic norms, and expecting full consensus between participants.

Participatory views of Bentham imply that his democratic views are closer to this “stronger deliberative” perspective than often thought. They liken some aspects of his democratic thought to the kind of deliberation espoused by Habermas and his followers. On this view, talking literally takes the place of counting, making Bentham something of a deliberative theorist. The participatory view rejects the idea that Bentham merely relies on the aggregation, through voting, of raw, self-interested, unreformed preferences. As a result, it suggests that public reasoning is a critical stage in Benthamite democratic theory.

Different accounts reach this conclusion for different reasons. Some participatory views imply that citizens should deliberate together according to the public interest rather than self-interest. Unlike the conventional view, the participatory view suggests that citizens should deliberate, or at least discuss policy, together: this plays an important role in decision-making. Each instance of the participatory view is different. Some examples imply that citizens should deliberate together regarding the public interest. Others, in contrast, do not depart from a conventional assertion that Bentham wished citizens to consider their own interests before voting rather than those of the community. This is an important distinction, and one which has not always received the attention it deserves.

Consider, for instance, Niesen’s recent writing (2011), which builds a case for Bentham’s deliberative credentials on a reading of his evaluation of parliamentary discourse. Niesen’s direct focus is parliamentary, rather than public, deliberation. In the account below, I use his analysis of the former to shed light on the latter. I am aware that doing so steps over an important distinction. My aim is not to evaluate Niesen’s interpretation of Bentham on parliamentary deliberation, but to see how far a view of public reasoning based on it can go. There are two key questions here. First, can this account—based on Niesen’s work—provide a coherent interpretation of Bentham? Second, can it answer the two challenges outlined above: defining the role of effective discussion before voting and reducing opportunities for manipulation of discussion? In considering those questions, we shall use Niesen’s analysis as a starting point. Having done so, we shall later return to the distinction between public and parliamentary deliberation, arguing that it undermines the degree to which Niesen’s reading of “Political Tactics” (PT) could ground an understanding of how Bentham believed citizens should reason together before voting.

Niesen provides an interesting analysis of how “Political Tactics” (PT)—Bentham’s analysis of political procedure and behaviour in the newly established French Assembly—evolved as a text.

The kind of parliamentary reasoning Bentham endorsed was to be open to the public. Citizens needed to see and understand the rationale for political action in order to hold their representatives to account.

Niesen does not spell out the details of how this interaction should take place. It is perhaps in the dynamic interaction between parliamentary representatives and public scrutiny that we may find productive grounds for understanding Bentham's views on public reasoning. In "Political Tactics", Bentham suggests that a form of deliberation creates benefits across all levels of society. Niesen attributes those benefits to the way parliamentary representatives should interact with the public they represent:

"...decision-making in parliament profits from the dynamic relationship...[and] the public will then profit from its exposure to high quality deliberations. Gradually '[a] habit of reasoning and discussion will penetrate all classes.'" (Niesen, 2011:7)

The benefit of this kind of reasoning was clear: it made "sound opinion" more common (Bentham, 1999:31). To vote before deliberating was, therefore, to decide on a question while isolating oneself from the means of doing so in an informed way; to conclude without hearing the wisdom of peers. This places Bentham alongside deliberative democrats who claim that the process of deliberation causes unacceptably partial or biased preferences give way to "those that can form the basis of a public consensus" (Gaus, 2003:148).

The context for PT is particularly interesting. Bentham's original "Essay on Political Tactics" was written in 1791, but not published then. The best known version of this text, however, was published in French in 1816 as part of Dumont's "Tactique des assemblees politiques deliberantes". Dumont was more than just a translator. Rather, as a recent editor of "Political Tactics" has remarked, Dumont's were "reworked versions which, while preserving the essence of Bentham's texts, were shorter and more concise...[and] encumbered with less detail" (Pease-Watkin, 2003:4). In 1843, Bowring, the executor of Bentham's will, combined Dumont's text with the earlier 1791 essay as part of Bentham's collected works. This creates some particular challenges. Using these texts, it is impossible to know exactly how faithful Dumont's 1816 French text was to the preceding work on which it was based. As a result, we do not have a single, authoritative English version of PT from which to discern Bentham's views. When an updated and re-edited version was published in 1999, it took account of available original manuscripts. For many sections of the text, however, original texts had not survived. Consequently, the editors often had to rely on Dumont's work from 1816.

Niesen tries to address some of these challenges. He examines differences between available manuscripts from 1789-91, and the text published by Dumont twenty five years later. Those differences are particularly interesting on the subject of deliberation, and how it is described. The 1816 text refers to different aspects of the process of a group of representatives debating and making decisions together as "deliberation". In contrast, Bentham's original table of contents for the work refers to "proposing", "debating", and "voting" separately; it does not use this catch-all term (PT, 217 in Niesen, 2011: 2). By

the time Bowring published an amalgamated version of this work in 1843, the terminology had changed again; to “debate”.

Niesen suggests that the importance of these differences goes beyond terminology. In earlier manuscripts, Bentham made a critical distinction between the discursive and decisive phases of parliamentary procedure. He believed that these separate activities had been conflated as a consequence of the imprecise language used in debates about the French Constitution. Commentators observing the French Assembly used a single word, “deliberation” to describe the entire decision-making process, rather than identifying specific activities required of legislators at different stages. As Niesen notes, it is therefore ironic that Dumont himself made exactly this “error” in introducing Bentham’s analysis (Niesen, 2011:3). By separating those two activities, Niesen suggests, Bentham wanted to establish that deliberation and decision were different functions of a legislator. The former improved exercise of the latter. Better decisions resulted from prior discussion and consideration of the arguments. By implication, therefore, worse decisions were made when there was less opportunity for motions to be proposed and debated first.

The importance of deliberation should also be reflected in parliamentary procedure. Indeed, a failure to acknowledge the need for proper deliberation diminished debate and led to worse parliamentary decision-making. Without the time to propose, discuss, and consider motions *before* having to vote on them, assembly members would not be able to learn from the arguments, reasons, and evidence presented by their colleagues. In this respect, the terminological error made in describing “deliberation” had real consequences. Bentham believed that analysts of French parliamentary procedure had diminished the role of deliberation by failing to correctly describe the procedure needed to facilitate it. To illustrate, Bentham cited past practice in French Estates as an example. In those bodies, each representative voted in sequence: after making his own contribution to the debate, but before hearing from others in the group. In doing so, each was “deliberating” in the broad sense used at the time; i.e. making a decision together as a group. In Bentham’s view, however, they were missing what was essential and valuable about parliamentary deliberation: learning from, and persuading, one another *before* deciding how to vote.

Indeed, voting before fully debating a question was inimical to the kind of parliamentary deliberation that Bentham considered essential. Conversely, allowing assembly members to debate questions in turn and exchanging and considering reasons and evidence together, is central to Niesen’s “participatory” reading of Bentham:

“...He who speaks last and "into whose lap the collected wisdom of the whole assembly is poured in a full tide, sees no one to whom he can give the benefit of illumination but himself."
(PT: 95 in Niesen, 2011:5)

This argument suggests something rather different from the “simple” or conventional view of Bentham considered in Chapter 2. “Conventional” Bentham assumes that isolated individuals express fixed preferences through voting, and remain independent of one another before this. In contrast, Niesen’s “participatory” Bentham suggests that representatives should vote only after proposing and debating together to consider the merits of relevant questions. Rather than merely expressing fixed preferences, Niesen suggests that representatives should be encouraged to change their minds during debate (PT: 94 in Niesen, 2011: 4). Democratic decision-making should thus be understood more as a forum in which participants learn from one another, than a marketplace in which they register individual preferences.

On this view, the merits of discussion between representatives before voting are those of the influence of “understanding on understanding”. We have seen that aggregative and deliberative accounts of public reasoning differ in a number of respects. One of these is the transformative power attributed to discussion before voting. Some aggregative theories suggest that voters should be considered as independent of one another before they vote. In contrast, deliberative theories actively encourage interaction between voters as part of democratic participation. A central part of this interaction is the capacity of participants to persuade each other to change their minds: to transform the preferences they previously held. The force of persuasion using reasons is thus principally a feature of deliberative, rather than aggregative, conceptions of public reasoning.

The features of Niesen’s reading of Bentham more readily fit into the deliberative side of this divide than the aggregative side. Participants are seen not as isolated individuals but rather as co-operators in a common “exercise of their intelligence” (Niesen, 2011: 5). Bentham famously distinguished between expository and censorial analysis. The task of expositor was to show and understand what had already been done by judges and legislators. The role of the censor was to demonstrate what they ought to do in future (Schofield, 2006: 51). In “Political Tactics” (1999), Bentham as expositor outlines some details of parliamentary procedure and how they interact practically when applied to political decision-making. Bentham the censor evaluates these by reference to a utilitarian definition of the public interest. As censor, he concludes that procedural restrictions, of the French Estates in particular, were a barrier to good decision-making.

The importance of representatives changing their mind before voting was thus central to the “common enterprise” of deliberation and it was critically important to reflect this in the design of parliamentary procedure. Without this, the community would not be able to realise decisions which reflect the shared wisdom of the community, realised through discourse: “Bentham's criticism is that the French practice rules out changes of mind on the part of earlier speakers and thereby disadvantages later speakers” (Niesen, 2011:5). Niesen’s work thus suggests that pre-voting discussion has an important role to play, and that this is related to participants’ capacity to change their minds before voting.

We saw in Chapter 2 that Bentham defined the interests of the community as those of the individuals who compose it (Bentham, 1970a:2). In these terms he was, as Postema (1986) defines it, an “individualist”. That individualism lends itself to the conclusion that the public interest can be reflected through an aggregation of votes at the ballot box rather than informed, or special forms of discussion before voting. How might individualism be squared with the participatory reading suggested by Niesen? Rather than disputing Bentham’s individualism, Niesen acknowledges that Bentham believed that the public interest could—in the abstract—be reached through an aggregation of individual interests. It was a product of *counting* of interests, rather than *talking* between deliberators. The importance of pre-voting discussion could, however, be understood in practical rather than abstract terms.

Niesen suggests that deliberative principles become vital to Bentham’s democracy not because the abstract aggregation of individuals’ interests was undesirable, but because it was difficult to achieve in practice. Bentham wished legislators to make rational decisions based on an informed assessment of the utilitarian impact of different proposals. Niesen’s Bentham also believed that they should, in practice, pursue utility through a form of deliberation. The importance of this form of debate, on Niesen’s view, was established by the necessary difficulties involved in conducting a comprehensive and complete calculation of relevant interests. While the latter was, for Niesen’s Bentham, ideal, the former would often be the best available route to maximising utility. Niesen’s interpretation of Bentham thus contends that the right kind of discursive interaction helps to secure the merits of “well-ordered deliberation” which can consequently “be seen as a functional equivalent to the elusive [felicific] calculus that Bentham had suggested” (Niesen, 2011: 13). Utility is thus served through public-spirited deliberation, practically akin to that advocated by Habermas (1997). This argument is instrumental: it suggests that deliberation is made valuable by the benefits it creates for public decision-making. It follows, therefore, that deliberation procedures between citizens are justified to the degree that they perform that role.

The right discursive conditions thus give participants access to reasons that accurately reflect the interests at stake. Presenting information and arguments during deliberation facilitates a decision that takes each claim into account, and assesses relevant evidence in light of available information. Indeed, contemporary utilitarians sometimes assert that this form of reasoning is implicit in the way Western polities debate policy. Riley provides one example of this. He cites Ely's claim that American government is "a sort of applied utilitarianism—unfortunately possessing utilitarianism's weaknesses as well as its strengths—an institutional way of determining the happiness of the greatest number". Ely insists that the root of the appeal of democratic government in America is thus utilitarian: "the formation of public policy, at least in this country, [America] begins with the questions how many are helped, how many hurt, and by how much" (Ely quoted in Riley, 1990:336). The claim here is not that utilitarianism should lead to aggregation of votes without prior discussion, but that discussion should be guided by utilitarian principles.

Similarly, Niesen's Bentham suggests that by exchanging proposals, reasons, and arguments with one another, representatives in the Assembly mimic the merits of an abstract utilitarian calculation of interests. Their discussions act as an approximation of exact utility calculations. They allow the interests of each individual to be appropriately considered as a constituent of the community's interest. Just as Bentham used an abstract calculus to count the interests of each person equally, so the "right" discursive conditions perform a similar purpose. They allow each party's needs and preferences to be presented, considered, and judged. The procedures followed in parliaments should thus shape and facilitate debate to play this role. They should prevent the interests of some individuals being afforded greater regard than those of others.

Niesen illustrates using a number of examples from "Political Tactics". For instance, there is to be no hierarchy among speakers. This is to be avoided, and mitigated, by rejecting permanent distinctions between frontbenchers and backbenchers (PT, 54 in Niesen, 2011: 9). Doing so would prevent the merits of arguments presented being distorted by received misled notions of authority and prestige. Similarly, Bentham wanted contributions to debates in the Assembly to actively respond to the contributions of others, in a creative way. This creativity resembles the progress that some contemporary theorists believe citizens achieve through, and during, deliberation. Bentham rejected the idea that speeches should be pre-written before they were presented (PT 77, 132 in Niesen, 2011:9). In contemporary debates, this is resonant of the claim, made for instance by Jon Elster (1986), that a public form of deliberation is likely to

encourage individuals to present public-interested rather than self-interested arguments, as they seek reasons which will garner support from the widest group of their audience.

We have not yet considered what Niesen's participatory reading implies about the orientation of debate. On this question, Niesen's references to Bentham's view on parliamentary procedure are instructive. Those procedures were to be designed to encourage impartiality from speakers. They were expected to consider and reflect the interests of the community; not merely or those of the groups they represented directly. For this reason, every speaker was to speak from a neutral tribune; in contrast with British parliamentary practice of the time (Niesen, 2011: 7). This is a key difference between Niesen's participatory view, and the "simple" or "conventional" reading of Bentham on public reasoning. That interpretation suggests that aggregating self-interested votes allows decisions which serve the public interest. Aggregation therefore makes public-interested deliberation redundant, and an impartial approach from deliberators unnecessary. "Conventional" voters are free to be "partial": there is no need for them to reflect anything other than their own interests at the ballot box. In contrast, Niesen's reading of PT suggests that Bentham wanted deliberators in the French Assembly to be impartial, in an important sense. They were to consider interests closer to those of the whole community, and further away from the preferences with which they initially entered the forum.

On Niesen's view, measures like this were intended help debate in the Assembly operate as a real world version of the felicific calculus. The inputs into that calculus were meant to reflect what Mill later called Bentham's dictum. This was his famous insistence that "Every individual in the country tells for one; no individual for more than one" (RJE in Bowring vii, 334).³⁰ Analogously, debate in the Assembly was to ensure equal consideration of the interests of all, and Niesen's argument suggests that this demands a measure of impartiality from deliberators.

Niesen's participatory reading thus reflects a kind of public justification that is also emphasised by contemporary deliberative democrats. This implies that "having to defend one's policy preferences in public leans one towards using public reason" (Chambers, 2004:390). The parliamentary representatives described in PT engage with the public when they speak in debate; not just with one another. Consequently, they speak to defend their positions to a wider group that "demands more from such debates in the way of justification than would be necessary in order to defeat the parliamentary minority (Niesen, 2011: 7). As Niesen points out, there are some important and interesting passages in PT. These

³⁰ RJE (Rationale of Judicial Evidence, specially applied to English practice). See (Schofield, 2006:84).

challenge the idea that Bentham believed in aggregation, to the exclusion of deliberation. Indeed, Bentham explicitly rejected the simple aggregative view sometimes ascribed to him by deliberative democrats. We have seen that independence of voters is a foundational assumption of much aggregative theory (Condorcet, 1785). Bentham explicitly rejects that assumption in “Political Tactics” (1999:145).

Rather than replacing discussion with aggregation, therefore, Niesen’s Bentham instead urges use of appropriate modes of discourse to reach decisions that best serve the community’s interest. Niesen’s understanding of Bentham thus resembles some claims advanced by contemporary deliberative democrats. Both espouse the benefits of deliberators having to justify their arguments to a wider (public) audience. For this version of the participatory Bentham, those benefits are exclusively instrumental. As we discovered in Chapter 1, different deliberative democrats ascribe different types of value to deliberation. Indeed, Niesen himself makes this point. He cites Gutmann & Thompson to suggest that they have more in common with Bentham than they acknowledge (Niesen, 2011:8).

Niesen does not himself specify the full details of the type of reasoning which should take place under this account. I believe, however, that his references to Bentham’s attempt to promote impartiality, are instructive. So too is his emphasis of the critical importance of deliberators changing their mind during participation before voting. Together, these arguments suggest something akin to public-orientated deliberation, with its potential to transform preferences; rather than aggregation using fixed preferences. This reading is thus closer to contemporary deliberative democracy than it is to the “conventional view” of Bentham as exclusively aggregative.

Having outlined Niesen’s reading of Bentham, we should now turn to consider its strengths and weaknesses. Niesen’s work has the great merit of challenging the idea that Bentham was simply an aggregative democrat with little to contribute to contemporary debates about the role of democratic discussion between citizens. There are, however, two reasons why it should not form the basis of a reading of Bentham on public reasoning. One relates to the editorial history of *Political Tactics*; the other to the scope of that work.

Since Niesen’s reading relies on “*Political Tactics*” (Bentham, 1999), it must necessarily focus primarily on parliamentary rather than public discourse. That work is not an explicit statement of Bentham’s views on public reasoning. Instead, it is more an explication of his beliefs about how a parliament should function. Above, we have attempted to consider whether that reading could provide the basis for an interpretation of Bentham on public reasoning. Unfortunately, the editorial history of *PT* presents some

significant barriers to us using the text in this way. First, consider the question of editorial history. As I have attempted to explain, the history and provenance of PT is complicated, even compared with Bentham's many other works. Editors have access to original manuscripts to corroborate, or contradict interpretations of many other texts. Following editorial work at the Bentham Project, we now have access to reliably edited versions of many key texts. These include "First Principles preparatory to Constitutional Code" (1989), the first volume of the "Constitutional Code" (1983b), IPML (1970a) and many other important and relevant texts. Each of these provides a clear insight into Bentham's views on important questions. Each broadly reflects the order and content of what he intended to publish. Editors of each text have made extensive use of original manuscripts throughout the editorial process to corroborate the decisions they made.

The same cannot, unfortunately, be said of PT. A re-edited text is now available to us, but original material used to produce it was necessarily limited. As a consequence, there is greater uncertainty around the reliability of "Political Tactics"; relative to other texts from Bentham's corpus. I do not claim that PT should be disregarded. Rather, it should be given less weight than other texts; particularly where apparent contradictions emerge between them. Given this, it would be counterintuitive to base a reading of Bentham's views on public reasoning—or any other significant matter—exclusively or predominantly on this text. After all, there is still real uncertainty around which parts of PT could reflect Dumont's views, and which faithfully report Bentham's. "Political Tactics", as it was published in 1816 then 1843, was somewhat detached from, rather than regulated by Bentham's personal oversight. Much of the text that we must still rely on remains an inexact re-translation into English from an earlier translation into French, which may or may not have been wholly accurate. It is possible that the original translation of this text into French, and its subsequent re-translation and editing back into English, both perfectly reflected Bentham's views. It is possible that they did not. While this editorial history does not invalidate PT as a source for Bentham scholarship, it should moderate the degree of confidence we place in this single work as a reflection of Bentham's beliefs.

My assumption here is that—in general—we may more reliably infer an author's intentions from works whose publication they oversaw, and from original manuscripts they wrote, than from texts which have been edited, translated, and retranslated in their absence. In recent years, Rosen (1983) and Schofield (2010) have demonstrated just how significant editorial decisions like these can be. Schofield showed, for example, that Bentham did not hold the positions attributed to him by Hart's attempt to describe a "utilitarian tradition" of jurisprudence (Lieberman, 2014:121) and apply the principles of this tradition to contemporary legal debates. Just as Hart's editorial decisions and subsequent interpretations had an

impact on how Bentham was understood in the past fifty years, it is likely that decisions taken by Dumont also had an effect in the 19th century.

Second, consider the scope of PT. Niesen himself is acutely aware that “Political Tactics” addresses a very specific question. It is not an attempt to prescribe general conditions for useful discussion, nor to explain how citizens should engage in deliberation and voting. Instead, PT provides an examination of how parties do and should behave within the constraints imposed on them by membership of an assembly. It considers how the design of that body—and behaviour of those within it—should be amended to better serve the community. This does not make the principles it describes, by definition, inapplicable to other contexts. Rather, we must consider how those principles could read across to the wider public, and whether Bentham intended them to be used in this way. Clearly, a simple read across from parliamentary to public deliberation could imply that Bentham wanted members of the public to behave like effective assembly members. That might entail that they should consider reasons to vote from a position of some impartiality; having regard for the public interest rather than just their self-interest, and engage in a form of open deliberation which can transform the preferences of those involved.

There are, however, clear grounds to conclude that Bentham held public and parliamentary discourse to be very different activities; with very different outputs and expectations. Consider some of the important differences between parliamentary and public participation. Bentham famously argued citizens should vote in secret to protect their ability to reflect their self-interest (Schofield, 2006:144-5). In contrast, he often wanted parliamentary representatives to vote in public, so that they could be held accountable for their service of the public interest (Ibid: 280). I believe that fundamental differences between parliamentary and public reasoning mean that “Political Tactics” cannot justify ascribing to Bentham a public-orientated, deliberative view of public reasoning. To claim, therefore, that Bentham is a “stronger deliberative” democrat who wished citizens to deliberate directly about the public interest would contradict, or ignore some of the most important arguments in his democratic theory. These include his view that public-interested voting was unnecessary and dangerous (Bentham, 1998b: 430-2); and his belief that expressed notions of the public interest were often delusory (Schofield, 2006: 264). Consider how fundamentally different the public discourse Bentham later advocated in the “Constitutional Code” (1983b:36) was from parliamentary discourse advocated in “Political Tactics”. In the former, discussion is defined as having an unlimited, open membership and an open agenda. In the latter, only assembly members are assumed to take part in deliberations, and Bentham defines strict procedures for setting the agenda for debate (Bentham, 1999:77). We shall return to these important differences in Chapter 4.

Niesen's analysis reinforces my scepticism about the potential for "Political Tactics" to be used as a basis for understanding Bentham on public reasoning. Later editions of PT did emphasise the importance of "deliberation". References in sections of the text which can be traced directly back to Bentham do not, however, give sufficient confidence that Bentham himself was referring to the kind of "stronger" deliberation espoused by contemporary theorists. Bentham himself did emphasise the importance of representatives changing their mind during debate (Bentham, 1999:72). Here too, though, he did not provide sufficient detail about the character of this process to suggest that this amounted to the kind of full preference transformation espoused by theorists like Habermas (1997). For example, it is possible that the "changes of mind" referred to by Bentham are the result of representatives becoming better informed about how to realise existing preference(s), rather than the consequence of transformation through "stronger" deliberation. Niesen's interesting analysis improves our understanding of how Bentham approached aspects of parliamentary procedure, as it aims to do. It would, however, be inappropriate to stretch that excellent analysis to explain Bentham's views on how members of the public should reason together. Indeed, there are significant reasons to reject such a move. To summarise, these include: important differences between Bentham's arguments for parliamentary and public discourse; the uncertain editorial history of PT; and the existence of more reliably edited texts elsewhere.

In normative terms, I believe that Niesen's view meets the first of our challenges: explaining and defining the importance of pre-voting discussion. It does so because it places discussion, and information-sharing between individuals, at the centre of democratic decision-making. If well-ordered deliberation is the best functional equivalent of the felicific calculus—and we accept the utilitarian basis from which Bentham approached political theory—then it follows that effective deliberation is essential to the kinds of democratic outcomes Bentham sought: those which maximise public utility. It is clear from the critique of the French Assembly put forward in PT that Bentham believed that expressed preferences better reflected the public interest after deliberation than before it. On some basic level, therefore, this contention could form part of an explanation of the value of pre-voting public discussion in a democracy. The conditions for deliberation, which PT sets out, speak to the same insight cited by some contemporary deliberationists. Fundamentally, this is that the consideration of interests and principles, through the right sets of discursive conditions, is an essential component of democratic participation and decision-making. Of course, for Bentham, individuals' interests alone ultimately define good outcomes. Those theorists who argue that ideally conducted deliberation embodies an intrinsically valuable virtue (Gutmann & Thompson, 1996) will, therefore, be disappointed by this utilitarian response.

This response will not satisfy deliberative theorists who insist that deliberation is intrinsically valuable; but no utilitarian account ever could do. For this reason I do not take this as a necessary condition of a successful response to this challenge. There are other deliberative critics that Niesen's participatory thesis

cannot satisfy. Those who believe that appropriately conducted deliberation embodies a deep and valuable commitment between individuals as free and equal citizens will be unlikely to be convinced. How could a merely instrumental argument capture the value of deep connections of that sort? Rawls, for example, conceived of public reason in those terms (1997). Rawlsians will thus deem inadequate a view of public reasoning which neglects intrinsic duties held by citizens towards one another. The participatory view of Bentham suggested by Niesen is thus illuminating in reference to Bentham's understanding of parliamentary deliberation. It cannot, however, help us by providing grounds for an interpretation of Bentham's views on public reasoning. Our search for an interpretation of Bentham's views on this matter must therefore move on to different primary texts and alternative analyses. Let us consider an alternative analysis of Bentham.

Ben-Dor has also argued that Bentham's democratic theory relies heavily on a developed conception of how public opinion should operate in a democracy. The argument is presented in *Constitutional Limits and the Public Sphere* (2000); a work which outlines various aspects of a view regarding how public discourse and public decisions should interact. This is an impressive and innovative reading of Bentham's political theory that provides a novel way to understand how Bentham wished citizens to reason together. Indirectly, the form of public reasoning that Ben-Dor advocates asks citizens to consider something more than their own self-interest. It does not thus share a "simple" self-interest orientation with the conventional reading. Equally, this form of reasoning does not have a simple public-interest orientation. To understand the relevance of this reading of Bentham, let us examine some of its key proposals.

Ben-Dor attempts to tackle some very wide issues in legal theory. Among other things, he is interested in the normative foundations for legitimate laws when citizens have fundamentally different—often incompatible—ethical views. In examining a range of answers to that question, he argues that an appropriate response should be procedural rather than substantive. Rather than stipulating a single set of reasons that grounds legitimate laws, theorists should instead identify the right procedural conditions for citizens to participate in their formation. Ben-Dor's version of Bentham describes appropriate discursive conditions for forming public opinion; not the substantive ends it should serve. Through this discursive process, citizens are supposed to question the very ideas of harm and benefit—broadly construed—as well as the activities to which they are applied in practice. Seen in this light, the objective behind Bentham's political theory is less to promote utility directly, and more to allow citizens to do so themselves, through discourse:

"Instead of justifying morality upon a priori substantive metaphysical ideals, his main goal was to liberate the public sphere and to unfold communicative possibilities within it" (Ben-Dor: 2000:237)

Individuals could thus only begin to engage in what Ben-Dor calls "utilitarian consensus-formation" (Ibid) once they accept that public discourse regarding the ends of public decisions was desirable, and necessary to democracy. Doing so, in turn, required them to reject the claims of religious and political authorities which demanded unquestioning acquiescence. An analogous position is now relatively common in contemporary debates on deliberation. Habermas, for instance, is often understood as part of a tradition which seeks to establish deliberation as a "sovereign guide" to resolving fundamental disagreements (Gutmann & Thompson, 1996:17). Ben-Dor's reading is unusual because it portrays Bentham as part of that tradition. In doing so, he is quite explicit that some aspects of his argument are inspired by Bentham, rather than attributable to him. Nonetheless, in large part, Ben-Dor's view is also a reading of Bentham on public opinion. It is that reading of Bentham (on public reasoning) that I shall outline and respond to here, rather than Ben-Dor's response to wider questions in political theory.

Unlike Niesen, Ben-Dor does not rely on "Political Tactics" as a primary source. Instead, he analyses how public opinion is supposed to operate in Bentham's wider constitutional writing. Also unlike Niesen, Ben-Dor does not suggest that deliberation should be construed as a proxy for the felicific calculus. Instead, he holds that effective public reasoning promotes convergence between citizens around the principles and norms that they believe serve the public interest. Ben-Dor focuses on the role of the "Public Opinion Tribunal" (POT) to explain how public opinion could promote political outcomes and behavior which served the community.

The POT was a "fictitious entity", which was officially recognised in Bentham's proposed constitution. It appears in various different texts; each time with a slightly different purpose. For instance, Bentham's writing on the Panopticon prison plan includes a form of POT: designed to aid the rehabilitation of inmates. The focus for Ben-Dor's book is however, more narrowly on how the POT operates as part of a constitution; how it contributes to the creation of justified laws. For that purpose, Ben-Dor focuses on the "Constitutional Code" (CC), which provides an interesting picture of how Bentham wished public opinion to contribute to democratic governance. This aspect of Bentham's corpus has received more attention since publication of the first volume of the Constitutional Code, edited by Rosen, in 1983.

According to Ben-Dor, the POT provides a template for how public opinion helps to make legitimate democratic decisions possible: by facilitating a special kind of consensus. He identifies building blocks of this position in the references Bentham made to two ideas: the possibility of agreement between citizens;

and his concept of the “social sanction”. We shall examine each below. On consensus, Bentham expresses a general expectation that citizens discussing joint ends in the right sorts of ways should be expected to reach agreement (Bentham, 1983b:36). This is perhaps surprising. We have seen that Bentham expected individuals to—predominantly—be guided and driven by their own pleasure and pain: not that of the community as a whole (IPML: 1). Given this, we might expect agreement between citizens to be elusive in cases where their individual interests diverge. Indeed, it is unsurprising that contemporary aggregative accounts that assume the self-interest of voters also tend to emphasise aggregation rather than—deliberative—consensus (Riker 1982).

For Ben-Dor, the POT addresses that challenge: it explains how Bentham expected predominantly self-interested individuals to reach agreement; despite their different interests and circumstances. Following principles set out in the POT, public discourse does not produce consensus by some happy accident; or because individuals disregard their previously divergent interests for the public good. Instead, the process of discussion is itself specifically geared towards generating consensus. It helps to produce agreement by creating an identification of interests between the individuals within a community and that community at large. Rather than simply allowing consensus to emerge from discourse, Ben-Dor’s participatory view makes use of Bentham’s extensive writing on persuasion and incentives, found in the “Deontology” (1983c) and elsewhere. This “participatory view” of Bentham on pre-voting discussion between citizens thus operationalises individuals’ ability to judge public ends in concert while still being motivated by self-interest. Public reasoning bridges the gap between citizens’ self-interest, and the need for political decisions that serve the whole community. It does so by forming a special kind of consensus, which helps to ensure that conceptions of the former become compatible with the latter. When successful, it causes self-interested individuals to consider “what is good for me?”, and produces answers which serve the whole community.

Ben-Dor’s Bentham believed that public discourse should take place regarding utilitarian ends: the harms and benefits created by different pieces of legislation. Seemingly, this is an obvious implication of Bentham’s utilitarianism. It also underscores much of Bentham’s constitutional writing. As Niesen notes, the basis for justification of political proposals in a parliament ought to be both public and utilitarian (Niesen, 2011:9). Since the rightful ends of political action concern the pleasures and pains they create for the individuals in a community, it followed that the rightful object of political discourse should also be a form of instrumental reasoning. This meant considering prospective pleasures and pains, and the probabilities of them being realised. Debate on these utilitarian terms would not, however, guarantee agreement. There could be substantial disagreement about which outcomes would help or harm individuals; even if there was consensus about the outcomes that a particular measure would create. After all, individual citizens could assess those outcomes in different—perhaps opposite—ways. The possibility

of this kind of disagreement is implied by Bentham's earlier writing on the principle of utility, and its alternatives. Bentham asserts in IPML, for instance, that observers should approve of an act in proportion to its tendency to produce to utility, but may not do so. In fact, he believed that approval was often wrongly attached to acts that destroyed pleasure and promoted pain. He labelled this attitude "asceticism" (IPML, 17-21). The possibility of asceticism, and similar evaluative errors, meant that it was far from certain that people would agree on particular policies; even when evidence pointed towards their utilitarian merits. For Ben-Dor, the role of the POT in Bentham's constitution was to ensure that approval and disapproval were allocated to promote utility, rather than negating it.

The use of powerful "sanctions" was essential to that endeavour. In IPML, Bentham described four of these, which could be used to shape the behaviour and attitudes of citizens: "physical, political, moral...and religious..." (IPML, 34). To understand the role of the moral (or "popular") sanction in Ben-Dor's account, let us briefly step back and explain these sanctions more generally. The physical sanction emerged in the course of nature; without any need for intervention from other people. In contrast, Bentham believed that the religious sanction related to the pleasures and pains that people expected to experience after death. He went to great length to analyse whether religious faith could thus be understood in utilitarian terms (Schofield, 1999). The political sanction was embodied in the rulings of judges, under the power of the state. The moral sanction, in contrast, was created by what an individual believed others thought. In particular, it had often been created by the opinions of "significant individuals" (Schofield, 2006:35). Ultimately, this moral sanction was grounded in the approval or disapproval of the community: their "sympathy and antipathy" (Ibid: 36). This sanction thus gave an opportunity to encourage utility-maximising motives and actions, and to discourage harmful alternatives.

Bentham believed that the same event, and the physical pain it created, could be an instance of any of these sanctions. He cited the example of a man whose goods or person are consumed by fire. That same event could be an instance of the physical sanction, if it occurred by accident. It could also be the result of the political sanction, if it resulted from punishment mandated by the state. The moral sanction may have been its source if the fire occurred as a result of a neighbour withholding help in disapproval of his moral character. Finally, it could be caused by agents being distracted by the thought of suffering in the afterlife: the religious sanction. Bentham thus proposed sanctions as a way of understanding the source of (dis)pleasure, rather than as a set of categories for utility itself (IPML: 36). While the political, religious and physical sanctions were predominantly operated through other means, the moral sanction relied on public opinion. Consequently, Bentham thus also termed it the "popular sanction" (Ibid). Whatever the source of the pleasure or displeasure in question, he believed that individuals naturally wished to seek out the former and avoid the latter, for themselves at least. The expectation that a sanction would operate in

response to an activity would thus help to shape how individuals behave. Famously, it was for those two twin masters of pain and pleasure “alone to point out what we ought to do, as well as to determine what we shall do” (Ibid:11).

Quinn has written extensively about how Bentham believed a legislator should encourage the rational allocation of sympathy and antipathy in her community, and respond to errors in these allocations: those that did not promote utility (Quinn, 2014:74-5). This work illustrates two things. First, that Bentham believed public opinion could err in its ascription of sympathy and antipathy to outcomes; second, that it could—nonetheless—play an important role in democratic debate. Predominantly in his earlier writing, Bentham termed this means of shaping conduct “indirect legislation”. He defined that term negatively. It was used to cover all measures that a legislator might take to prevent misconduct, without explicitly prohibiting it:

It was “whatever...can be done in the way of law’ which is not direct legislation, which does not, that is to say, prohibit acts identified as sufficiently mischievous as to be designated offences”. (Quinn, 2014:82)

Thus, while direct legislation was concerned with the punishment of harmful actions through prosecution, indirect legislation also allowed the encouragement of beneficial actions by non-legal means. It was a powerful tool placed in the hands of the legislator: a “subtle engineering of conduct based on the understanding of the motives for action, in order to manipulate...an individual’s perception of what is his or her best interest” (Brunon-Ernst, 2012:68). A public-spirited legislator could thus use indirect legislation in pursuit of the public interest while also satisfying the self-interest of individuals in a community. It is here—in the critical use of the moral, or popular, sanction—that Ben-Dor suggests the Public Opinion Tribunal plays a crucial role. The POT is thus a way to understand how citizens can make use of the sanctions Bentham described: to make future agreement possible (Ben-Dor, 2000:108). It does so by promoting an “intersubjective consensus” between citizens considering public decisions (Ben-Dor 2002:109). This utilitarian mode of thinking thus allows individuals to co-ordinate their interests with one another.

Effectively, Ben-Dor’s reading suggests that the POT makes indirect legislation a vehicle of public opinion, rather than one of fan enlightened legislator. It places the power to influence the social incentives offered to others in a community in the hands of the people rather than those of a legislator. It allows citizens, together, to shape their social environment: the background against which preferences are determined. Doing so makes agreement possible. Effective public reasoning could, therefore, encourage the growth of a group of egoists who genuinely believe that their own interests are served by actions that

help the community. It encourages consensus on the consequences of political decisions which are harmful or beneficial. Critically, however, this type of reasoning demanded an unlimited group of participants. Constraints on the types of people able to take part in the POT were—Ben-Dor’s Bentham believed—likely to produce results that favoured sections of a community rather than the public. Consequently, the POT could not be constituted by political parties or other sectional groups. That would lend itself to discussion reflecting their interests rather than those of the public. The tools of persuasion, through the moral sanction, which could help to shape citizens’ conception of their own interests were powerful; only a public unencumbered by party loyalties could be trusted to deploy them to proper effect (Ben-Dor, 2000: 108).

By discussing the consequences of various political decisions—the possible benefit or harm which might result from different possible outcomes—citizens begin to equate the terms of their own welfare with those of the whole community. The moral, or popular, sanction is essential to this process. It allows a community to promote views which contribute to the public welfare—by expressing its approval—and dissuade those which detract from it, through disapproval. Ben-Dor thus reconstructs Bentham’s beliefs about how the moral sanction could operate to explain its role within the POT. In doing so, he relies predominantly on Bentham’s “Table of the Springs of Action” rather than IPML. That work acknowledges the importance of a “principle of sympathy” which motivates individuals to help others. When allied to rationally-justified moral norms—expressed through public opinion—therefore, “love of justice” could be a powerful tool to promote utility. It could include: desire of self-preservation (if justice involves the individual concerned); sympathy for someone suffering injustice, now or future; antipathy for someone profiting from, or causing, injustice (1817:97).

Just as someone usually motivated by self-interest might act on her “love of justice”, so she might also act on a desire for public approval: in conformity with the moral sanction. A predominantly self-interested individual considering a practice which harms the public interest must consider the disapproval she will receive from the community. As “someone profiting from, or causing injustice” she will be subject to “antipathy” from the community and experience pain, and the absence of pleasure as a result. In this sense, it could be in her self-interest to abstain from actions which harm the public good if this is what the moral sanction, expressed through public opinion, dictates. In this way, just as Adam Smith (1759) believed that the sentiment of sympathy could promote moral behaviour, analogously, the (dis)approval of the community can materially shape a citizen’s interests. Thus although Bentham argued that public opinion needed to be liberated from irrational political and religious influences to which it had been enslaved, he also argued that social influences on opinion could be harnessed to promote progress. If repeated consistently and persistently, this process could fundamentally shape how individuals conceive of their own interests. The approval of her peers therefore provides the individual with an incentive to

gradually conform her preferences to the public utility, because only by doing so could she also gain the public approval from which she derives further utility.

Ben-Dor's version of the POT uses this principle. It starts with largely self-interested individuals with different ideas about what serves utility and—over time and through dialogue—gradually helps create a group of citizens who agree about what serves the public interest. The POT also encourages a long-term consensus because, by carefully and dynamically allocating approval and disapproval, it encourages citizens to understand their own interests in a way which serves rather than detracts from that of the community. This form of public opinion thus aligns individual and community interests *without* expecting citizens to consciously become any less self-interested. Just as Bentham believed that a wise legislator should use punishments and indirect legislation to deter crimes which harmed the public interest so, Ben-Dor's POT uses the incentives created by the moral sanction—operated through public debate—to promote agreement on the public interest. A simple example may illustrate this principle. A community might encourage the perception amongst peers that it is in one's own interest to hold and discharge responsibilities towards others: to vote; to donate to charity, to vote in elections, or volunteer for "good" causes. Over time, the POT may consistently express approval of actions like these which—for argument's sake—serve the public interest. In response, self-interested individuals will gradually begin to factor the approval of their peers into their decisions on how to act, and which activities to value. Exercising public opinion in this way thus makes egoists more likely to perform public-serving acts. More broadly, and indirectly, it also makes a public orientation from individual egoists more likely. If I derive utility from social approval, and can gain that approval from valuing the same ends as my peers, I will begin to factor that public reception into my plans: to incorporate the social sanction into my thinking. The incentives provided are thus analogous to those cited by Elster (1986).

Having outlined some features of Ben-Dor's interpretation of Bentham, we must now consider its strengths and weaknesses. As many commentators have remarked, Ben-Dor's fascinating and ambitious argument challenges common views of how Bentham should be understood in relation to normative debates about how justified laws are formed. Some aspects of his view are, however, controversial amongst Bentham scholars (Harrison, 2003:257). I believe that some of this controversy is well-founded. There are three particular reasons to question Ben-Dor's argument as a basis for understanding Bentham on public reasoning. First, Bentham's substantive arguments do not sit well with a procedural interpretation. Second, there is no extensive evidence to confirm that the POT's principal role—as described in the "Constitutional Code"—was to apply the moral sanction to citizens, rather than rulers. Third, the terminology of "utilitarian consensus" is misleading. It was never used by Bentham, and carries connotations which conflict with important aspects of his work. Let us examine each of these objections below, in turn.

First, I do not believe that Bentham can be understood as a purely procedural theorist of public opinion. Ben-Dor's aim is to elucidate the discursive ideals of Bentham's constitutional theory without relying on Bentham's own substantive conception of morality (Ben-Dor 2000:237). That objective risks neglecting some of the—explicitly utilitarian, substantive—reasons why Bentham himself came to advocate democratic governance. Ben-Dor advocates a view of justification based on the conditions under which citizens can deliberate together. A more standard interpretation of Bentham would start with his utilitarian definition of value, and analyse why democratic procedures are valuable, in those terms. As Harrison notes, this more traditional reading of Bentham has some distinct advantages over Ben-Dor's procedural alternative. In particular, it explains why Bentham made such extensive efforts to improve the substantive conclusions reached by public opinion, rather than just the discursive conditions producing it:

"What Bentham does to natural rights, Ben-Dor, on Bentham's behalf, does for all evaluative foundations. In so doing he thinks that he has made his hero a greater Samson. But what he has got may be a blinded Samson with no other option than to pull down the house on both himself and all critical and legal theory" (Harrison, 2003:257).

Disregarding Bentham's substantive arguments grounded in utilitarianism therefore risks disregarding his substantive grounds for the other conclusions, including democracy. It is clear that this aspect of Ben-Dor's argument is a significant departure from Bentham's own views. It is one of the reasons that Ben-Dor himself describes parts of his account as Bentham-inspired rather than directly attributable to Bentham himself.

This departure has other implications. One is that Ben-Dor's Bentham would be unable to critically evaluate the substantive ethical conclusions reached by citizens in discourse. Rather, he would focus exclusively on the way discussion takes place. I do not believe that this kind of view is compatible with the way public opinion is characterised in much of Bentham's work. Bentham was not ambivalent or agnostic about the substantive conclusions reached by public opinion. Rather, he believed that the individuals who generated public opinion could err. He explicitly encouraged legislators to deter them from doing so:

"Every nation is liable to have its prejudices and its caprices, which it is the business of the legislator to look out for, to study, and to cure" (IPML, p.183 quoted in Quinn, 2014:85).

A purely procedural account would undermine the substantive beliefs which motivated Bentham's reformist zeal. These were underscored by what Bentham believed were the dictates of utility. From his intent to root out and correct delusions in "First Principles..." (1989) to the notion of error in the "Constitutional Code" (1983b: 265); from the capacity for something close to false consciousness in chapter 9 of IPML (1970a), to the argument that the ability to amend a constitution was essential to correct for mistakes made by representatives and the public (1998b: 276); Bentham relied on a substantive evaluation of the outcomes of debate. I believe that he did so too often, and too insistently, to permit us to now portray him as an exclusively procedural theorist of public opinion with credibility.

Second, while the early Bentham emphasised how important indirect legislation should be to legislators who seek to shape the behaviour and opinions of the community, it is not clear that the POT was intended to do so. There is significant evidence that the importance of indirect legislation gradually diminished as Bentham's political theory became more democratic. The early Bentham believed that an enlightened legislator could help to achieve optimal outcomes through such means; but the later Bentham was intensely distrustful of political elites (Schofield, 2006:109-114). Ben-Dor reflects this by construing the POT as a vehicle for public rather than elite control of the moral sanction. What is less clear, however, is that the POT Bentham describes in the "Constitutional Code" is meant to apply the moral sanction principally to citizens rather than rulers, as Ben-Dor implies.

Consider how the Public Opinion Tribunal appears in the 1983 edition of the "Constitutional Code", edited by Rosen. We shall return to this question in more depth in Chapter 4. It is, however, worth noting two brief pieces of evidence which count against Ben-Dor's interpretation of that text. First, the POT is explicitly listed as a check against misrule by those in power, under "Securities against Legislative, and Judiciary" (1983b: 39). Second, throughout the text, there are no explicit references to the POT applying the moral sanction to citizens. Again, in "Political Tactics", Bentham explicitly describes the POT as a vehicle directed toward the powerful rather; not the population at large. It "always decides the destiny of public men; and that the punishments which it pronounces are inevitable" (1999:29-34 in Schofield, 2006:260). As we shall see in Chapter 4, I believe that POT is better understood as being directed toward holding public officials accountable. This sits more easily with Bentham's rationale for democracy than the claim it should engineer consensus.

Finally, alongside these substantive challenges to Ben-Dor, there is a linguistic one. Throughout "Constitutional Limits and the Public Sphere", he repeatedly attributes to Bentham a belief in "utilitarian

consensus” (Ben-Dor, 2000:112). Ben-Dor cites the “Constitutional Code” as evidence for this conclusion. He is right to say that Bentham anticipated agreement between citizens who reason together clearly. Bentham believed that much disagreement was the result of the use of unclear, non-utilitarian terms of debate which obfuscated relevant interests, as his arguments on fallacies suggest (1989:65).

It is misleading, however, to describe this agreement as a fully-fledged “consensus”. Bentham never used that term. Throughout available texts from the “Constitutional Code”, he refers to the possibility of agreement rather than a full “consensus”. This may sound like a trivial distinction, but I believe it is significant. The contemporary *meaning* of consensus is absent from Bentham’s work, as well as that particular term. In contemporary political theory, “consensus” has a specific and significant meaning, with connotations that Bentham himself would have been unlikely to subscribe to. Consider one prominent example. For Rawls (1997:766), consensus conveys something far beyond mere agreement. It means not merely that citizens agree a response to “fundamental political decisions” but that they do so for the *same*, shared reasons: those which are justifiable to all reasonable persons. This is a world away from the kind of utilitarian agreement Bentham mentions in his political writing. Like Hume, Bentham famously opposed the idea that the authority of the state could be justified by a hypothetical contract between citizens (Wolff, 1993:87). That argument was unnecessary, because direct utilitarian justification was available by appeal to the interests of those concerned. It would be odd, therefore, to ascribe a proto-Rawlsian view to Bentham by supposing that he wished to promote some form of consensus, which is defined in reference to reasonable persons rather than the actual persons who constitute a community.

One implication of Rawlsian public reasoning is that citizens have a duty to make relevant arguments on the basis of reasons that are justifiable from within all reasonable persons’ reasonable fundamental beliefs: what he calls “reasonable comprehensive doctrines” (Rawls, 2005:214). Now, there are some references in Bentham’s work that make an analogous appeal. For instance, in “Political Tactics”, Bentham encourages parliamentary representatives to justify their arguments in general rather than sectarian terms (Bentham, 1999:122). The rationale for this view is, however, instrumental. Making a political case in broad, non-sectarian, terms could increase the likelihood of its success. Bentham’s claim is therefore very different from contemporary Rawlsian arguments. Indeed, as a utilitarian, Bentham would have no reason to necessarily prefer an agreement reached on political principles as a “modus vivendi” (in Rawls’ terms) to one reached for shared reasons. In contrast, that distinction is fundamental for Rawls. For him, public reasoning is part of a process whereby principles “initially accepted for instrumental gain... [become] an essential part of individuals’ worldview” (Stears, 2010:221).

Thus, while Bentham expresses his expectation of agreement in the CC, and elsewhere gives parliamentarians pragmatic advice to avoid sectarianism, there is little to suggest that he intended those references to convey adherence to a belief in full consensus, for the same shared (Rawlsian) reasons. This point about the meaning of “consensus” thus complements our observation about Bentham’s use of the word. The term appears nowhere in his vast body of work. Indeed, it was popularised in the nineteenth century, after his death. In that context, we should be careful in attributing to Bentham a term or idea which developed later. Without strong evidence that Bentham believed in the same ideas as contemporary theorists of consensus, it would be wrong to attribute the term to his thinking on public opinion. I do not believe that such strong evidence exists. Consequently, attribution of this particular term to Bentham’s writing thus risks projecting (deliberative) assumptions and beliefs onto his work without justification for doing so.

Interpretive barriers to Ben-Dor’s reading are significant; so too are the normative challenges. According to Ben-Dor, public reasoning is more than just the most efficient means available of satisfying existing interests in a society. It is also a way of reshaping the way individuals conceive of their interests to promote socially desirable ends. Just as Habermasian citizens shape norms for their society through the act of deliberating, so Ben-Dor’s citizens deliberate, and mutually shape their own interests by doing so. The moral sanction thus operates through public opinion. It serves a similar function to that proposed for discourse by contemporary “stronger deliberative” democrats. Of course, this gives pre-voting discussion a far more important position in democracy than that afforded by a conventional reading of Bentham. Perhaps this understanding of public reasoning is closer to the contemporary deliberative conception of preference change. It allows citizens to move beyond an initial impression of what their self-interest demands. Rather than departing from self-interest, Bentham instead believed that citizens—on Ben-Dor’s view—help to engineer conditions in which they, and others, conceive of it in more socially useful ways.

Public reasoning thus induces a convergence of interests. Bentham’s famous belief that individuals are, in general, motivated by self-interest, thus does not entail a “rational choice” aggregative view. That view expects that individuals should vote according to self-interest and we should aggregate results to reach the public interest. Instead, Ben-Dor’s argument accepts egoism and claims that it makes discussion essential rather than unnecessary: it is the only means by which individuals’ conception of their own interests can be aligned with measures that serve the whole community. Without this kind of reasoning, some self-interested individuals would not be motivated to perform some of the basic civic and social duties necessary to maintaining the stability and security of a society. In this sense, public reasoning can be said to be essential to the maintenance of a flourishing Benthamite society: one which serves the public interest, defined as the interests of individuals who compose the community. This is an instrumental explanation of public reasoning. It is one which, however, makes the criterion against which the outcomes

of political action are assessed itself responsive to a discursive process. Political action should serve the community's interests, which are composed of the interests of individuals. In turn, individual interests are themselves shaped by public discussion.

There are close parallels between Ben-Dor's and contemporary discourse theory. Like Habermas, Ben-Dor suggests that the norms which define the value of democracy are themselves defined—in part—through a procedure of deliberation between citizens. This is a process through which “participants question and transcend whatever their initial preferences may have been” (Habermas, 1997: 449). This is a convincing response to the challenge of explaining the normative value of pre-voting discussion. It makes that act integral to the aims of a democracy. Without a public process for deliberating political ends, there might be little hope of reaching a collective decision which serves the public interest: individuals might approach, and exit, the voting booth with fundamentally divergent preferences.

Ben-Dor and Niesen perhaps offer the most explicit departures from the conventional view. There are, however, other perspectives from which it is possible to interpret a view of public reasoning. One such potential avenue is provided by Lyons in his 1973 work “In the Interest of the Governed”. In that text, Lyons outlines and then departs from a common understanding of Bentham's utilitarianism. The opening chapter of “An Introduction to the Principles of Morals and Legislation” (IPML) is usually understood to state Bentham's view that there is a single, simple standard by which all actions can be assessed: the principle of utility. That principle “approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question” (1970a:12).

This single standard is usually understood as universal in scope. As Bentham implies, it applies to every action whatsoever and “therefore not only of every action of a private individual, but of every measure of government” (Ibid:12). Thus, at a basic level, the rightness or wrongness of the actions of an individual can be evaluated by reference to likely consequences. Acts of parliament can also be subjected to this same, single standard. This “universalistic” reading of Bentham entails that the utilitarian evaluating options must consider “Whoever is likely to be affected by an action (or by its alternatives, since an action is to be compared with its alternatives)” rather than a particular or parochial group (Lyons, 1973:23). Lyons rejects that conclusion. Instead, he contends that the standard for evaluating laws and actions put forward by Bentham differed fundamentally according to whose interests were at stake. In turn, this means more than simply considering who is likely to be affected by a particular act. Thus while the

principle of utility is Bentham's ultimate standard, which determines whether an act is right or wrong; it "alone does not suffice to tell us which acts are right or wrong" (Lyons, 1973:20). To do that, we need to assess the likely consequences of an action on relevant interests. Only then can it be evaluated according to the principle of utility. The relevant interests will, however, be different for "public" and "private" actions.

As a consequence, the "first principle" of utility translates in two distinct and particular ways of evaluating actions, "with community interest as the test within the public or political sphere, while self-interest is to rule in 'private' matters" (Lyons, 1973:20). In effect, this reduces the scope of both public and private action. The test for the former is condensed from universal interest into a test for serving the interests of a defined and relevant community. Similarly, the test for the private action reduces radically from maximising the welfare of anyone who could be affected, to simply serving one's own interests. Lyons' main project was to propose a revision of Bentham's psychology and ethics. "In the Interest of the Governed" does not, therefore, explicitly consider public reasoning or the role of active citizens at and before the ballot box. It would not, however, take a huge leap to see the implications that a Lyons' dual standard could have for this question. If "public" actions are to be assessed by their tendency to promote the public interest, and individuals' private actions should be evaluated against their service of self-interest, which category could voting fall into? As Lyons points out, Bentham consistently argues that the role of government is to serve the interests of the whole community; not those of particular individuals, or groups.

The objectives of political action, voting, and public deliberation may thus be considered public rather than private. From this, we could conclude that citizens should vote according to the public interest. This might then lead to the conclusion that voters should deliberate according to the public interest before voting. An evaluative standard for voting seems to imply a set of expectations about pre-voting discussion. The dual standard thus potentially suggests something akin to "stronger" deliberation which is directed towards the public interest, and conducted in the expectation of consensus. Public-interested deliberation of this kind could then create a set of subsequent demands and obligations for citizens. When engaging in the "public" activity of deliberation they could, for instance, be obliged to appeal only to those reasons—whatever they might turn out to be—which promote an agreement which serves the community's interest.

As we have seen, an argument of this kind could potentially be justified by appeal to “Political Tactics”. In that text, Bentham argues that an explicitly atheistic argument for abolishing blasphemy laws would be unnecessary. Rather, the same conclusion could be argued for in more general terms which do not alienate theists in the community (1999:122 in Niesen, 2011:10). Bentham does not himself describe these as “public” reasons; but we could understand these kinds of general, rather than sectarian, reasons—which appeal to a broad base of perspectives—in those terms. Lyons’ distinction between public and private actions could thus suggest a similar distinction between the private actions people take, for themselves and their family and friends, as individuals, and those they engage with as citizens, for the community.

Aspects of Lyons’ dual standard are controversial amongst scholars; particularly in reference to Bentham’s ethics (Hume, 1978). It is not necessary to accept these critiques, however, to see that a more participatory reading of Bentham—along the lines suggested—based on Lyons’ interpretation would be misplaced. Suppose that Bentham did subscribe to two different evaluative standards for public and private actions, as suggested. It is possible that voting and deliberating could be considered public actions: the appropriate ultimate objective for each should be the community’s interest rather than the individual’s. It would follow from this that the actions taken by citizens as part of a process of deliberating, then voting, should be assessed by their tendency to promote the public interest.

On utilitarian grounds, however, it does not follow from this that the deliberative participation of citizens should be consciously directed towards the public interest, as some deliberative theorists suggest. The numerous, extensive contemporary and historical debates on indirect and direct utilitarianism illustrate this point emphatically. They demonstrate just how aware many utilitarians are of the (possible) divergence between principles and motives used to make decisions, and the underlying rationale which ultimately justifies those principles (see Lyons, 1994). Indeed, to reach the conclusion that public-spirited deliberation is necessary, as a result of Lyons’ dual standard, a further assumption is necessary. We must also assume that direct pursuit of a particular goal, like the public interest, is the best way to realise it. Like many subsequent utilitarians, including Mill (Alexander, 1985), Bentham rejected that assumption. Consider two examples from his work. First, Bentham’s view of judicial adjudication is usually thought to imply “judges would not decide at case-level by applying directly the principle of utility, but they would rather apply general rules which, in turn, were enacted on utilitarian grounds” (Ferraro, 2010:1). Similarly, Bentham famously provided “subordinate ends” as a guide for legislators. In doing so, he explicitly acknowledged that principles grounded in utility often served that end better than its direct pursuit (Quinn, 2008). These examples confirm that even if Bentham did believe that public deliberation and voting should be evaluated against its tendency to promote the public interest, it does not follow that participants must explicitly pursue that end. It therefore also does not follow that deliberation which

focuses on public ends, or particularly public reasons is necessary to promote the public interest. A separate argument would be required to show that this is the case.

A dual standard reading of Bentham thus does not count in favour of public-orientated deliberation before voting. Indeed, there are further reasons for rejecting that view. As we have seen, Bentham explicitly wished citizens to vote in a private, individualistic, and self-interested manner (Schofield, 2006: 263). His arguments for franchise reform and the secret ballot thus—while seeking to promote the public interest—aim to protect voters from the dangers of public influence rather than exposing them to its benefits. If Lyons' grounds for a dual-standard reading of Bentham are questionable, these are further and compelling reasons to reject his view as a basis to interpret Bentham as a public-interest deliberative theorist. Lyons' argument provides a fascinating, if debateable, revision of the basis of Bentham's utilitarian views. For the reasons outlined it cannot, however, be stretched to suggest grounds for a more deliberative reading of Bentham's views on public reasoning.

Unlike Lyons, Crimmins has explicitly suggested that Bentham can be likened to contemporary deliberative democrats like Habermas and, therefore, does provide something close to an account of public reasoning. Crimmins has briefly examined Bentham's views on the Public Opinion Tribunal (Crimmins, 2011:154-5) as part of a wider survey on Bentham's utilitarianism. His interesting revision of Bentham recasts Bentham's famous dictum that "everybody to count for one, nobody for more than one". The original source of this dictum, Schofield has established, was the "Rationale of Judicial Evidence", which was edited by Mill and published in 1827 in five volumes (Schofield, 2006:84). Bentham's original wording in that text was that each is "to count for one and nobody for more than one" (Mill, 1861 in Schofield, 2006:84). This statement is usually understood as a substantive claim about the outcomes that the principle of utility demands. Crimmins, however, reinterprets Bentham's commitment. He suggests that it could relate to democratic participation rather than the outcomes it creates.

Citing Habermas, he contends that "Bentham's theory of representative democracy bears more than a passing resemblance to the deliberative model...[in part, since] it implicitly recognises the need for public reasoning, for reasons that are acceptable to all as free and equal citizens" (Crimmins, 2011:155). Crimmins suggests that this similarity is grounded in the six conditions of ideal Habermasian deliberation: unconstrained participation by all citizens; participants expressing their views and preferences in a truthful way; assertions must be presented in a logically coherent way; the principal objective of participants must be the common good; participants respect, and carefully listen to, one another; participants are open to

changing their mind in response to a better argument. Without elaborating much further, Crimmins then suggests that Bentham's democratic views are "open to the same objections as the deliberative model"; principally that citizens may well struggle to communicate in a free and open way without a degree of material equality to mitigate the potential undue influence of money over on debate. He thus portrays this obstacle as an objection to both Habermasian deliberation and Bentham's own views on public reasoning.

I think that this accusation is misplaced for a number of reasons. First, while it is possible that some deliberative democrats are not sufficiently aware of the potential influence of material inequality on discursive conditions, this does not apply to Habermasian deliberation. Indeed, in his early writing, Habermas paid great attention to the interaction between economic conditions and discursive equality. It is precisely because he acknowledged the capacity of socio-economic conditions to impede proper deliberative participation that he assumed substantive social equality as a starting point for discussion (Habermas, 1996:308). Subsequently, a number of deliberative democrats have followed that lead. They argue that an existing distribution of power and resources must not affect chances to contribute to public deliberation. Egalitarian conditions of interaction must be satisfied in ways which are manifest to citizens in order for deliberation to function properly (Cohen, 1989:18).

Of course, it may be that these conditions of substantive equality are difficult to realise, but "stronger deliberative" theorists usually account for this. They accept that real world conditions can be assessed by their proximity to the ideal. The stronger deliberative thesis is therefore not usually a binary choice between ideal conditions for interaction and rejecting the value of deliberation altogether. Consequently, it is not fair to accuse Habermasian deliberative theorists of failing to take account of—at least some of—the real world conditions that could impede effective deliberation. Perhaps it is difficult to assess exactly how far real world circumstances are from this deliberative ideal. Making that assessment might itself be highly contentious and difficult; to such a degree that the practicality or cogence of the idea of a deliberative ideal is itself undermined. That is not the objection that Crimmins suggests, however. His argument is rather that "[i]t is easy to see that making the ideal a reality is fraught with obstacles" (Crimmins, 2011:155). Of course, it is possible that the benefits of achieving an ideal form of Habermasian deliberation justify the difficulties and costs of overcoming those obstacles. This is a possibility that Crimmins neither entertains nor responds to.

More importantly, Crimmins provides little evidence to support his claim that Bentham's democratic views can be equated with the Habermasian project. His version of Habermas insists that "the principal objective of the participants must be the common good" (Ibid: 155), but we have seen that there are good reasons—rooted in the argument for the secret ballot—to believe that Bentham did not demand a public

orientation from citizens during pre-voting discussion. Similarly, Crimmins echoes arguments put forward by Gutmann (1980:26) to imply that Bentham's views on public reasoning construe that process primarily in terms of its ability to hold those in public office to account; to prevent abuses of power. This is quite different from the Habermasian approach. Ordinarily, this attributes more ambitious normative claims to the deliberative process. For instance, Habermasians ordinarily suggest that deliberation has a transformative power in reference to participants' preferences; and that it can be expected to produce full consensus between participants. Clearly, for Habermasians, deliberation is more than just a "negative check" against abuse of political office. Rather, it is a positive, enlightening and transformative process that leads individuals towards a fuller appreciation of relevant interests and arguments.

It is in light of the apparently negative character of discourse in Bentham's democratic theory that Gutmann (1980:26-7) casts Bentham as an opponent of participatory deliberative democratic theory, and as a traditional aggregative democrat. Crimmins acknowledges the negative role of public opinion as a check against abuse. In doing so, he has a wealth of textual evidence to call upon. In contrast, his effort to ascribe to Bentham a positive, quasi-Habermasian form of deliberation, which teaches citizens about the public interest, has little textual basis. This short part of Crimmins' work is interesting. It is a rare attempt to understand some important ideas in Bentham's democratic thought in reference to a contemporary debate about the merits of public reasoning. Its cursory nature, however, illustrates how significant the gap between those two fields can be.

Finally, there are many instances of recent scholarship which do not ascribe to Bentham a public-orientated view of deliberation in order to explain the importance of public reasoning before voting. Consider Lieberman (2014:119-142), who has provided a rich analysis of how Bentham's theory of public opinion interacted with his programme of legal reform. Like Schofield (2006) and Rosen (1983)—to whom we shall return in the next chapter—Lieberman approaches this question in terms of the development of Bentham's political theory. He primarily emphasises how Bentham's ideas about public opinion interact with his work to develop a comprehensive body of laws. Lieberman sees these as complementary strands in Bentham's thought. The proper operation of publicity would help to assure that the body of law Bentham sought to define would operate in public. Unlike the common law, that body of law would be understood by those subject to it. In the long-term, that publicity would enable citizens to have more secure expectations about how the law would affect their own welfare.

The common law was established through a series of individual precedents, rather than explicit principles. This had made it mysterious and unpredictable to ordinary citizens. In place of the clear, explicit and predictable principles that they needed to enable them to plan their future, citizens were instead provided with a series of implicit and unclear precedents and tacit assumptions. Each judicial decision that helped to create precedent was itself a result of how a particular judge's adjudication happened to progress. Critically, this undermined the security with which individuals could make plans to maximise their happiness. Of course, Bentham wanted to address this fundamental problem. As Lieberman emphasises, he was determined that laws should be more predictable and widely understood in the future than they had been in the past.

Moving beyond the opacity of common law was a central objective of the comprehensive body of law that Bentham tried to create. In place of murky precedent and unpredictable application, it was to guard against the abuse of power by “furnishing a structure of publicly articulated and maintained securities, which enabled individuals to chart their futures, undertake complex cooperative ventures, and realise their plans and expectations for happiness” (Lieberman, 2014:138).³¹ This project of legal reform was itself intimately related to the idea of publicity. The transition from common law to a new, comprehensive body of law was also a transition from secrecy to publicity. It would accompany a move from implicit assumptions to explicit principles; from remoteness to accessibility. In his inimitable style, Bentham summarised by saying that his objective was to marshal the law so as to “maximise cognoscibility” (Lieberman, 2014:139).

Lieberman's detailed and interesting work on how Bentham's thinking developed over the course of his career does much to explain some of his influences, and how they fit together in reference to publicity. It does not, however, aim to explain what Bentham's views on public reasoning were, or how these might contribute to contemporary normative debates on the role of discussion between citizens before they vote (Lieberman, 2014:119-42). Similarly, Schofield has developed a comprehensive and rich exposition of how publicity functions in Bentham's democratic theory (2006:250-71). He highlights the central place that Bentham gives publicity in assuring the accountability of public officials, and the many measures which were necessary to citizens having access to the information required for this. Reading this account, along with Lieberman's, in the context of contemporary debates about public reasoning and deliberation,

³¹ Clearly, Bentham did believe that this security of expectations was critically important. Indeed, separately, Kelly has argued that it was, for Bentham, the single most important factor in determining utility over the long term. Without the security of understanding how the law would interact with their own lives, individuals would inevitably lack an important and “necessary condition of personal continuity and of interest formulation” (Kelly, 1990: 73).

however, it is striking that publicity is portrayed almost as a feature of the institutions of the state; rather than as an account of citizens' behaviour and obligations. Following this trend, Schofield's comprehensive analysis outlines how Bentham believed publicity and responsibility should operate in a democracy, and the features of the architecture and processes of government necessary to making publicity an effective catalyst for making governors responsible to those they govern.

This account reflects some aspects of how Bentham believed the rationale for laws should be publicised, and how public opinion could go wrong when infected by delusions caused by fallacies. These are fundamental components of Bentham's views about how citizens should interact with one another and their government. Indeed, Bentham clearly links the delusions evinced by fallacies as an instrument of misrule with prodigious harm to the political process. By acquiescing to the unjustified claims of authority, he believed, passive citizens played their part in impeding progress. Whilst Bentham believed that a rational and informed debate would attribute approval and disapproval appropriately according to the pleasures and pains at stake, the manipulated and delusory discourse had prevented this from happening. There was thus a direct link between the delusions generated by a discourse replete with the fallacies that Bentham identified, and outcomes which served sinister rather than public ends: "Wherever the moral branch of the field of thought and action is on the carpet, they are continually occupied, with or without intending it, in drawing portraits (1989:266). Bentham believed that "sinister" ends, or interests, were served when the public interest was sacrificed in service of the particular interests of a group or individual.

Bentham also believed that the positive and negative implications of the language used in debate had been used to reinforce a false impression that members of the governing elite were benevolent and infallible. This failure of public discourse provided cover for sinister abuses because it made members of a governing elite immune to the kind of public scrutiny which was needed to make them dependent on the people they governed. It was this kind of unthinking reverence for putative authority that Bentham wished to demolish. He believed that public opinion could only operate successfully in a democracy once it had been liberated from these limitations. An appropriate process of reasoning between members of the public was part of the process through which that demolition of delusions and corruption could be promoted. In the long term, this would deprive sinister public officials of the opportunity to "reap the united fruits of selfishness and malevolence" (1989:266).

What is perhaps less clear in this account, however, is whether and how those views put forward by Bentham can be understood to provide an account of public reasoning, and how that account might interact with contemporary debates on this question. Contemporary deliberative theorists grapple with questions like: "what duties do citizens owe one another when deliberating on certain questions? What

kinds of reasons should they appeal to or ignore? How does deliberation change the character of its participants?”. There are many valuable contributions to Bentham scholarship which explain how publicity should function, but they tend not to directly address questions like these.

Conclusion

In Chapter 2, I argued that the conventional view can offer no convincing explanation of the value of discussion between citizens before voting. This is principally because its premises preclude many of the plausible grounds on which contemporary theorists conceptualise the value of this discourse. Participatory views do not fall into that trap. Where “conventional” readings of Bentham leave us with a view of democracy heavily reliant on aggregation, and relatively neglectful of public reasoning; participatory readings explicate ideas necessary to explaining the value of public reasoning. Both Niesen’s and Ben-Dor’s views, therefore, offer a fuller account of the value of effective discussion between citizens before voting than the conventional view, but encounter interpretive issues in doing so.

To summarise, the participatory views examined in this chapter offer a better response to the challenge of explaining the role of pre-voting discussion than “simple” or “conventional” readings considered above. Broadly, however, there are two reasons why they cannot provide grounds for a Benthamite account of public reasoning. Some of these accounts do not provide a plausible interpretation of Bentham’s views. Others do provide a plausible account, but do not relate it to the ideas which are relevant to contemporary debates. Let us recap these two matters in turn.

There are a few reasons to conclude that the accounts examined do not accurately reflect Bentham’s views. For instance, some scholars—like Crimmins—draw a direct equivalence between Bentham and Habermasian deliberation. That allusion could answer the normative challenge of explaining the value of pre-voting discussion. But it sits uneasily with Bentham’s legal and political writing. There are few reasons to believe Bentham intended pre-voting discussion to be transformative in the sense suggested by contemporary Habermasians. Other participatory readings would suggest that Bentham endorsed a public-spirited form of deliberation. That conclusion could, for instance, be drawn from Lyons’ (1973) dual-standard. This conclusion is mistaken, in my view. Bentham endorsed the secret ballot. His rationale for doing so was to protect individual voters from public influence, rather than exposing them to it. Further, Bentham portrays self-interest rather than the public interest as a predominant—if not comprehensive—motivation in individual behaviour.

The interpretive objections to other participatory views are more nuanced. Both Niesen and Ben-Dor consider Bentham’s later democratic writing in some detail. Niesen offers a strong justification for pre-voting discussion within a parliament. His excellent analysis of parliamentary discourse cannot, however, be extrapolated to explain Bentham’s views on *public* pre-voting discussion. Bentham believed that parliamentary and public forums were significantly different: consequently he stipulated different voting

rules for each. Second, the complex editorial history of “Political Tactics” limits how confident we can be in using that text to explain Bentham’s views; particularly where it seems to diverge from other texts.

Ben-Dor’s argument instead relies on the view that public discourse makes consensus possible by aligning incentives between citizens through the social sanction. Taken to its logical conclusion, this argument suggests that it could be possible for full consensus to be reached through discussion rather than the aggregation of votes. Although rich and fascinating, this interpretation must also be challenged. The early Bentham wrote extensively about how an enlightened legislator should use indirect legislation. That line of argument is less prominent, however, in Bentham’s later, democratic writings. Indeed, the text on which Ben-Dor principally relies—the first volume of the “Constitutional Code”—characterises the Public Opinion Tribunal as a security against misrule by governors; not an instrument for engineering agreement between citizens. In “Securities Against Misrule”, as elsewhere, Bentham makes the link between publicity and security against abuses by governors explicit:

“...the greater the number of the members of the whole community to whom the existence of an act of oppression has been made known, the greater is the number of those by whom, on the occasion of an endeavour to exercise other acts of a similar nature, supposing the past act notified to them, not only may obedience withholden, but resistance opposed”

(Bentham, 1990:30)

Bentham did not emphasise utilitarian consensus in the way Ben-Dor suggests. He did not use that term, or describe anything similar to the consensus described by contemporary theorists, including Rawlsians. Indeed, Bentham’s analysis of fallacies and delusion suggests a scepticism about the provenance of many of the (putative) agreements of his time.

The second reason why some views we have examined do not suggest an interpretation of Bentham on public reasoning is simply that they do not directly address the question. Scholars like Schofield and Lieberman offer an accurate and detailed analysis of Bentham’s views on the constitutional role of publicity. In doing so, however, they do not explicitly address the questions asked and answered by contemporary theorists of public reasoning. There are, however, reasons for optimism. Interpretations like these show that publicity, if not public reasoning, was central to Bentham’s democratic theory. It is therefore possible that the arguments presented by these scholars can be re-framed to address the contemporary debate. I attempt to begin that process in the next chapter. By using primary texts—along with insights from Schofield, Rosen and others—we shall investigate how far the views Bentham himself expressed provide grounds for an account of public reasoning which responds to the two challenges we have outlined.

Chapter 4. Bentham reconsidered

“It is well known that for Bentham the public interest is the sum of the interests of the members of the community, but it does not follow that the aggregate of whatever the individual members see as being in their interest at a particular time is in the public interest” (Rosen, 1983:29)

Introduction

Bentham’s rationale for democracy gives context for how he wanted public opinion to operate. Democracy was necessary because, without it, public officials consistently lacked the “aptitude” needed to serve the public interest. Creating that aptitude meant making them dependent on the people whose interests they should serve. Creating this dependency, in turn, required the use of clear language, rooted in individual interests. That required an empowered, critical, public opinion with the ability to exercise both social and political judgement (“sanctions”) against public officials. This is the context for public opinion in Bentham’s democracy. It must reflect individual interests and, in doing so, the public interest, to hold officials to account. It could therefore create “official aptitude” through a particular type of public engagement. I shall refer to this as “public aptitude”.

I have outlined and responded to a number of contemporary accounts of public reasoning, and some interpretations of Bentham on this question. Clearly, there are some normative “gaps” in that debate; particularly in the boundaries between views which have often been termed “deliberative” and “aggregative”. Some contemporary views provide a compelling explanation of why pre-voting discussion may be valuable to a democracy. In doing so, however, I believe they create greater opportunities for the outcomes of discussion to be manipulated. Indeed, it is alleged that some of the deliberative conditions defined by theorists are open to manipulation, because they imply a form of paternalism about citizens preferences; or the conditions under which they are formed (Posner, 2004). The strength of deliberative theory may also be its weakness, therefore. Other accounts avoid this danger. They offer fewer such opportunities for discursive outcomes to be manipulated because they utterly reject any paternalism, or even evaluation at all, over citizens’ preferences. Doing so, however, weakens their account of the value of pre-voting discussion.

In this chapter, I re-examine Bentham’s writing to determine the extent to which his work provides the basis of an account of public reasoning. In doing so it would, of course, be tempting to say that all of the gaps we have identified can be neatly filled by a proper understanding of Bentham’s actual views: that he anticipated and responded to all of the challenges posed by contemporary debates in this area.

Unfortunately, I do not think that a faithful reading of Bentham's work does answer every aspect of the challenges mentioned. Such a reading does, however, go far beyond some of the crude generalisations often made by contributors to that debate; particularly some of the deliberative theorists mentioned in Chapter 2. Bentham was not the blinkered subjectivist that Hart suggested, or the unthinking aggregator that Gutmann & Thompson imply. As the quotation which begins this chapter suggests, he did not believe that legislators should unthinkingly maximise *whatever* citizens happen to value at a particular time (Rosen, 1983:29). He consistently argued that the public interest should be understood in aggregative terms, but also placed public opinion—and an active and informed citizenry to generate it—at the centre of his democratic theory. In this sense, Bentham's views were both participatory and aggregative. We should not dismiss Bentham's potential contribution to the contemporary debate lightly: the simple premises of his views do not preclude a meaningful explanation of pre-voting discussion.

Of course, one major barrier is that Bentham wrote in a radically different context. Many contemporary theorists follow Rawls (1997) in seeking principles which are in some sense justifiable to a wide range of reasonable ethical perspectives. One would not expect Bentham to have done so. Rather, Bentham famously believed that he had defined—in the principle of utility—an evaluative standard against which all moral and political claims could be explained, understood, or repudiated (IPML, 12). Bentham's views are clear. It is important not to skew those arguments in an attempt to apply them to questions of contemporary interest. Indeed, it is partly an attempt to address contemporary constitutional debates that has led to interpretations which overplay similarities between Bentham and deliberative democrats. As we have seen, Ben-Dor (2000) brilliantly and originally applies Bentham's views to important current legal matters. On occasion, however, he does so by departing from what Bentham actually wrote. As one well-known delegate at a recent conference said of Ben-Dor's text, "it's beautiful, but it's not Bentham". As we saw in Chapter 3, others have made a similar argument. Crimmins (2011), for instance, suggests an affinity between Bentham and contemporary Habermasians, implying that both are committed to a small group of key concepts. Interpretations like these often overlook critical differences over questions like the orientation of debate.

The aim of this chapter is to outline the basis of a view of what Bentham did believe about public reasoning, and the importance of public opinion in a democracy. Clearly, there is no single text on this matter which covers every question of interest. Nonetheless, I believe that Bentham does provide many of the resources needed for a developed response: relevant principles become clear when we examine a number of themes in his writing together more closely. The result may not be "beautiful" in the sense described above, but my present aim is that it should "be Bentham".

To summarise, I believe that Bentham's logic for democracy makes official aptitude essential to good governance. In turn, the moral aspect of official aptitude requires a particular form of reasoning and engagement from citizens. I term this "public aptitude", and explore some of its features by drawing on a variety of different strands of Bentham's thought. Public aptitude has some important features which tell us about Bentham's views on public reasoning. It is critical of authority rather than acquiescent to it; it focuses on citizens' own self-interest rather than the public interest. It examines language to test its relation to pleasures and pains; and expects to reach decisions through majority voting rather than deliberative consensus. The practice of public reasoning both requires *and* cultivates the reasoning capacities needed to hold public officials to account.

This implies that public participation in pre-voting discussion is vital. But that discussion does not resemble the contemporary Habermasian ideal. Participation directed toward "public aptitude" is focused on clarifying individuals' particular interests rather than transforming them; it holds no intrinsic moral or democratic status independent of the benefits it confers on decision-making; and takes place with no expectation of full consensus between participants. Let us begin by taking a look at the context for Bentham's views on public reasoning: the roots of his belief in democracy. To understand Bentham's views on public reasoning, we must first revisit the underlying logic of the argument he made for democratic reforms in his later years, and how this relates to his notion of the public interest. Critically, this argument emphasises the importance of "official aptitude" which, itself, relies on a form of public participation. As I hope will become clear, Bentham's democratic views are therefore heavily reliant on a particular type of participation and reasoning from citizens: "public aptitude" is central to his account.

Bentham's democratic logic

It was not until 1817 that Bentham published an unequivocal endorsement of the case for radical democratic reform. Chapter 5 of Bentham's "Parliamentary Reform Catechism" was published in the text of the "Plan of Parliamentary Reform". It summarises the argument succinctly:

"The immediate cause of all mischief of mis-rule is, that the men acting as representatives of the people have a private and sinister interest, and sufficient power to satisfy that interest, producing a constant sacrifice of the interest of the people...the cause of this immediate cause is this, that these same agents are in one case unduly independent...of their principals—the people; and dependent upon the Corruptor General" (Bentham, 1818a:24)

The position, power, and incentives of officials caused damage to the public interest. Officeholders had historically depended on the King (“Corruptor General”) for their office and the money and status that came along with this, but depended on the general population for nothing. The rewards offered were corruptive: they influenced the will of officials in a way which damaged the public interest (Bowring viii, 545-6 in Schofield, 2006: 142). Radical reform was needed to reverse this situation; to make representatives dependent on the people, and independent of the monarch. Along with delusion, which operated on the intellect (Ibid), corruption was a principal threat to the happiness of the community. It was promoted, not negated, by the secrecy of the parliamentary system.

In part, this conclusion was the product of Bentham’s definition of the public interest. As we have seen, he held that the public interest was the legitimate end of government (IPML:12), and was part of what distinguished political action from private conduct in ethics:

“Of this constitution, the all-comprehensive object, or end in view, is, from first to last, the greatest happiness of the greatest number; namely, of the individuals, of whom, the political community, or state, of which it is the constitution, is composed; strict regard”.
(Bentham, 1983b:18)

Of course, the wording that Bentham used to describe the end of government changed slightly over the course of his career. He began by referring to the “principle of utility” in “Deontology”, and then clarified that this entailed achieving the “greatest happiness”; a phrase which was then expanded into “the greatest happiness of the greatest number” (Schofield, 2006:38). Bentham rarely referred to the public interest without also tying that idea back to its roots: the pains and pleasures of individuals. The care that he took to emphasise that individualistic definition of the public interest was significant. He believed that a failure to attend to the exact definition of the interests at stake in political debate caused immense damage. In the “Book of Fallacies”, for example, he insists that the pleasures and pains attributed to any proposal must be assignable in order to produce useful debate. In his essay on the “Principles of International Law” he reiterates that to attribute a character to nations would be harmful; for analogous reasons. Here, again, referring to general ideas rather than identifiable causes of pleasures and pains clouded debate rather than clarifying it:

“A nation has its property—its honour—and even its condition. It may be attacked in all these particulars, without the individuals who compose it being affected. Will it be said that it has its person? Let us guard against the employment of figures in matter of jurisprudence. Lawyers will

borrow them, and turn them into fictions, amidst which all light and common sense will disappear; then mists will rise, amidst the darkness of which they will reap a harvest of false and pernicious consequences.”

(Bentham, *Principles of International Law*, Bowring II, p.544)

This emphasis on assignable references to identifiable interests is also evident in Bentham’s constitutional analysis. He denounced the authors of the French Constitution for using “terrorist language” (2002: 330). They were guilty of using abstract propositions that diminished the utility of discourse, and made good decision-making less likely. Indeed, Bentham suggested a direct link between how abstract the language used was, and how liable it was to mislead. He suggests that it is precisely because the language used made imprecise reference to the public interest, and related concepts, that it was prone to misleading citizens. That claim was generalised, to suggest that “the more abstract—that is the more extensive, the proposition is, the more liable to involve a fallacy.” (Bentham, 2002:320). In place of a clear explanation of the relevance of prescribed measures to the self-interest of individual citizens; broad and ambiguous reference was made to the welfare of the nation. Like vague references to the “character” of the nation, claims like these were “beside the purpose” (2002: 321) of real political discourse. They obscured rather than clarifying citizens’ view of their own interests.

Through operation of the common law, Bentham also believed that references like these had undermined the law itself. As we saw in Chapter 3, the remedy was clear justification, explained in public according to the assignable interests concerned. In the “Codification Proposal”, for example, Bentham insisted that laws must be accompanied by a rationale, which constantly connects them to their rightful end; in reference to the real entities of pleasure and pain (Ibid: 245-60). In the absence of this kind of justification, he thought that the practical application of the law had become arbitrary. Without justification, the law could become a product of the way individuals felt about particular questions: sympathy, antipathy, or caprice. Those feelings were often shaped by their particular interests and prejudices. The abuse of language was an important part of the sinister sacrifice of the public interest which Bentham wished to prevent. It maintained a deluded, and corrupted public; without which those abuses could not continue. The logic behind Bentham’s belief in reform was thus that democracy was needed to prevent this sinister cycle. It introduced an essential measure of public scrutiny and justification. Bentham believed that “sinister” ends, or interests, were served when the public interest was sacrificed in service of the particular interests of a group or individual.

It is worth exploring the role of language in allowing sinister abuse a little more. In the “Book of Fallacies”,³² for example, Bentham identified some of the abuses of language and logic that he thought were to blame. There, the “fallacies of confusion”—explained in Chapter 4—defines the language that Bentham believed had been used by a ruling minority to conflate their own interests with the professed ends of public institutions. Thus, an indiscriminating use of general terms like “government” was observed. This permitted the service of sinister interests above the public interest, since: “This impropriety of speech being thus happily and successfully established, the fruits of it are gathered in every day” (Ibid, 260).

Bentham’s views on political language are important. They provide a significant reason to believe that he was sceptical of public-interest-based public reasoning, under many conditions. He was suspicious of the potential for abuse of many abstract terms, including “government” itself. If, as seems likely, an account of public reasoning which aims to help voters learn about the public interest involves more extensive use of language which refers directly to the public interest—rather than its component parts—then, on Bentham’s logic, this would be more likely to deceive voters. In the first chapter of IPML (paragraph V), for example, he insists that “It is in vain to talk of the interest of the community, without understanding what is the interest of the individual. A thing is said to promote the interest, or to be for the interest, of an individual, when it tends to add to the sum total of his pleasures: or, what comes to the same thing, to diminish the sum total of his pains.”

This idea permeates much of his career: from his constitutional writing, to his efforts to achieve practical political and social reforms. In August 1793, for instance, Bentham wrote to Earl Spencer in support of his ill-fated plan for a Panopticon prison. Here, he again reiterates the importance of understanding the public interest in terms of the assignable interests of individuals, rather than more general terms:

“No Man...can be more sensible than I am to the abuses to which the maxim that private interest should give way to public, is liable, and is but too frequently made subservient: as if the public were made up of anything but individuals”
(Bentham, 1981: 472)

This individualistic, aggregative definition of the public interest remained relatively stable through Bentham’s career. So too did his emphasis on the importance of precise language; which picks out individuals rather than abstract entities. The wording that Bentham used, however, varied significantly

³² <http://books.google.co.uk/books?oe=UTF-8&id=KuEqAAAAMAAJ&q=abstract#v=snippet&q=abstract&f=false>

(Gunn 1968:4). A consistent theme is that Bentham is sceptical about the usefulness of general references to “the public”. Through his works, he refuses to give the public or community a particular nature or character; independent of those of the people who composed it.

From this individualistic approach, and the aggregative definition which followed from it, we might expect Bentham to argue that a simple aggregation of self-interested votes alone reflects the public interest in political decision-making. The reality of Bentham’s views was more complex. He thought that the public interest would be served in a community when the agents making decisions had the ability, information, and integrity needed to perform that role effectively. Bentham termed these properties “aptitude”. As Rosen (1983) and Schofield (2006: 272-303) have established, the aptitude of officials became an important feature of Bentham’s later writing, particularly from around 1809. The ideas it signified also featured in Bentham’s earlier work. For instance, the first section of the first chapter of the “Handbook of Fallacies” provides an analysis of claims to authority. Here, Bentham’s suggests that putative authority should be understood in relation to the qualities of the agents involved. The reliability of a putative authority on any particular question could be ascertained by understanding their intelligence, probity, proximity to the matter under discussion, and how faithfully their opinion had been communicated. Relative and adequate intelligence was required to use and process information reliably. Being distant from the matter in question lent itself to deficiencies of intelligence because it meant that the information used could be less reliable. Probity, in contrast, was a moral rather than epistemic quality. It reflected the ability of a person or group to pursue the principle of utility rather than interests which were particular merely to themselves or specific groups of people.

Bentham believed that the most common deficiency of probity was simple insincerity: discrepancy between an opinion expressed and that genuinely held. Authority could also be undermined in deeper ways. Where a person’s interests were served by reaching one conclusion over another, he may pay greater attention to facts supporting one side than another. The will to favour one side of a debate was a failure of probity; the misjudgement it created a deficiency of intellect. These deficiencies meant that those who could control public decisions, and stood to gain from them, were likely to serve themselves rather than the community; unless given incentives to do otherwise. As we shall see later in this chapter, this idea of (“sinister”) persons or groups being able to sacrifice the interests of the community was central to Bentham’s view of democracy. I believe that it is also critical to understanding his views on public reasoning.

Aptitude and utility were intimately, perhaps even proportionally, connected. Bentham's later constitutional writing, for example "Economy as Applied to Office", suggests that the value of a government was determined by the usefulness of its laws. In turn, this depended on the aptitude of the officials who created them:

"The goodness of the government will be as the aptitude of the portions of law enacted by it and the operations performed by it: their aptitude with relation to that same end...[and] the aptitude of the operations will be as the aptitude of the operators "

(Bentham, *Economy as Applied to Office* in 1989:4)

Official aptitude was therefore critical. Bentham developed a fuller picture of how it should be defined and cultivated in some of his later writing. For instance, he paid great attention to the characteristics needed to make good decisions, and how they could be identified through measures like public examinations. As Rosen has noted, this reflected Bentham's intention that the decisions made by officials must be more than just an unthinking reflection of the will of the majority (Rosen, 2003:196). Rather, competent officials would consider the longer term interests of the whole of the community; including the security of individuals which Bentham prized. Ideas of probity and talent in Bentham's early writing thus gave way to a more developed later account of "official aptitude". Like earlier work on probity, talent and intelligence, this reflected a range of different attributes required to maximise public utility. As mentioned briefly above, official aptitude had three aspects: moral aptitude, intellectual aptitude, and active aptitude.

Moral aptitude (which was similar to earlier notions of probity) was a negative quality: the absence of a motivation to act in one's own interest at the expense of all other interests. Interestingly, moral aptitude was not an attribute of rulers' personality or character. Bentham's view was not that we should expect, or hope for, self-sacrificial valour from officials. Rather, moral probity was a function of the situation that public officials found themselves in; the environment and incentives to which they were exposed. Indeed, Bentham believed that it was delusory and dangerous to expect good conduct from officials who were offered the opportunity to serve their own interests above the public interest (Bentham, 1989:266). Assuming their public-spiritedness in this way reduced citizens' capacity to be aware of relevant interests; and to reason critically about political decisions.

Bentham's mission was thus to create arrangements whereby rulers' natural inclination to serve their own interests was used to maximise—not sacrifice—the happiness of the community:

“...to place them in such a situation as that while they have no prospect of being able to encrease each of them his own happiness by diminishing the happiness of the greatest number, they have each of them a prospect of giving encrease to his own happiness by giving encrease to the happiness of the greatest number”

(Bentham, *Economy as Applied to Office*, in *First Principles*...1989:14)

At the heart of Bentham’s constitutional theory was the assumption that self-interest was a predominant motivation for individuals. Securities were therefore necessary to ensure that this motivation served rather than undermined the public interest.

Functionaries also needed the ability to understand what would further the public interest; as well as the inclination to do so. Intellectual aptitude was thus essential, and was divided into two sub-categories. These were the scientific—necessary for an official to perform his public duty—and judicial, which gave him capacity to make correct decisions. (Schofield, 2006: 274). Finally, Bentham believed that officials needed active aptitude: the physical activity of performing their public duty, rather than simply understanding it or being inclined to act. Like doctors, whose “uninterrupted endeavour to afford relief” (Bentham, 1989:87) was fulfilled through an obligation of attendance, public officials should also be given minimal scope to fail to attend their office. Of these three branches of official aptitude, moral aptitude is most interesting in reference to public reasoning. Bentham famously proposed that an artificial identification of interests between the governing and the governed was necessary to maximise public utility (Harrison, 1983:232). The particular need for this identification of interests arose from what Bentham believed was the structure of interests in society, as well as an assumption that of predominant egoism. As we shall see, this became central to his thinking about democracy.

Of course, Bentham was not always a democrat. Much of his earlier writing was directed at legislators rather than democratic voters. It sought to furnish them with qualities needed to make good decisions in the expectation that a sufficiently wise and well-informed elite could systematically serve the public interest. His notions of probity and talent were not, therefore, intrinsically democratic. Rather, they were technocratic; Bentham wanted to analyse and improve the quality of decision-making on a practical level. It was in this efficiency of office that he believed that greatest happiness of the greatest number would be realised. There was, however, a significant change in Bentham’s thought over the course of his career.

From 1808-9, Bentham's commitment to fundamental constitutional reform was in place, but he did not publicly argue for near-universal suffrage until the 1817 "Plan of Parliamentary Reform". The logic for the later Bentham's democratic views did not hinge on an unconditional right of citizens to participate under all circumstances. Instead, he came to the conclusion that rulers in the real world did not have requisite moral and intellectual aptitude to serve the ends of the community. Misuse of power was the central issue. Unless subject to the incentives created by a democratic system of government, governors would fail to maximise public utility. To serve the people, they had to be made dependent on the people.

Thus, while the early Bentham placed some trust in the possibility of an enlightened legislator, the later Bentham became deeply distrustful of the ruling classes. He suggested that to believe that the ruling classes wanted to serve the public interest would be like supposing "that it is for the happiness of negroes that planters have all along been flogging negroes" (Bentham, 1993:254). That colourful quotation captures two things. First, it reflects the transition from trust to distrust of elites, which accompanied Bentham's transition to democracy. Second, it reflects the importance of class in Bentham's democratic thought. He came to believe that interests were starkly divided in society: between opposing classes. That opposition was critical to the need for democratic accountability. The exact timing of Bentham's conversion to democracy has been the source of much speculation (Mack, 1962: 409; Halevy, 1952: 147; Dinwiddy, 1975:683-700). In recent years, however, our understanding of how Bentham's political thought progressed has been advanced substantially. Schofield's authoritative analysis in "Utility and Democracy" (2006) suggests that there was no single point at which Bentham suddenly became a democrat. Rather, he developed constitutional theory that he started to outline in the 1820s, and the calls for radical parliamentary reform published in 1817 were the product of reflection on views that Bentham had held for decades. Using a variety of unpublished texts, Schofield (2006:109) has argued that Bentham's conversion to a form of democracy was intimately related to the idea of "sinister interest", which first developed rigorously in or around 1804. He describes sinister interest as bringing about "The critical development of Bentham's political thought...from which he eventually drew the conclusion that democracy was essential to good government" (Schofield, 2006:109).

Whereas moral aptitude was achieved where governors' particular interests were aligned with the public interest, sinister interest was served where these two diverged and self (or group) interest was served at the expense of the community. The public official was peculiarly exposed to the potential for sinister abuses of power. Bentham explained why in Chapter 3 of the "Handbook of Fallacies":

"The mind of every public man is subject at all times to the operation of two distinct interests—a public and a private one. His public interest is that which is constituted of the share he has in the

happiness and well-being of the whole community, or of the major part of it: his private interest is constituted of, or by, the share he has in the well-being of some portion of the community less than the major part; of which private interest, the smallest possible portion is that which is composed of his own individual—his own personal—interest”

(Bentham, Ch.3, Book of Fallacies).

As this quotation—and the later idea of moral aptitude—imply, the potential for sinister abuse was largely a feature of the situation that public officials, rather than a product of individuals’ failures of character. Bentham says that, when left without recourse from others, individuals will almost always prefer their particular interests to those of the community. Indeed, this was something in which “there should be found no regret”. Rather, it was a simple feature of human behaviour. In the majority of cases, he believed, there would be a divergence between officials’ particular and general interests. This caused government to be used to serve the (sinister) interests of an elite, and for the public interest to be sacrificed to that end. As we have seen, he termed that sacrifice “sinister”. Bentham also saw sinister interest at work in the legal system. Here, too, the “sacrifices of public welfare to private convenience” were commonplace (Bentham, 1821a:80). There is a striking similarity between Bentham’s diagnosis of how jurors were manipulated, through pecuniary incentives and deception, and his analysis of rulers’ conduct.

In both cases, he assumed that we cannot expect individuals to serve the community by consistently diverging from their own particular interests. To do so would be to neglect a “fundamental axiom”: particular interests shape behaviour and judgement. As a consequence, controlling a man’s prospects of future welfare conferred the ability to influence his actions: material dependence shaped judgment and action. Thus, possess “in your hands...the faculty of influencing that interest or welfare...[and] the exercise of... [their] power is...at your command” (Bentham, 1821 a:16). This principle caused both public officials and jurors to fail in their roles. Both were placed in a situation where particular and public interests diverged: both pursued the former at the expense of the latter.

This “fundamental axiom” of behaviour was not a significant problem for most individuals. Bentham believed that this group, which he termed the “ruled”, or “unopulent” at different points in his writing, had particular interests which largely coincided with the public interest. It was, however, a colossal problem for another class. Bentham held that there was much in common between the majority of the population, who were ruled. There was, however, very little in common between the ruled and the ruling. In different

texts he uses slightly different language to describe what we might now call a “ruling class” or elite. They were sometimes “opulent”; on other occasions they were described as “great characters in high situations” (1821a:263). In “First Principles...” two classes in society are characterised as “productive” and “unproductive” (1989: 71-2). The former referred to the working majority and the latter to their political rulers. In this text, the importance of these groups to Bentham’s understanding of how the public should reason together in a democracy begins to become more obvious: he begins to relate class to the operation of public opinion more explicitly.

Bentham argues that the Public Opinion Tribunal (POT)—the fictitious entity which represents and operationalises public opinion in a democracy—is composed of two distinct segments: the Democratical, and the Aristocratical sections. The former constituted the majority of the Tribunal, and the latter the minority. Since the community was the sum of the individuals who composed it, this difference in numbers mattered. The opinion of the Democratical Section was always more likely to be “in the highest degree contributory to the greatest happiness of the greatest number and therefore in the highest degree accordant with the universal interest” because it comprised the majority of individuals. Critically, however, Bentham caveats this claim. He says that the opinion of the majority will promote public utility only “in so far as the conception entertained by the several Members in relation to their respective interests is correct” (Bentham, 1989: 68-9).

Two aspects of this caveat are key to our current purposes. First, Bentham acknowledged the capacity for error in how interests were identified. This provides additional evidence that we should reject any simplistic “naive subjectivist” reading of public opinion in his political theory. It therefore also provides grounds to reject that aspect of the “simple” or “conventional” view described in Chapter 2. That important point is confirmed by Rosen’s analysis of the “Constitutional Code”: Bentham did not believe that the public interest was whatever public opinion happened to dictate at a particular time (Rosen, 1983:29). Rather, its efficacy depended on the quality of the judgements made by individuals. As we will see, that quality was enhanced by a particular form of democratic institutions, and a particular type of reasoning by the people who composed the public.

The second thing about this quotation that we should note is that Bentham wants members of the Democratical Section to have an accurate conception of their *particular* interests; not of the universal (or public) interest as a whole. The simple fact was that they—as the “unopulent”—composed the majority of the community. As a result, individuals within this majority could reflect on, and understand, their

particular interests. By doing so, they would, by weight of numbers, also reflect the public interest. This followed since: the unopulent had largely similar interests in light of their similar circumstances (relative to those of their rulers); and the interest of the public was nothing more than the sum of the interests of the individuals who composed it. This meant that the interests of the majority could not meaningfully diverge from those of the public. In contrast, the interests of a minority—particularly one with different relative circumstances—could readily diverge from those of the public. Consequently, for the majority at least, effective public opinion requires a genuine reflection of particular interests; not deliberative transformation away from them. This is perhaps to be expected. After all, we have seen that Bentham found no cause for regret in individuals' preference for their particular interests over the universal interest.

In contrast, individuals who composed the Aristocratical Section were a mere minority of the whole community. They came from “the ruling and otherwise influential few: of the highest ranks of functionaries in the state, with the addition of such other classes whose particular interests are in a league with theirs” (Bentham, 1989:69). Bentham assumed that they, like anyone, would form opinions and behaviours which reflected their particular interests. Their actions will “...therefore be in each instance determined by the interest common to the Members of this Section”. The great danger was that the particular interests of the Aristocratical Section radically diverged from those of the rest of society:

“...in a great, not to say the greatest, part of the field of morals including that of legislation the interest common to the Members of this narrow Section is in direct opposition to the interest of the more comprehensive Section” (Bentham, 1989:68-9)

Bentham thus asserted a “diametrical opposition” between the interests of these two classes (1989:70). The interests of the Democratical Section generally aligned with those of the community as a whole, by weight of numbers. In contrast, those of the Aristocratical Section usually ran against those of the community: they were a minority, in unusual circumstances.

In order to serve the public interest, therefore, effective public opinion required an accurate conception of respective interests from the (Democratical) vast majority of the community and far less input from the (Aristocratical) ruling elite. For this reason, Bentham suggests that Aristocratics may contribute some intellectual aptitude to the operation of public opinion (as more formally characterised by the later Public Opinion Tribunal). They should, however, have no significant say over its judgements. Aristocratics thus remained part of the wider public; but they were a small part whose interests were likely to be outweighed. Thus, when Bentham considers the function of the Public Opinion Tribunal in his

“Constitutional Code”, he implies that Aristocrats could contribute intellectual aptitude to debates, but were—by definition—unlikely to contribute moral aptitude: their interests diverged too radically from the public interest. This principle was reflected in the constitution of sub-committees of the Tribunal.

When these sub-committees were formed, Bentham prescribed that they should include only a minimal, minority representation for the Aristocratical Section. This would ensure that the community could benefit from whatever knowledge and judgement that group could bring to bear. Minimal representation would also ensure that Aristocrats could not affect or direct the POT’s overall decisions (1989:74). This measure thus reflected the importance of both intellectual and moral aptitude to good decision-making. Bentham’s aim was to ensure that the findings of the POT served the public interest as reliably as possible, despite a pronounced divergence of interests between classes of citizens.

The fundamental opposition of interests that Bentham identified between classes thus meant that there was not so much a single “public”. Rather, there was something closer to a majority “public”, and an elite. Different discursive rules applied to each group, to reflect the fundamentally different interests that they held. In describing the role of different groups in the operation of public opinion, the later Bentham never lost sight of his view that members of the ruling elite may only be expected to prefer their own particular interests when these diverged from those of the community. In the “Constitutional Code”, for instance, Bentham expected each member of the Aristocratical Section to behave in ways which served their particular interests; to “give his support to all things the tendency of which is to give augmentation to that quantity—to all laws, institutions, customs, opinions, prejudices, desires, hopes, fears, in so far as the tendency of them is to give existence to that effect” (Bentham, 1989: 73).

These two classes marked out the background of those who constituted the Public Opinion Tribunal (POT), but they also influenced how they should contribute to its effective operation. It would have been very strange if Bentham had expected the (Democratical) majority of the public to transform their preferences through participation in the Tribunal. After all, he believed that their particular interests already coincided with the general interest. Society thus required from them an accurate conception and expression of their own interests, rather than a public-spirited transformation away from them. Of course, the same could not be said for the (Aristocratical) ruling few. As a consequence, different rules promoted the public interest in respect of these two different classes. Where Bentham wished to adduce a genuine reflection of particular interests, he accorded secrecy. When he wished to push particular interests toward greater conformity with the general interest, he advocated publicity. Voting in secret was useful in the

former case; public voting in the latter. When thinking about the rules of discourse and reasoning that Bentham thought appropriate to the Democratical Section, his view on the secret ballot is particularly instructive. That argument reaffirms just how far Bentham's views were from what those endorsed by many Twentieth Century "stronger" deliberative theorists. He suggested that the POT should address "public" matters rather than the private conduct of individuals (1983b:37).

As we have seen, Bentham's advocacy of secret voting is well known. Participants in contemporary debates do not, however, always appreciate just how insistent, consistent, and important his arguments on this matter were. Through the 1790s and 1800s, his views on many important questions changed. In arguing for parliamentary reform, for example, Bentham endorsed different configurations of qualifications for the right to vote, and a significantly different scope for the franchise. His belief in the value of reform itself was radically affected by developments in France following that nation's revolution (Schofield, 2006:78-109). It is well known, too, that Bentham's beliefs on the rights of women (Ball, 1993), the desirability of democracy (Mack, 1962), and many other matters changed over time. It is therefore striking that Bentham—as far as we can see—never advocated public voting. This comes despite the fact that it was a popular view at the time (Park, 1931). The rationale for the secret ballot is, therefore, absolutely central to understanding Bentham's view of democracy. It is essential to his conception of the value of voting that "public happiness...is in this instance better promoted in every point of view by keeping the votes secret than by rendering them public." (Bentham, 2002:430)

Bentham's reasons for objecting to public voting echoed his view of class, and of participation in the POT. The standard objection to secrecy at the time was that:

"The practice of voting by ballot has an immoral tendency: it affords a screen to cowardice. A free man and Englishman ought to declare his choice boldly, without respect of persons, or fear of the consequences to himself" (Bentham, 2002:429).

He objected to the idea that the practice of voting should be designed to assume that voters would act as they should, by voting for the public interest, rather than how they could act. The belief that individuals should vote on the basis of the public interest—irrespective of the consequences to themselves—thus supposedly leads to the conclusion that voting in public is justified; to allow individuals to discharge this 'duty'. Bentham's rejection of that view is clear. He insists that procedures like voting should be designed around the likely behaviour, not optimistic hopes around the way people might ideally behave:

“In our dealings with mankind we ought to square our measures not by what they ought to be but by what they are. Am I to leave my house-door perpetually open, because all men ought to be honest... because no man ought to enter my house and steal? Merely because such is a man’s duty, ought I to place a man in a situation in which it is certain that he will not do his duty?” (Bentham, 2002:430)

Thus, even if voting to reflect the public interest were desirable, Bentham did not believe that this would justify a public ballot. This would present citizens with unnecessary temptation to neglect their duty. Instead, he believed that we should design constitutions round realistic expectations of individual behaviour; to mitigate the harm that abuses could cause.

Of course, “realistic expectations” of citizens can be understood in a variety of ways. Bentham’s own view is clear. As he famously argues at the start of IPML (Ch1, XI), assuming behaviour on the basis of self-interest is a more reliable basis for designing institutions and processes than alternatives. Indeed, Bentham was equally clear that the individual citizen had a duty to serve the public interest. Before voting, he gained the status of “being a trustee for his Co-electors as well as for himself”. This entailed that he should “act in conformity to the public interest in preference to any private interest of his own”. Bentham believed that this duty was often useful in many circumstances. It was, however “perfectly useless in all cases where either it is his own private interest only that the case requires he should pursue, or whereby acting in pursuit of his own private interest he acts as much in conformity to the public as if the public interest was the direct object of his pursuit” (2002: 431).

Voting was exactly such a case. By acting in pursuit of his own interests, the voter advanced the public interest; as much as he would have done by directly pursuing the public interest itself. For Bentham, the connection between self-interested voting and public-interested outcomes was no accident. It was a necessary consequence of the fact that (IPML:1), the interest of the community is nothing more than the interests of the individuals who compose it. As a consequence:

“... It is *impossible* the majority of grown persons in a nation can have each of them an interest distinct from that of the whole. They have not naturally of their own any such separate interest, and in whose power should it be to give it them?” (Bentham, 2002: 431, italics added)

Bentham’s views on this question were clear: the necessary connection between self-interested voting and public interested outcomes made public interested voting unnecessary:

“While he acts [votes] in secret he is at free liberty to do what he thinks right and for his interest (which in this case are but two phrases for the same thing)” (Ibid).

Further, he argued, encouraging voting in public—in an attempt to promote public-minded participation—would be dangerous. It would expose voters to influences that work against the public interest:

“Expose him to the public eye, what will be [the] consequence? You expose him at the same time to the eyes of some individual or individuals who have an interest of *their own opposite to the interest of the public* and who have it in their power to make you suffer if you do not give preference to theirs, or what comes to the same thing, an opinion or caprice of theirs” (Bentham, 2002:431, italics added).

Public voting thus made the mass of ordinary—“unopulent”—voters, who by dint of their similarity in circumstance and scale of numbers, had much in common, vulnerable to the influence of others who could make them vote in ways which do not reflect their own interests. In the terms later put forward in the “Constitutional Code”, publicity at the ballot box would make the Democratical Section subservient to the Aristocratical Section. It would enable the rich and powerful to manipulate the votes of those who depend on them for a living.

As we have seen, the threat was similar to that Bentham had diagnosed in juries. Jurors could be corrupted if placed in a situation where serving their particular interests meant neglecting the public interest. So voters could also fail to serve the public interest when given incentives to do otherwise. There was, however, a crucial difference between voters and jurors. Voters could serve the public interest by considering their own interests, but jurors needed to consider the merits of a case impartially to do so. The key, in both cases, was to ensure that incentives offered to the individual charged with making a decision coincided with measures that serve the public interest, rather than diverging from it. In each situation, Bentham wished to design procedures which worked with the natural self-interest of individuals; producing better outcomes by taking account of their circumstances.

Effective voting, for Bentham, required the millions of people who shared common, or ordinary, circumstances to vote according to their own interests. Aggregation would then recognise what they had in common, and produce results which serve the public interest. Barriers to effective voting—like the public ballot—placed obstacles between the self-interest of “ordinary men” (1821a:221) and the ways in which they could vote. These barriers could be understood in terms of the principles that affected individual conduct:

“Three principles, the selfish, the dissocial, and the social, share the dominion of man’s conduct. Where the first two are out of the way, the latter will carry everything before it. No matter how weak it is, it will shape completely the course of every man’s conduct, while it acts alone. What

should hinder a man from voting for the candidate whom he deems the fittest, when there is no advantage to be got by voting for any other? What should induce him to vote for any other candidate than him whom he really deems the fittest, when there is no advantage to be got by it?" (Bentham, 2002:431).

The central task ahead of those designing democratic institutions and procedures was thus to mitigate the influence of selfish and dissocial influences on actors' conduct. Doing so would enable actions, and outcomes, which serve the public interest.

Specifically, those who depended on landowners for their living would feel pressure to vote in accordance with those "sinister" interests rather than their own; for fear of the consequences of doing otherwise. Public voting thus pressurised the ordinary voter to neglect his own interests at the ballot box for fear of retribution. This meant that those with financial status were likely to benefit from public voting: not the majority or, consequently, the community. As a result, "the only security for genuineness of suffrage" was secret voting (Bowring, iii, 599). Note that unlike "stronger" deliberative democrats of the 20th Century, Bentham wanted votes to be *genuine* rather than the product of transformation of preferences.

He believed that effective voting reflected voters' particular interests rather than being the product of a type of deliberation which transformed underlying preferences beyond self-interest. The vital question was therefore whether individual voters had a correct conception of their particular interests, and whether they were put in a position (secrecy) to genuinely reflect them; not how public-interested they could become before voting. This account of secret voting thus complements Bentham's later writing on the Public Opinion Tribunal (1989: 68-9). In both cases: the individual's actions were directed by "his own interest meaning according to his own conception of it, to the conception correct or incorrect entertained... at the moment of action" (Bentham, 1989:68).

The idea of "correct" and "incorrect" conceptions of interests runs through many strands of Bentham's work. It underscored his rejection of any claims to infallibility in his political and religious writing, and was grounded in his epistemology. Bentham believed that the indirect or inferential relation between physical entities and our understanding of their existence was an inevitable source of error. It relied on "judgement, ratiocination, which is liable to be erroneous, and in experience is very frequently found to be so" (Bowring, viii 224). Fundamentally, therefore, since inference was required to gain an awareness of even the simplest external event, and inference was prone to error, so even these simple claims had to be treated with care. Their relation to the real (external) world had to be questioned and verified: "Simple perception is not capable of erring, no, nor sensation neither. But judgement is, on the part of every person, and on almost every occasion exposed to error" (Bowring, viii, 320). The number of "infallible"

perceptions that could take place was, therefore, extremely limited. The only simple perceptions in which error could not be found were those regarding the direct perception of one's own experience; including one's own utility: the direct sensation of pleasure or pain.

In contrast, aspects of language that supposedly refer to "real" aspects of the external world had to be criticised by those using language in order to test their relation to the real world, and hence—given the assumption regarding true beliefs' utility above—their usefulness. On questioning the inferential process by which a supposedly "real" entity referred to an external object or event, if one was unconvinced of the inference's validity, the noun-substantive "in question [would]... not represent a real entity, but a non-entity" (Schofield, 2006:17). On these grounds, error was to be found in many places. One of Bentham's reasons for railing against Church of England doctrine was that its authors "as *men*...are but fallible" (1818b:36). In part this was inevitable. As Bentham asserted in the "Table of the Springs of Action", "no human act ever had been or ever can be *disinterested*" (1983c: 99), and "judgement—opinion—is liable to be acted upon, influenced, and perverted, by interest" (1983c:111). Bentham generalised that claim in the "Rationale of Judicial Evidence":

"I have no great opinion of infallibility; and if it were necessary to believe in it, I would go to work by degrees, and begin with the pope"

(Rationale of Judicial Evidence in Bowring, vi,3, 5, quoted in Schwartzberg, 2007:564)

This idea of fallibility of judgement was therefore intimately related to whether individuals held "correct" or "incorrect" conceptions of various interests; including their own. It also produces a fascinating tension between Bentham's epistemology, and his writing on public opinion. On the one hand, he suggested that all judgement is inevitably fallible. Consequently, he rejected the idea that any individual could be infallible. On the other hand, he had great optimism about public opinion in general; particularly in the "Constitutional Code". In places, that optimism seemingly overwhelms the potential for error identified in his epistemology. For instance, Bentham asserts that the variance between the dictates of public opinion and those of the Greatest Happiness Principle will become less and less numerous over time until they are in complete conformity with one another:

"Even at the present stage in the career of its civilisation, [public opinion's] dictates coincide, on most points, with those of the greatest happiness principle; on some, however, it still deviates from them: but, as its deviations have all along been less and less numerous, and less wide, sooner or later they will cease to be discernible; aberration will vanish, coincidence will be complete". (Bentham, 1983b:36)

Bentham is thus acutely aware of the capacity for error in individuals' judgements, but also extremely optimistic about the public opinion they created, in aggregate. He is at once wary of the fallibility of public opinion, yet reliant on its apparent infallibility.

On this question, as on others, there is no reason to assume that all of the views Bentham expressed over the course of his career were necessarily consistent with one another. After all, Bentham changed his mind about a number of important matters. These included the extent of the franchise, and even the necessity of democracy itself. That said, this particular tension is perhaps less substantial than it first appears. Consider Bentham's suggestion that public opinion is infallible in the "Constitutional Code". It would be a clear contradiction if Bentham had said that a single person, like the monarch, generated public opinion, and was incapable of error. As we have seen, his epistemology implies that this would be impossible in reference to judgments about the external world, if not sensory experience. It was for precisely this reason that he mocked authors of the French Constitution in "Necessity of an Omnipotent Legislature" for their tacit assumption that no amendments to their work would be required: "Blessed fruit of a wisdom which proclaims itself infallible! Blessed state of the government of a great empire" (Bentham, 2002:270).

Of course, public opinion was not, for Bentham, like this. It was not generated by—or ascribable to—a single person. Rather, public opinion was what individuals believed the whole of their community to think. Bentham also believed, as we have seen, that it was impossible for the majority of individuals in a community to have an interest which was opposed to that of the community: after all, the former defined the latter. His claim that public opinion was infallible could be understood in the same way. If a community's interest was nothing but the sum of the interests of the individuals who composed it, then how could the opinion of the majority of that community—properly constructed, informed and expressed—possibly diverge from the interests of the community?

Elsewhere in the "Constitutional Code", Bentham also says that public opinion is "incorruptible". This might sound like the statement of a naive relativist; and reinforce some of the criticisms that Hart lays at Bentham's door. Again, this claim is, however, less surprising if we reflect on what corruption meant for Bentham. He famously castigated the ruling elite as a source of corruption, and described the King as the Corruptor-General (Schofield, 2006, 42). Bentham criticised the monarch for distributing peerages to promote his own particular interest.

This was sinister interest made manifest in corruption: “reward, becomes corruption when applied to a sinister purpose, when applied to a man, in such a manner as to direct his endeavours to the doing good to the one or to the few, at the expense of preponderant evil to the many” (Bowring viii, 545-6 in Schofield, 2006: 142). In a sense, this kind of patronage was the opposite of the secret ballot. Voting in private enabled a genuine expression of individual interests by the ordinary majority of voters. In contrast, corruption by patronage made the votes and decisions of influential men subservient to those who had patronised them. The former removed—or mitigated—the dissocial or selfish principles governing individual behaviour; the latter exploited and exacerbated them. The idea of corruption was thus related to incentives. It entailed their used to benefit a “sinister” few rather than to the community at large.

In this context, it is perhaps less surprising that Bentham later deemed public opinion incorruptible. If it reflects the interests of the individuals who compose the community, what incentives could be offered to its constituent members to corrupt them? A relatively small payment or bestowal of prestige might be enough to corrupt a single peer. The corruption of public opinion at large, however, would require a majority of people in the community to be given an incentive to serve a particular interest. Of course, if the interest of the community is nothing more than the sum of the interests of the members who compose it, then the act of providing a strong incentive to a majority of individuals—who themselves create public opinion—would itself render a measure in the public interest, by definition. Thus, while public opinion might be subject to delusion and confusion, it could not be “corrupted” in the way that a single peer, receiving prestige from the Monarch, could be.

It is clear from the role described for the Public Opinion Tribunal (POT) in the “Constitutional Code” that Bentham believed public opinion was absolutely critical to the success of a democracy. It operated in parallel with voting, as well as through it. He officially recognised public opinion as a formal part of the constitution, and lauded its influence. Public opinion is said to be progressive and incorruptible; to be the best hope of good governance; essential to the success of democracy. Bentham believed that history showed public opinion had the capacity to influence the behaviour of public officials. Its effects could be observed even where formal democratic institutions were absent. Thus while “rulers are by the unalterable constitution of human nature, disposed to maximize the application of the matter of good to themselves” (1983b:20), historic abuses had not always been as great as they might have been. This applied even to states ruled by monarchs with unlimited constitutional power, and is worth quoting at some length:

“As to moral responsibility, imperfect as it is, this species of security against misconduct is the more necessary to be brought to view, inasmuch as, in monarchies in general, were it not for this, there would be no responsibility at all: and, in other words, the monarch would be altogether without motives for compliance with the laws, even with those of his own making, which are, at all times, such as, and no other than, it is agreeable to him to make. It is by this source of restraint alone, that the English form of government—a mixture, composed of monarchico-aristocratical despotism with a spice of anarchy—has been preserved from passing through the condition of France, Russia, and Austria, into that of Spain and Portugal. Even without the assistance of a posse of his own creatures, acting under the name of a parliament—he may kill any person he pleases, violate any woman he pleases; take to himself or destroy any thing he pleases. Every person who resists him while in any such way occupied, is, by law, killable, and every person who so much as tells of it, is punishable. Yet, without the form of an act of parliament, he does nothing of all this. Why? Because by the power of the Public Opinion Tribunal, though he could not be either punished or effectually resisted, he might be, and would be, more or less annoyed.” (Bentham, 1983b:25)

This clearly confirms that Bentham saw the outputs of public opinion as more than just whatever voters happened to express at the ballot box. It was a critical and essential security against misrule: a protection against abuse of public office. Whilst imperfect, and insufficient by itself, the opinions that the masses held of their rulers could exert some influence over their conduct. By doing so, it could improve governance. That security was principally exercised by encouraging the moral responsibility of rulers. As we have seen, Bentham understood this responsibility or “moral aptitude” in terms of an alignment between self-interest and the community’s interest; not as the cultivation of officials’ virtuous character. Public opinion was important for precisely that reason: it was able to produce greater moral aptitude in public officials through what Bentham called the moral sanction, which operated through public opinion.

There was therefore a clear relationship between the responsibility with which governance was conducted, and the power of public opinion: “appropriate moral aptitude must be considered as exactly proportioned to the strictness of the functionary's dependence on public opinion” (1983b:174). Good governance had to be induced via a situation which made officials dependent on the public. This required publicity around government and the effective operation of public opinion. The Public Opinion Tribunal (POT) illustrates this point. The POT was to advance the happiness of the community through public meetings and publications; the most important of which were newspapers. It had a series of sub-committees to reflect different aspects of its role. As we have mentioned briefly above, its contribution to making officials dependent on the people was made through four distinct functions: Statistic, Censorial, Executive, and

Melioration-Suggestive (1983b: 36-8). Each of these tells us something about the kind of role that Bentham wished public opinion to play.

In its Statistic capacity, the POT collected facts and evidence to inform the judgements its members made about the conduct of representatives. It is fair to conclude from this that Bentham placed significant importance on the need for public opinion to be sufficiently informed. His writing about the Tribunal does not suggest he held a view that expressed preferences can be blindly aggregated at the ballot box. Rather, he was attentive to the need for a sufficient volume and quality of information to be available before citizens made political judgements.

The Censorial function was a matter of forming judgements; principally about officials' conduct, policies, and other relevant questions. Members would then express approval or disapproval—approbation or disapprobation—according to whether conduct or policy was considered useful or detrimental to the greatest happiness of the greatest number. Bentham believed that the judgements formed by the Tribunal had great potential power. In “Political Tactics”, for example, he suggests that the “tribunal... is more powerful than all the other tribunals together... [and] always decides the destiny of public men” (Bentham 1999: 29-34 in Schofield, 2006:260). It was therefore very important that the Tribunal accurately understood whether officials' conduct was advancing or retarding the interests of the community. A failure to understand pleasures and pains, and how they related to political conduct and policy could cause public opinion to impede progress rather than promoting it. The judgement of the public was thus critical. It held the capacity to direct the power of the POT; either accurately or inaccurately. The concept of error in the formation of individual judgements, and public opinion was thus very important to Bentham's account of democracy. We shall return to that important question later.

The Executive function enacted the judgements of the POT. It created rewards and punishments by enforcing the will of the people through the moral sanction. As we explored in the previous chapter, Bender advances a very particular interpretation of this function. He suggests that the POT was to apply the moral sanction to all members of the community. By doing so, it would promote consensus by better aligning the incentives of individual citizens with those of the whole community. I do not believe that this reading is supported by references to the POT; either in the “Constitutional Code” or in Bentham's correspondence and other work. Some examples may illustrate this point. In 1823, Bentham wrote to Simon Bolivar to recommend constitutional reforms for Columbia. This was one example of the significant efforts Bentham made to ensure that his ideas for reform had a practical influence around the world. This letter was written relatively late in Bentham's career: six years after the publication of the “Plan for Parliamentary Reform”. Consequently, it also provides an interesting insight into how Bentham

thought his ideas should be applied in practice. Addressing Bolivar as Columbia's liberator, Bentham argued that a free press was essential to good governance because:

“Only from such sources can the citizens... form any tolerably well-grounded judgement on the question whom to choose for Representatives, or, in their quality of Members of the Public Opinion Tribunal, on the question of what measures and arrangements to disapprove, what to approve, what persons, and in particular what functionaries, to punish by their disesteem, what to reward by their esteem. But, as the principal use of all such publications is, to serve as a bridle to the force of sinister interest in the breasts of rulers”. (Bentham, 2000:254-5)

For Bentham, the “principal use” of publicity was to restrict the operation of sinister interests in government rather; not to create deliberative consensus. Thus, while Ben-Dor portrays the POT as akin to the conditions for the formation of Habermasian consensus, I do not believe that the way Bentham explains its workings corroborates this reading. In this quote, as elsewhere, we see that Bentham emphasises the importance of *both* citizens' votes, channelled through formal democratic institutions, *and* the application of the moral sanction through the Public Opinion Tribunal as a security against misrule. It is clear that the operation of the POT has a significance beyond its influence on the way that citizens vote. That significance—through use of the social sanction—was however described as a security against the abuse of power; not as a driver of consensus.

There is more evidence to reject a stronger deliberative interpretation of the POT in “Economy as Applied to Office”. There, Bentham paints the Public Opinion Tribunal as a majoritarian numbers game; not a deliberative forum that generates consensus. The argument Bentham makes here to describe the value of the POT resembles the aggregative logic of thinkers like Condorcet (1785) rather than the deliberative views of later Habermasian thinkers. He suggests that the decisions of the POT should be understood as those of the majority of its members, and distinguished from the contrary views of a minority.

Bentham suggested that no correct decision could ever be known with certainty. The value of the POT could thus be understood in terms of probability. Given appropriate conditions, the probability of a majority of participants being right, was generally greater than that of a dissenting minority being correct:

“The tribunal of public opinion having for its members the majority of all such individuals as take cognizance of the individual case in question, the decision capable of being in a state of contrast

with and opposition to the decision of such majority can not be any otherwise than the decision of the smaller number". (Bentham, 1989:57)

Again, there is little here to suggest that Bentham expected members of the POT to achieve the kind of full consensus later described by Habermas. Rather, the implication is that the POT will always reflect the numbers involved and, perhaps even that a residual minority view is to be expected, not discouraged. Bentham therefore describes public opinion as a more reliable source of correct opinion than individual judgement. The basis for its reliability was to be found in provision of appropriate information—as the Statistic function of the POT implies—but principally in the weight of numbers held by the majority. Contemporary writers have referred to a similar principle as the “miracle of aggregation” (Page & Shapiro, 1993).

Similar arguments can be found elsewhere in Bentham’s constitutional work. In the “Necessity of an Omnipotent Legislature”, for example, he again implies that the value of public opinion is in part derived from the weight of numbers producing it, and the increased chance of correct opinions that this generates:

“...he who, in judging of the course that will be taken by public opinion, follows the light of reason, will but too often miss the mark: but his chance of hitting it will at least be less bad than were he to take for his guide the ignis fatuus of caprice”

(Necessity of an Omnipotent Legislature, in RRR, 2002:276).

While imperfect, public opinion thus provided more reliable judgements than available alternatives. The rules that Bentham outlines for the Tribunal tell us more about how he thought the public should reason together. The POT was to operate as a source of moral aptitude in public officials by exercising the moral sanction. It was to have an unlimited membership; including anyone who could be affected by, or be cognizant of, the issue in hand. Bentham therefore believed that individual qualifications were unnecessary for members to make a valuable contribution to the Tribunal’s judgements.

This is an interesting view; particularly since Bentham did not say the same of voting. In the “Parliamentary Reform Catechism”, he began with an assumption that suffrage should be universal, because all interests must be advanced and considered. Quickly, however, he conceded that “if it be sufficiently clear that any class of persons neither are nor can be in such a state of mind as to possess the

intellectual fitness necessary to exercise the right of suffrage with advantage, then such a class may be excluded” (Bentham, 1818a:35). The right to vote was thus conditional on persons’ ability to use that right productively. Illiterate citizens were to be temporarily excluded while they were taught to read (Ibid: 37).

On these pragmatic grounds, Bentham also concluded that minors should be excluded, and that there were no clear grounds for excluding women. Bentham’s views on this question oscillated over his career. In “*Considérations sur la Composition des États-Généraux*”, he was swayed by five pragmatic arguments that women should not be enfranchised. One of these was that they “already possessed sufficient influence over males by the greater need of the latter for sexual gratification” (Schofield, 2006:86). Whilst Bentham’s starting assumption or “supreme principle” implied universal suffrage (Ibid), therefore; different groups had to be assessed according to evidence about their aptitude and, therefore, the usefulness of their participation. In the Public Opinion Tribunal, we find no such distinctions or exclusions. The “Constitutional Code” does not assess how qualified different groups are to participate, or how those without the required aptitude should be trained. Instead, it operates on a simple principle: accessibility to all concerned.

The POT was inclusive in terms of reasons as well as persons. Bentham does not stipulate any detailed rules about the types of reasons that would qualify for consideration by the Tribunal. In fact, he explicitly acknowledges the importance of the POT using a wide variety of forms of communication. Rosen has noted that these included relevant artistic performances, as well as speeches, and the operation of the newspapers (Rosen, 1983:27). This inclusive description of the POT provides some important indications of how Bentham wished public reasoning to operate. It was open to all. It should: focus on the conduct and policies of government officials; collect information; and use the weight of numbers involved to progress toward correct judgements. To that end, Bentham also provided some indications about what he expected the Tribunal to achieve. Through the Statistic, Censorial, Executive, Melioration-suggestive, functions, the POT will speak for the public interest. The remarks I have cited suggest, however, that he did not expect it to do so by reaching full consensus through deliberation.

So how were members of the POT to speak for the public interest, if they were primarily self-interested individuals? This question is not quite answered in the first volume of the “Constitutional Code”. The role that the later Bentham gives public opinion in speaking for the public interest thus seems to conflict with his conception of citizens as individualistic utility maximisers. Each individual approaches discourse—

and voting—with the intention to maximise his own utility. How then might a single, coherent view of the public interest emerge from pre-voting discussion? Rosen (1983) acknowledges this as one of the significant unresolved questions in Bentham's text. From the mass of egoistic voices in a society, Bentham believes that a 'Tribunal' can be formed of individuals who collectively speak for the public interest. Bentham is clear that the POT should voice its views through a liberal press. But he does not explain exactly how the published views of a (middle class) writer could be expected to represent fully the interests of masses of working class readers. Can a self-interested journalist "speak for" the public interest?

How might Bentham have responded to that challenge? For members of the POT it is possible that there is a general coincidence between their particular interests and those of the whole community. For the same reason that secret voting was appropriate for the Democratical Section, it may be that almost all members of the Tribunal can be expected to have particular interests which serve the community. For this reason, Bentham had conceived of voting as a reflection of particular interests rather than an effort to abstract away from them. That simple answer cannot, however, answer the question. Although Bentham expected the majority of the POT to be comprised of the Democratical rather than the Aristocratical Section, he did not advocate it being exclusively so. Many of the people who play critical roles in the POT—journalists and editors—could be middle class. After all, the expertise required to perform those roles, alone, might make their circumstances and interests significantly different from those of the majority of workers.³³ The role of persons like these in representing the public interest is difficult to make sense of; given Bentham's insistence that every individual would largely be directed by his own particular interests rather than an impartial regard for the community. There is thus an unanswered question in the "Constitutional Code" about exactly how we might expect the POT to speak for the public interest when it relies on particular people, with particular interests.

Although Bentham does not explicitly answer this constitutional question, I believe other themes in his writing provide some potential answers. None of the individual members of the Public Opinion Tribunal is expected to move beyond her own interests and consider the interests of all in an impartial way. Bentham only made that demand of public officials. When it came to public officials, he proposed that legislators should make an inaugural declaration. This included a public commitment to be impartial. Future legislators in Bentham's democracy would declare:

"sincere and anxious shall be my endeavour, to keep my mind as clear as may be, of undue partiality in every sense... in favour of any class or individual, to the injury of any other... through

³³ See (Rosen, 1983: 27-35). This issue is covered in more detail below.

self-regarding interest... against partiality in favour of superiors, to the prejudice of inferiors: of superiors, in whatsoever scale of comparison: opulence, power, reputation, talent, natural or acquired” (Bentham, 1983b:144).

Here, we can see the marked difference between Bentham’s standard for members of the public and the officials who represent them. The former could usefully reflect and express their particular interests. Often, the latter should publicly commit to impartiality. Doing so would encourage them to move beyond their own particular interests and instead prioritise the public interest. This public commitment was designed to influence officials’ behaviour. It would create greater scrutiny and accountability. Similarly, the conduct of officials was to be the principal object of the POT; not reshaping the opinions of citizens to foster consensus. There is a sense here that public officials must be open to unconstrained critical scrutiny; this was not necessary or useful in respect of private citizens. Officials were paid for being “written at”, in the same way that soldiers were paid for being “shot at” (1983b:40). This is consistent with Bentham’s arguments on liberty of the press. As Rosen has noted, Bentham:

“...proposed that in place of the prevailing practice, where severe punishments were handed out for criticism of public officials and much smaller ones for the defamation of private individuals, he would reverse this order so that the private individual would be secure against libel and defamation, but no punishment could be imposed for criticism of a public official except for false or groundless defamation” (Rosen, 2003:243-4).

Freedom to criticise public officials was thus critical to the success of the POT. This provides another reason to reject Ben-Dor’s consensus-based reading; a conclusion that is confirmed by Chapter 9 of “Economy as Applied to Office” (1989). There, the POT is introduced as a way of maximising responsibility from officials. They were expected to be subject to “reproach at the hands of the Public Opinion Tribunal, by which the power of the moral or popular sanction is applied as a counterforce to the legal power of the state”. The POT “shot at” officials; not citizens.

There is reason to believe that Bentham cast the value of the POT in terms of the numbers participating rather than the merits of deliberation. Relatedly, individual POT members were not expected to consider the public interest directly when participating in discussion, or when voting. This complements Bentham’s argument for the secret ballot. Bentham did, however, believe that every individual had shared interests in the “subordinate ends” he had identified. There was an inevitable disharmony between the Aristocratic and Democratical sections. There remained, however, grounds to expect some harmony

between the interests of all individuals. Whatever interests or preferences individuals chose to develop in their lives, these would be likely to be served by: security, subsistence, abundance and equality. In effect, these were universal interests, shared by every individual.

Whilst Bentham believed that these universal interests were a reality, he did not suggest that public opinion would inevitably discover and reflect them in a rational way. So whilst Bentham was generally, perhaps too, optimistic about the progress that public opinion would make, he identified some critical barriers to individuals having an accurate understanding of the interests that were relevant to political decisions. The efficacy of government itself depended on language being used in a way which accurately reflected relevant interests rather than obscuring or conflating them, since “perverting the import of words... undermines and destroys whatever security can be afforded by the laws” (Necessity of an Omnipotent Legislature, in 2002: 269).

As we have seen, the abuse of language had the capacity to disarm public opinion and enable sinister abuse of power. Critical public engagement was, therefore, needed to counter this threat, and to maintain a form of discourse which reflected interests rather than distracting from them. This would help to ensure that the POT allocated approval and disapproval appropriately: in line with the interests of the community. To this end, Bentham thought that fallacies were often used to deceive citizens. He defined a fallacy as “any argument employed, or topic suggested, for the purpose, or with the probability, of producing the effect of deception” (Book of Fallacies, Section 1.A). Here, too the idea of class is important. Bentham argued that there were distinct classes of men with the motivation to deceive others in pursuit of their “sinister” interests. Fallacies could operate on people’s affections, judgement or imagination. Bentham suggested that the principal course of action used by those who were opposed to measures which would serve the greatest happiness of the greatest number was “in the first instance, to endeavour to repress altogether the exercise of the reasoning faculty, by adducing authority in various shapes as conclusive upon the subject of the measure proposed” (Part 1, Book of Fallacies). Public opinion was thus impotent to the degree that citizens accepted these kinds of claims to authority, and empowered to the degree they rejected them and instead engaged critically. In reality, Bentham believed that the appropriate weight or authority given to a statement on a particular matter should depend on key factors. These were: the relevant intelligence of the person producing it; information available to them; their probity, and the fidelity of the means of communication used (Book of Fallacies, Ch1, 11). The probity of the conclusion expressed could thus show itself through the insincerity of the speaker, which Bentham thought was a common failing.

Beneath stated opinions, however, the way opinions were formed was also susceptible to improbity since “either the relevant means and materials are kept out of the mind; or...the attention is kept from fixing upon them with the degree of intensity proportioned to their legitimately persuasive force” (Ibid). In the *Book of Fallacies*, Bentham argued that the harm which fallacies could create falls into two distinct categories. “mischief within doors” and “mischief without doors”. The former took place through the deception of legislators within parliament. The latter operated through deception of “any person whose station is among the people at large” (*Book of Fallacies*, Ch. 2). Although the “*Book of Fallacies*” was meant to address deception within a parliament, therefore, it is clear that Bentham also wanted it to liberate members of the wider public too.

Barriers to the effective public opinion could thus be considered in terms of the fallacious logic Bentham diagnosed. Delusion, he suggested, “has place and operates in so far as, by some erroneous conception or opinion, individuals, who but for the error would not be, are operated upon in such sort that they give support to misrule” (1989:261). Delusion thus had two vital characteristics: it operated on the will via the understanding and in doing so, it facilitated misrule. As we have seen, Bentham understood sinister interest as a source of misrule, and the sacrifice of public for particular interests as essential to its operation. Its effects were often similar to those of corruption; which operated instead on the directly on the will by offering pecuniary or other rewards or punishments.

Over the course of his career, Bentham came to think that these “false lights” of corruption and delusion were near universal in reach. They operated in many areas of life, and across society. They could be found, for example, in his analysis of: religion, political elites, and the legal profession. Each shared a common object: “causing men to regard those by whom the powers of government [are possessed] as being every where possessed of a degree of appropriate aptitude beyond that which is every where possible” (Bentham, 1989:263). This unquestioning overestimation of public officials and others with claims to authority was inimical to the accurate understanding of interests that Bentham wished to promote. It created the foundations for a deluded and corrupted public that allowed sinister interest to dominate politics, and beyond. In the legal, religious, and political professions, the acceptance of claims to infallibility by members of the public disarmed and diminished the power of critical public opinion; allowing sinister abuses to go unrecognised and unchecked.

Official aptitude was Bentham’s antidote to this problem. As we have seen, however, a key part of official aptitude this was the moral aptitude of public officials. Bentham believed that this depended on the situation in which they were placed. As a result, the later Bentham argued that we could only expect this kind of aptitude from officials when they were given incentives to serve the community. Famously, this

required an “artificial identification” of interests (Halevy, 1952: 404-6). Publicity helped to achieve this identification because it made the personal fate of officials dependent on the public they served. Bentham placed great faith in the potential of public opinion to curtail official abuses of power. He suspected that, when subject to appropriate publicity, public officials would rarely, if ever, be immune to its influence. Consequently, he suspected, even officials who appeared to be acting virtuously and ignoring the public repercussions of doing so were often, in reality, secretly counting the personal benefits and harms of doing so by “weigh[ing]... in secret the suffrages of those who resemble himself” (Bentham, 1999:29-30).

Official aptitude has been the subject of extensive recent scholarship. What is sometimes given less emphasis, however, is that the later Bentham believed that moral aptitude could only be achieved through a particular type of publicity, and reasoning from members of the public. When describing how a community should promote the aptitude of its officials, Bentham describes how he believes they should discharge that duty. In effect, this is a form of public reasoning. It enhances the quality of governance. Some features of this account are evident in the problems Bentham describes. Public opinion would only play its proper role where individuals had a correct conception of interests. That required a discourse cleansed of delusion, and a community free from the corruption which had emasculated the public, and previously made them unaware of their own true interests. In “First Principles...”, a deluded public of this sort is likened to a baby, who responds unthinkingly to immediate surroundings:

“The baby who, as he is carried past the pastry-cook's, feels the saliva as it [is] called forth by the sight of a favourite cake, what should lead him to speculations about the past or future rise in the price of com to which the cake, if bought for him and eaten by him, will according to the value of it have contributed?” (Bentham, 1989:264).

Critical and informed public opinion thus guarded against sinister abuses; deluded and acquiescent public opinion allowed abuses to continue unchecked. In the former case, it was informed by an open government. This provided citizens with the information they needed to form a “proper conception of their interest” (Schofield, 2006: 252). In the latter case, it lacked this information, and was led—without critical reflection—like a child; by the will of a “sinister” minority. Grounds for understanding Bentham’s views about effective public reasoning can thus be found in the contrast he drew between effective and ineffective public opinion.

It was the purpose of constitutional law to replace the natural opposition of interests between political rulers and the public with an artificial identification of interests (Schofield, 2006:272). Creating this

identification—through critical public scrutiny—was the central task of effective public reasoning. The elements of public reasoning noted above should be understood in this light. They contribute to scrutiny of public officials. In doing so, they limit or prevent sinister influence. Public reasoning is thus one part of Bentham’s wider democratic principles. Principal among these was the attempt to create a “limitation of power held by those with charge of the governed” (de Champs, 2008: 122). This conception of public reasoning thus complements readings of Bentham’s democratic theory—advanced by Rosen, Schofield, de Champs, and others—which define Bentham’s identification of the operation of sinister interest as “the pivot of a utilitarian analysis of the mechanisms of government” (de Champs, 2008: 233). Without sinister interests, there would essentially be no need for public reasoning; or democracy.

We have investigated some of the problems that Bentham identified. The abuse of power by an influential elite was made possible by secrecy, and a lack of scrutiny. In turn, that lack of scrutiny was made possible by deception, corruption, and the abuse of language. Each of these problems points towards a potential democratic solution: official aptitude, fostered by public scrutiny. Official aptitude—particularly moral aptitude—relied on the public; not just officials. For this reason, I will refer to the public characteristics which help to create official aptitude as “public aptitude”. Its aim is to mitigate the causes of sinister abuse that Bentham had identified, chiefly by making effective use of a dependence of public officials on the public they serve.

Having examined some related themes in Bentham’s writing, we may now ask “what does public aptitude look like in practice?” The texts cited above offer some clear indications. The first volume of the “Constitutional Code” does not explicitly state the shape that Bentham wanted public reasoning before voting to take. There are, however, some important clues elsewhere in his writing. Several things are important here. Public aptitude is critical of—rather than acquiescent to—putative authority; it focuses on participants’ self-interest rather than the public interest; resolves disagreement through majority votes rather than full consensus; and attempts to clarify the terms of public debate by testing the language used by participants. The principal aim of public aptitude is to promote the official aptitude of those in government. Its scrutiny facilitates the alignment of particular and general interests that Bentham thought was essential to the moral aptitude of governors. It also has other benefits. Citizens reasoning in this kind of way also learn by doing so. The object of their learning is their own self-interest rather than the public interest.

Let us summarise four key characteristics of public aptitude in turn. Public aptitude:

- ...criticises authority and focused on participants' self-interest.
- ...analyses legal and political language.
- ...promotes learning by citizens and encourages them to make suggestions on policy.
- ...requires and defines no account of admissible persons or reasons.

The first notable characteristic of public aptitude is its orientation and nature. Public aptitude is critical and focused on the self-interest of participants. Its approach to putative authority should be understood as questioning rather than acquiescent, and self-interested rather public-spirited. This kind of thinking is the natural starting point for public discussion, given Bentham's analysis of the danger of sinister delusion and corruption of the public. It also follows the role that he gave public opinion in his constitutional theory: that of a key security against the misrule. Bentham wanted the public to emerge from its docile state of being "Habituated to receive and enjoy the effects of the corruption...[and] delusion... without reflection and almost without thought" (1989:264). From the role that he gives publicity in his developed constitutional thought, we can also adduce two things. First, he believed members of the public had to liberate themselves from corruption and delusion in order to reason effectively. Second, he also believed that the right kind of public reasoning, and publicity around parliamentary procedure, could itself help to improve the efficacy of public opinion in this capacity.

Confirmation of the first of these points is to be found in various places, including many sections of the first volume of the "Constitutional Code". There, Bentham contends that "Without publicity, no good is permanent; under the auspices of publicity, no evil can continue" (Bentham, 1999:37). To this end, the individuals who compose the public were expected to be critical and reflective about their own interests and, in doing so, promote the aptitude of their rulers. This was central to identification of interests on which Bentham believed the value of democracy depended when exercised both through the moral sanction and via voting.

On the second point, Bentham sometimes makes a further claim. In "Political Tactics", for instance, he suggests that publicity around political hearings—and appropriate deliberation within a parliament—can promote a more useful form of reasoning by the public. When those in parliament deliberated in the right kind of way, and appropriate publicity was given to their hearings, Bentham believed that citizens, too, would begin to reason more productively in service of the public interest in the long term:

"Among a people who have been long accustomed to public assemblies, the general feeling will be raised to a higher tone—sound opinions will be more common—hurtful prejudices, publicly

combated, not by rhetoricians but by statesmen, will have less dominion... A habit of reasoning and discussion will penetrate all classes of society” (Bentham, 1999:31)

Bentham believed that the right kind of deliberation within a parliament could have a lasting, transformative impact on the “national spirit” (Ibid). It is in the references to the habit of reasoning that this would create that Bentham provides evidence of his views on what would now be termed public reasoning. Given the multitude of references to the negative power of publicity, particularly in the “Constitutional Code”, Bentham clearly believed that its role as a security against misrule was central.

While reservations remain about how far “Political Tactics” can be relied upon as a reflection of Bentham’s views, there are passages in that work which complement this idea in the “Constitutional Code”, “First Principles...”, and elsewhere. As the POT suggests, the later Bentham was insistent that parliamentary procedures should be exposed to a great deal of publicity. He also implied that, in time, a culture of scrutiny and publicity would cultivate useful democratic habits and traits among the population. The fruits of publicity would thus be the opposite of the products of secrecy. Secrecy denuded the public of its aptitude: the characteristics necessary to hold officials to account. But publicity could reverse this by making the “multitude more secure from the tricks of demagogues” (Bentham, 1999:31).

Rather than appearing in isolation from Bentham’s earlier work, his later democratic views can be seen as an extension of his earlier writing on deception. The idea of a ruling class with interests opposed to those of the community also complements other ideas in Bentham’s writing; including his analysis of ideas of taste and the use of language. Delusory language was used to diminish public aptitude. It encouraged a passive and subservient public to ignore or acquiesce to the abuse of power. So too was the idea of public approval or taste. Bentham believed that ideas of good and bad taste were exploited by a ruling elite in order to delude the subject many:

... “into believing that they “(the aristocracy) were superior, and hence entitled to rule and to enjoy disproportionate quantities of wealth, power, and esteem. In other words, it was the interest of the ruling few to appeal to ‘taste’ in order to maintain their dominance over the subject many, and to provide an apparent justification of the oppression which they exercised... The principle of taste was adopted in order to subvert the principle of utility.”
(Schofield, 2014:97).

Taste was, therefore, one way in which individual judgments, and public opinion, could be led away from utility. It was another tool—used by the influential few—to diminish and dilute public aptitude. Doing so allowed greater opportunity for sinister abuse of power.

Bentham's argument for the secret ballot thus fit into wider themes in his thought. In particular, it complements his critique of delusions and fallacies, and the way this played into his conception of class. Together, those strands of his thought suggested that the disproportionate influence of a ruling few over the terms of public debate needed to be mitigated and controlled. This view had a specific linguistic element. Beyond the idea of taste, Bentham believed that elites held influence over the type of language used to describe different sets of preferences: this has framed debate for sinister ends. Bentham's views on taste also provide another reason to reject suggestions that he wished citizens to adopt a public orientation in their reasoning before voting. To do so would be to expose them to a wider range of delusory influences—including predominant notions of taste—and misleading abstract references to the welfare of the community (Bentham, 2002:320).

Several related themes in Bentham's thought come together in the ideas examined above. He believed that corruption and delusion: stymied progress; led citizens away from a proper understanding of their interests; and allowed a powerful elite to sacrifice the public interest in pursuit of their own particular interests. In short, these abuses blunted public aptitude, and reduced the accountability of public officials. Systematic delusion and corruption was only possible because the public lacked that aptitude. Public opinion was in a baby-like state, and could be easily manipulated. From Bentham's divisions between two classes in society, and his views about the incentives needed for each, we can see that he broadly advocated two different forms of reasoning. The Aristocratical Section was already dominant in parliament. There, they were to be held to account by the public. Their discussions were to be subject to publicity, which provided a security against sinister misrule. Their role was to be as impartial as possible and to consider the public interest. "Political Tactics" thus describes something akin to a contemporary deliberative account of parliamentary discourse. Representatives were to learn about the public interest from one another's contributions, and to collect the information necessary to doing so through debate and reflection. For this reason, the expressed preferences on which they voted should often change during debate. The success of decisions of this sort was dependent on the moral and intellectual (official) aptitude of those taking part. Bentham spent a great deal of time specifying the types of institutions, procedures and architecture necessary to maximising aptitude and minimising expense.

There are some similarities between the parliamentary reasoning advocated by Bentham and the "stronger deliberation" of contemporary theorists. Parliamentary representatives should move beyond their self-interest; which was not inevitably aligned to the public interest. This required that they should learn about

the public interest, and have the intellectual aptitude to do so. It also demanded that they should have the moral aptitude needed to serve the community. The public was thus to act as a security against misrule by officials. This crucial safeguard against sinister abuses could only be secure if the public themselves had a type of aptitude: that required to entertain a correct conception of their own interests, and scrutinise officials' conduct accordingly.

Effective reasoning within parliament required publicity. Its aim was to induce an artificial identification of interests between representatives and the public interest. In contrast, reasoning by members of the public required secrecy. This was because those taking part were in the overwhelming majority: their particular interests were usually already aligned with those of the community. This contrast suggests that Bentham thought the type of reasoning required from a group of individuals was a function of their likely interests; not substantive claims about the intrinsic value of types of participation. Official aptitude was thus essentially public; but public aptitude was private, and focused on self-interest.

Thus, effective public reasoning did not require the same kind of publicity. It aimed to produce genuine votes and opinions rather than an "artificial identification". This mirrored Bentham's argument for the secret ballot. The aim of private voting was to secure a "genuineness of suffrage" (Bowring, iii, 599). That enabled individuals to vote in a way that genuinely reflected their own self-interest. It gave them protection from the pressure created by the expectations of others in the community, particularly those—landowners—with power over their pecuniary interests. Together with Bentham's views about delusion, corruption, and the divergence of interests between the two "sections" or classes of society, this implies that pre-voting discussion amongst the Democratical Section should remain focused on the particular interests of "Democratical" individuals. Moral aptitude was thus a standard feature in the Democratical Section; but it had to be artificially induced among officials.

These arguments imply conclusions about the type of language that should be used during public reasoning. Bentham thought that delusory language, instigated by Aristocratals, was a major feature of much public discourse, and abstract (non-assignable) references to the public interest were prone to such delusion. It is, therefore, unlikely that he would have endorsed a public-interested deliberative view of pre-voting discussion. More generally, Bentham suggested that English law owed its obscurity to an "abstract way of speaking". Of such terms, he urged his readers to "avoid them or explain them by the relation they bear to the real ones" (Bentham, in Harrison, 1983:252). Indeed, Bentham explicitly related that linguistic point to the orientation of debate in the POT. In the "Constitutional Code", he wrote that

each member acts on his “primary concern [which] is to secure *himself* against oppression and depredation” (Bentham, 1983b:100; italics added).

For these reasons, the kind of reasoning required as part of the Public Opinion Tribunal cannot—contrary to suggestions from Crimmins (2011:154-5)—be equated with the 20th Century deliberative theory of Habermas. This point is not always made as clear as it could be in the secondary literature. Despite providing the authoritative analysis of the first volume of the Constitutional Code Rosen, for instance, also suggests: “Bentham believes that through the...secret ballot in legislative elections, electors will tend to consider the public interest rather than sinister interests” (Rosen, 1983:31). It is worth clarifying, I think, that Bentham expected voters to serve the public interest because there was a harmony between their own interests and those of the community; not because he wished them to disregard the former in preference to the latter. This clarifies and confirms that attempts to characterise Bentham as a public-interested deliberative theorist—proposed by Crimmins (2011) and Ben-Dor (2000)—are misplaced.

Rather, the POT could be said to “speak for” the public interest in a different sense. It gives individuals the information and understanding they need to understand their own interests clearly, and to vote accordingly. Doing so serves the public interest because of the weight of numbers involved. I believe that this view of POT is more compatible with wider themes in Bentham’s work—the predominance of self-interest; his deep scepticism about abstract references to the public interest—than more deliberative interpretations. The POT was principally a security against misrule; not an engine of consensus. It “shot at” public officials; not citizens (Bentham, 1983b:40). Indeed—contrary to Ben-Dor’s (2000) view—Bentham’s work on indirect legislation was a more prominent feature of his earlier (non-democratic) work than his later democratic writing. I believe that the idea of “public aptitude” better sums up Bentham’s aims. He wished public opinion to emerge from its “baby-like” state of unquestioning acquiescence to the authority of Aristocratals; and for members of the public to hold a more accurate understanding of the interests at stake when political decisions were made. A critical form of reasoning, directed toward the self-interest of participants, was best-suited to that end; not public-spirited deliberation.

Let us turn to a second characteristic of public aptitude in more detail. It should analyse language in an attempt to create what Bentham called an “enlightened public”. The “Book of Fallacies” shows that Bentham believed that false arguments could easily be accepted by members of the public and lead them to misunderstand their own interests and those of the community. Clearly, Bentham wanted public opinion to remedy these corruptive and delusory influences. The POT reflected his belief that it could do

so. In part, that could work through the attribution of approval and disapproval by members of the public to behaviour, tastes, and motivations. In reality, he believed that the opposite had historically happened. Disapprobation and approbation had been attributed to people and pursuits according to the interests of a dominant class; at the expense of a suffering subject class. Thus while the powerful were exalted in life, and in death, with plaudits and hagiographies:

“...the unopulent, the undignified, are spoken and written of as less benevolent, less knowing, less judicious, less able, than they are. Why? even that while they have thus been rendered objects of contempt, less sympathy and antipathy may be excited by the oppression, depredation and contempt of which they are the destined objects: less sympathy towards the sufferers by the injury, less antipathy towards the authors of it” (Bentham, 1989:266).

As seen, Bentham aimed to promote the development and enlightenment of public opinion. In future, with the right information and opportunities, the approval and disapproval of the public was to reflect an accurate understanding of the interests concerned; not just the prejudices of a self-serving few. Delusory language was an obstacle to that enterprise; and it is reasonable to suppose that Bentham wished public reasoning to address this. He believed that the inaccurate use of language in this way prevented citizens from having a clear and correct conception of their own interests, and those of others. Discourse thus needed to be cleansed of the myriad of delusory inaccuracies in language.

One way to understand how the Public Opinion Tribunal “speaks for” the public interest is thus to say that it contributes to citizens’ understanding of their own interests. By doing so, it makes them better placed to use their vote effectively. Part of this role involves providing information to help people to vote “better”: the Statistic Function of the POT. Another part involves expressing approval and disapproval as a check on the ruling elite: the Censorial Function. In much of his correspondence, Bentham refers to these two roles as separate and complementary. Both should be directed towards the same ultimate end. But this was only possible if citizens understood the interests at stake; if they are protected from deluded conceptions of their own interests, and those of the community. Citizens themselves needed the aptitude to understand their interests in order to induce official aptitude in their representatives.

We could thus understand the POT in terms of Bentham’s earlier work on language. It might seem an undue revision to put together Bentham’s work on constitutions and language in this way. This is not the case. Schofield (2006) has shown that Bentham’s attempts to purge legal, and political language of error were fundamental to his project of democratic reform. It was to combat the usurpation of the public political institutions that Bentham thought democratic institutions and procedures were necessary. Further, we have seen that there was an intimate relation between Bentham’s understanding of the delusion—through the abuse of language—and the political operation of sinister interest. Bentham

believed that delusions could be used by rulers to usurp public opinion in a similar way. This illustrates, for example, why delusion features so prominently in the “First Principles” Bentham wrote in preparation for the Constitutional Code.

In that context, it would not make sense to merely allow people to vote without first giving them the resources to check that they can do so in a way to benefit themselves rather than their rulers. The POT is part of this effort. Bentham believed that a process of linguistic clarification was necessary to understanding actual interests; rather than misunderstanding fictitious entities. In his earlier writing on language, he explained that two processes were essential to understanding language, and using it correctly. He termed these “phraseplerosis” and “paraphrasis”. These complementary techniques aimed to explain the meaning of nouns which referred to fictitious entities in discourse by demonstrating their relationship to those which described real entities. He then used a technique which he termed “archetypation” to reveal the way in which he believed all language was rooted in a physical image (Schofield, 2006:23).

As a first stage, phraseplerosis sought to put words in the context of a proposition. Bentham thought this was essential to understanding their meaning. Paraphrasis could then take place, by which the meaning of that sentence was then translated into another sentence, which expressed its meaning using only simple ideas. The aim was to turn general and abstract ideas—references to “fictitious entities”—into the simple images which originally gave them meaning. Ultimately, Bentham believed, those general terms could either be understood “to raise *images* either of *substances* perceived, or of *emotions*” or they lacked meaning altogether (1977: 496n. in Schofield, 2006:24). Bentham’s references to the obscurity and opacity of the law; of public discourse; and of delusory influence should be understood in this context. It was partly in the failure to expound and understand the relationship between common abstract terms and the simple ideas which ultimately gave them meaning that these errors found fertile territory. Eradicating these kinds of mistakes—and purging language of error—was essential to creating an enlightened public: to realising, and practicing, public aptitude. In turn, this enlightened public was necessary to foster official aptitude. It is thus a reasonable assumption that this kind of linguistic analysis should be part of the critical analysis Bentham wished members of the public to exercise as members of the POT. Reflecting on language in this way, and exposing the inaccurate use of fictitious entities, would also help to guard against the use of fallacies. These also relied on the inaccurate use of references to fictitious entities; citizens who understood the real pleasures and pains at the root of political discussion would not so readily acquiesce to fallacious logic.

It was thus important that citizens should use this type of analysis to test and correct the assertions of public officials. Doing so would guard against the errors which had disarmed public opinion for so long. It would enable the POT to attach approbation and disapprobation to their conduct and policies on the basis of a genuine—rather than delusory—evidence. Bentham believed that intra-parliamentary debate had an analogous problem: it, too was replete with dangerous fallacies. Thus, the usefulness of intra-parliamentary debate was reliant both on proceedings being publicised, and on members of the public engaging with proceedings in a critical and engaged way. Publicity and criticism would, therefore, improve the quality of debate within parliament, and the results of its decisions:

“In a corrupt Assembly, a debate is a mass composed of one-tenth of silver drowned in nine-tenths of an alloy composed of lead and arsenic. To this compound, a Book of Fallacies aptly constructed would be an appropriate test. Applied to it, the arsenic would fly off in fumes: the lead would sink: the lucid bullion would be left pure.”

(Bentham, “First Principles...”, 1989:65)

The process of linguistic clarification could be construed as part of a wider learning process. Bentham was not an advocate of public-interested deliberation; for the reasons stated. There is evidence, however, that he wished citizens to learn from one another before voting. It should thus be noted that public aptitude helps citizens to learn from one another, and to make suggestions regarding policy. Thus while some aggregative democrats follow Condorcet (1785), and model voters as independent inputs to a democratic equation, Bentham—in *Political Tactics* (1999) at least—explicitly rejects that approach:

“This idea of absolute independence in the voters is absurd...the secret mode of election does not diminish the influence of mind on mind” (Bentham, 1999:146)

Although Bentham accepted the utility of this interaction, it was critical that it should take place through the influence of “mind on mind”. Unlike corruption—influence upon the object’s will—or delusion—influence on the will through deception of the understanding—persuasion was an acceptable and useful part of debate. Public opinion should have a positive role in a functioning democracy. The POT thus operated in a “melioration-suggestive” capacity; making positive proposals, as well criticising official conduct. Less is said about this positive role in the “Constitutional Code” than its critical capacities. I think that it is important to acknowledge, however, that Bentham’s views did not preclude citizens learning from one another through discussion before voting.

There is another reason why learning through public reasoning could be important. Individual citizens could have privileged information about their own motives and interests that could only be shared through such interaction. One of Bentham's arguments against colonialism illustrates this point. In "Emancipate your Colonies!" Bentham advanced a series of arguments for decolonisation. One of these was an appeal made to the National Convention of France in 1793. It asked readers to consider whether colonial rule could be better for those ruled over than self-government. For Bentham, the answer was a definitive "no". In part, he had an epistemic rationale for that conclusion. Colonial rule asked colonisers to "govern a million or two of people you don't care about...a set of people whose beliefs you know nothing about" (Bowring, II:4). In this sense, Bentham implied that government has to be close to the individuals it affects, to serve their interests. Decisions taken without sufficient information or understanding—like those of French legislators over the affairs of distant sugar colonies—would fail to promote utility.

Finally, unlike many contemporary deliberative accounts, public aptitude notably does not define the "admissible" persons or reasons required for public reasoning. As we have seen, Bentham saw the POT as an inclusive body with universal membership and no particular requirement for the types of reasons it should consider. The description of the POT relies on people participating to judge the relevance or admissibility of arguments rather than stipulating those ideas from the start. Bentham does not define the types of reasons that are (in)admissible for the POT's consideration. Rather, he says that anything with relevant content will contribute to the Tribunal's judgements. A broad and inclusive notion of admissible modes of communication is adopted. This includes plays and works of art. Consequently, there is no prescribed test of "reasonableness" or rationality for citizens to consider before they reason together.

Of course, Bentham did believe that the POT should reach conclusions based on the pleasures and pains at stake. He did not, however, pursue that goal by stipulating that citizens had a duty to only present certain types of reasons; or that only suitably qualified persons should participate in pre-voting discussion. Bentham's views on public reasoning are thus permissive in a way that few contemporary deliberative democrats are. Indeed, his argument "On the Liberty of the Press" suggests that he generally believed in the utility of false arguments. Thus when rejecting state censorship (Bowring i: 575) he cited grounds later made famous by Mill:

"The true censorship is that of an enlightened public, which will brand dangerous and false opinions, and will encourage useful discoveries" (Bowring 1, 538)

Conclusion

To recap, I believe that the conventional reading of Bentham underplays the importance of pre-voting discussion in his work. In contrast, the participatory reading does acknowledge the significance of public participation before voting. In doing so, however, it departs from some important principles in Bentham's writing, particularly the predominant self-interest of citizens participating which lends itself to a self-interest-orientated view of public reasoning.

In this chapter, I have laid out some themes in Bentham's thought, and tried to suggest that a different understanding of public reasoning is tenable. This "hybrid view" sits between the conventional and participatory interpretations, and shares some characteristics with each. It concurs with the conventional view that that public interest can be reached through aggregation. Consequently, citizens should principally seek to understand their own interests before voting. It also agrees with the conventional view that public reasoning should reflect the public interest primarily through weight of numbers; not via consensus reached through transformative deliberation. Indeed, the idea of deliberative consensus runs counter to Bentham's belief that abstract references to the public interest were liable to deceive.

The hybrid view also agrees with participatory accounts that there is an important place for pre-voting discussion in Bentham's account of democracy. That place can be understood in terms of what citizens can learn from one another through participation. Unlike participatory theorists, however, the hybrid view does not expect discussion between citizens to create consensus before they vote. Nor does the hybrid view expect that individuals should have to "move beyond" their own interests in pre-voting discussion. Instead, it is driven by the need for citizens to clarify their understanding of their own interests before they vote; for them to have the "public aptitude" needed to secure the official aptitude of their representatives. This complements Bentham's view that public officials can only be held to account by individuals who understand their own interests. Together, these individuals compose a "tribunal...which decides the destiny of public men" (Bentham 1999: 29). That Tribunal should be critical of how power is used in the public's name. This principle reflects the story of Bentham's conversion to democracy. That transition itself hinged on the idea that sinister interest must be negated through accountability to the public.

Under this "public aptitude" view, citizens should examine the conduct of public officials and ask whether that conduct serves *their* interests. Their answer to that question should then be expressed through opinions shared with peers, and through political action; including voting. Public reasoning thus reflects a kind of public aptitude. It actively criticises the conduct of public officials, and the language used to describe political action. It seeks to remedy the ills which Bentham thought plagued government. Its aim

in doing so is to make officials dependent on the public they serve. Having outlined key features of this view, we shall—in Chapter 5—assess some of its strengths and weaknesses.

Chapter 5. Implications of the hybrid view

“The military functionary is paid for being shot at.[the public official] is paid for being spoken and written at”

(Bentham, 1983b:40)

Introduction

So far, we have examined principles found in Bentham’s work without considering their strengths and weaknesses fully; or the contribution they could make to the contemporary debate on public reasoning. In this chapter, I want to make progress in that regard by taking a closer look at these principles and some potential objections to them. I believe that the “hybrid” account corrects some misunderstandings about Bentham’s own views, but it also has some implications for contemporary theory. In particular, it illustrates how much common ground there is between deliberative and aggregative democratic types of theory. Understanding this point should lead us to question some of the critiques of respective theories levelled by writers from these competing approaches. Of course, the conclusions drawn in this chapter will depend on precisely which construal of Bentham we consider. In particular, it is important to distinguish between the views that Bentham himself outlined, and broader “Benthamite” principles inspired by them. I will consider these two views separately: those *of Bentham*, and those which are *Benthamite*.

What is relatively novel about the broad “hybrid Benthamite” approach I have outlined is that it is aggregative and also suggests that preferences can be improved through dialogue before voting. I think that this points to a wider trend in the contemporary debate: there is important common ground between aggregative and deliberative theory. Additionally, it also illustrates that an account of public reasoning is possible, which both explains the value of pre-voting discussion; and escapes accusations of deliberative paternalism in doing so (Posner, 2004). Thus, deliberative theorists traditionally explain the importance of discussion by recourse to some fundamentally moralised notions; or by stipulating procedural conditions for preference-(trans)formation. Doing so convincingly explains why one should consider pre-voting discussion an important feature of democracy: it “improves” preferences in relation to these prescribed standards, or through these procedures. The danger of such standards is, however, that they can sometimes imply that citizens’ preferences are merely the result of deliberative failings (Waldron, 1999:111). Expectations of this sort may, therefore, contribute to something of an elitist perspective on citizens’ preferences, and a related potential for them to be manipulated during pre-voting discussion. In this thesis, elites are considered to be individuals or groups with influence on debate that cannot be justified by the

number, or expertise, of people involved. The hybrid reading of Bentham is unusual because it provides an explanation of the value of pre-voting discussion, but does so in reference to citizens' self-interest, rather than moralised terms of discourse, or the public interest. As a result, it may be less susceptible to the critique of deliberative democracy proposed by writers like Posner.

Common ground in the contemporary debate is easy to miss. Some aggregative democrats deride deliberationists as naïve believers in a set of ideal discursive conditions which may never obtain (Posner, 2004). Deliberative democrats often claim that their approach is essential to understand all but the thinnest conceptions of democratic procedure (Barber, 2003:198; Fishkin, 2008). I think that those arguments oversimplify the differences between these two approaches. This is perhaps a result of how deliberative theory evolved over the Twentieth Century. Theorists like Habermas defined their views in contrast to a "liberalism" which espoused the aggregation of self-interested individual preferences. Part of his enterprise was thus to overcome the "egocentric viewpoint" implied by those theories through a process of appropriate deliberation (Habermas, 1989: 45). A somewhat binary distinction is thus sometimes implied by Habermasians; between the more generally justifiable political decisions produced through deliberative participation, and aggregative theories which are content to rely on individuals' pre-deliberative preferences as grounds for collective decisions.

This form of deliberation was often seen as a response to the failings of rational choice, which made unjustified assumptions about individual behaviour and motivation. It was sometimes understood to assume, for example, that the preferences individuals express should be held as fixed in respect of collective decisions. The role of democratic process was thus to aggregate individuals' existing preferences rather than to reform them. Usually, Bentham is seen as part of this tradition.

Despite this, many theories of deliberation and aggregation are perfectly compatible with one another, despite their development in opposition to one another. The often-polarised debate between these two approaches misses some important distinctions. For example, many deliberationists propose objections to aggregative theory based on its assumption that voting should be a self-interested activity. As we have seen, however, aggregative theories may or may not hold this view. The assumption of self-interest may apply to some rational choice views, but not many social choice approaches (List & Goodin, 2001). Consequently, objections to aggregative theory based on this assumption alone can be, at best, incomplete. The assumed motivation for voting cannot, alone, distinguish between aggregative and deliberative approaches.

Similarly, the expectation of consensus—while important—is also not enough to distinguish between aggregative and deliberative democratic theories. From the aggregative perspective, it may be supposed that all deliberative theories risk false consensus by expecting participants to reach full agreement through discussion. Of course, deliberative theories are more varied than that. Many Habermasians expect full consensus. However, other contemporary deliberative democrats—like Gutmann & Thompson—construe deliberation differently. They describe participation as a way of clarifying arguments and information, rather than resolving fundamental disagreements. This particular objection therefore applies to some Habermasian views, but it cannot provide a blanket objection to deliberative theory in general. These are just two examples to show that the distinction between aggregative and deliberative approaches is less binary, and less clear than often supposed. In this context, Bentham's views could meet a number of objections from contemporary theorists. Below, I outline four of these challenges: the free rider problem, the danger of sinister interests within the Public Opinion Tribunal, questions over Bentham's optimism about public opinion; and a challenge of how to deal with what Dworkin (1978) calls "external preferences".

I believe that Bentham's rich and nuanced views provide a compelling response to some of these objections. The "free rider" problem is a good example. Some contemporary theorists allege that an account of public discourse premised on the self-interest of participants would be susceptible to allowing some participants to gain the benefits of public reasoning without the costs of participating. However, Bentham himself offered an understanding of self-interest which is expansive enough to respond to that objection. Other challenges, in contrast, are more difficult to counter using principles Bentham outlined. He did not, for example, comprehensively explain how the press—on which the Public Opinion Tribunal depended—could itself be protected from corruption by sinister interests. Given how the ownership of the press has developed, this is an important, though surmountable, oversight.

Second, I assess some objections to a broader Benthamite view. This approach is "Benthamite" rather than Bentham's. It follows the principles Bentham outlined, but does not exclusively reflect positions that he held. A view of this sort can be understood in relation to the four key questions, outlined in Chapter 1: focus, transformative power, status, and output. For the reasons outlined in Chapter 4, I believe that a Benthamite view should focus public reasoning on the self-interest of participants; reflect and clarify preferences rather than transforming them through deliberation; ascribe no particular moralized status to discourse, and should not expect full consensus. These characteristics mean that a "hybrid" Benthamite view could contribute to the contemporary debate on public reasoning. In particular, it illustrates the

inadequacies of the binary distinction—between aggregative and deliberative theories—used by so many modern theorists.

In a contemporary context, this broader view would also be subject to some important challenges. I examine three of these, regarding: power structures, the aggregative mechanism, and potential for a tyranny of the majority. Again, there is a convincing Benthamite response to some of these objections. It has been suggested that aggregative views which reflect rather than transforming individuals' preferences reinforce existing power structures (Gutmann & Thompson, 2004). I defend a Benthamite view against this objection. The kind of inclusive and critical dialogue that Bentham wanted to create provides opportunity for citizens to question and reform existing power structures. Indeed, the quotation at the start of this chapter illustrates that this critical approach is fundamental to Benthamite public reasoning.

Challenges to Bentham

The principles I have drawn from Bentham have some obvious implications. They challenge the idea that aggregative accounts of democracy say little about the importance of pre-voting discussion. They should also make us think carefully before assuming that supposedly aggregative democrats have little to say about the context for preference formation. Bentham himself distinguished between an accurate and inaccurate conception of individual interests. Nonetheless, self-interest-based views of democracy are relatively rare in the contemporary debate. One reason for this is the “free rider problem”, which has been a prominent issue in both contemporary economics and political science.

It is argued that when a group creates for itself a shared or common good, “each member of the group will have a strong tendency to be a ‘free rider’... to contribute little or nothing toward to cost of the good, while enjoying its benefits as fully as any other member of the group” (Kim & Walker, 1984:3). This is said to be a particular problem for democratic accounts which argue that individuals have self-interested reasons for participating in the democratic process. Many contemporary accounts of the free rider problem focus on pollution or voting (Downs, 1957). If individuals vote for self-interested reasons and, by doing so they help secure some common goods, it is problematic that they could free ride. They may benefit from the results of democratic governance without contributing, for instance by voting. The objection here is that a self-interested account of voting cannot plausibly ask voters to do something they have no self-interest in doing (vote), particularly when that activity is so fundamental to democracy.

The same accusation could be made against Bentham's view of public reasoning, with two implications: incoherence, and impracticality. First, a coherent democratic theory which presumes individuals' self-interest cannot, perhaps, ask individuals to participate in a political process when they have no self-interested reason to do. Second, free-riding might be a practical problem. It could make the public disengaged and thus prevent effective public reasoning, in Bentham's terms.

Consider how the question of coherence applies to the hybrid Benthamite view. I have argued that the grounds for public reasoning on Bentham's account is self-interest. In the "Constitutional Code", Bentham states that each participant in the Public Opinion Tribunal takes part to protect their own interests, particularly security (Bentham, 1983b:100). This complements his famous assumptions about the importance of self-interest to human motivation and behaviour:

"On every occasion, the conduct of every human being will be determined by his own interest... meaning according to his own conception of it, to the conception correct or incorrect entertained in relation to it by himself at the moment of action" (Bentham, 1989:68)

Assume that Bentham's views are correct on two matters: the Public Opinion Tribunal (POT) helps to assure the security of all; and participation should be voluntary. In Bentham's account of public reasoning, the good in question is somewhat unusual: security. The claim is that all citizens benefit from living in a society where public reasoning take place. This is because it holds sinister interests in check, and assures the security on which each individual relies in order to maximise their own utility. This security against misrule is a public benefit. Assume too that the group of citizens reasoning is likely to be large: Bentham gave the POT a universal membership. In large groups, it is likely that overall outcomes will not depend on the input or participation of any single individual. By the same token, an election result is very unlikely to be determined by any single voter (Downs, 1957).

It is likely that assuring the security of all through public reasoning will depend on *enough* citizens participating to make public reasoning effective, rather than on every individual citizen necessarily doing so. Bentham himself anticipated that the number of people participating in the POT would fluctuate: "the number may be of any magnitude not exceeding the sum total of the adequately adult members of the community" (Bentham, 1990:121). Each individual citizen therefore faces a choice: participate in public reasoning at a particular time, or decline to do so. From a self-interested perspective, there could be no incentive to participate. After all, group outcomes will not be affected by any single individual. The benefits of others' public reasoning will be available, irrespective of whether I take part. Participation also

has costs for the individual. Collecting relevant information, then making judgments and expressing decisions takes time and effort which could be spent pursuing other personal goals. Oscar Wilde reportedly said that “the problem with socialism is that it takes too many evenings”. We could forgive a member of the Public Opinion Tribunal for thinking the same of public reasoning. Bentham believed that the judgment of others could sometimes be trusted in debates, but never unquestioningly so (Schwartzberg, 2007). The costs of collecting evidence may thus be mitigated by others doing some of this work for the community, for example POT committee members. But forming a judgment *myself* as a member of the public will always take some time, effort, and engagement. A purely self-interested citizen may therefore choose not to participate in public reasoning. It may even be rational for them to do so: non-participation saves personal costs and leaves benefits unaffected. Free-riding is considered a significant problem for rational choice approaches; including those described in Chapter 1. It could also be a problem for the “hybrid” reading of Bentham.

The hybrid reading relies on public discussion and reflection taking place to protect and serve the self-interest of those who participate. As I have set out, this complements some wider themes in Bentham’s thought. In particular, it matches Bentham’s assumption that self-interest is a good predictor of behaviour in aggregate. It also fits with Bentham’s individualism, and his theory of subordinate ends. As we have seen, IPML outlines Bentham’s famous claim that the interests of the community are nothing more than the sum of the interests of the individuals who compose it. This implies that citizens reasoning about their own interests may contribute to the public interest by laying foundations for an appropriate aggregation of their views. Similarly, Bentham’s extensive work on subordinate ends suggests that individuals inevitably have important interests in common. These can be enunciated through discussion. This, too suggests that clarification—not transformation—of self-interested preferences is an appropriate aim for pre-voting discussion. Citizens should not move beyond their interests before voting. Rather, they should clarify those interests and, in doing so, discover what they also have in common. While it might be better for the community as a whole to benefit from the security against misrule which public reasoning assures, each individual also has a personal incentive to free ride, enjoying those collective benefits without incurring the costs of participation. There is thus potential incoherence in Bentham’s account. It assumes that self-interested individuals should participate in public reasoning, but cannot justify their participation in those terms. This mirrors accusations made against rational choice theories of democracy in contemporary debates.

Free-riding also has practical implications for the hybrid view. Assume that self-interest is a good predictor of individual behaviour, as Bentham suggested. If there is no self-interested reason to participate in public reasoning, it could be that no citizens choose to take part in this essential process. Critics could

thus suggest that the hybrid reading of Bentham on public reasoning is unrealistic in its own terms, unless participation is made compulsory for a sufficient number of people. Imagine that newspapers do, as Bentham wished, expose malpractice amongst the governing classes as a component of the POT. That publicity, in itself, would not protect the community from misrule unless sufficient numbers of citizens take notice. Unless they do so, and are prepared to change their political opinions and voting intentions accordingly, the Tribunal on which Bentham relied would have no way to use its political and social sanctions. It would be rendered toothless. Thus, while Bentham argued that public officials were “paid to be...written and spoken at” (1983b:40), communicative attacks could only provide effective scrutiny in the presence of an engaged public; one which noticed, and helped produce, criticism of officials’ conduct.

Free-riding could thus make the POT impractical; as well as incoherent. Bentham explained that the Tribunal should have committees with particular roles, but these alone would not be sufficient to assure its effectiveness without wider public participation. Both incoherence and impracticality are serious objections. I believe, however, that Bentham’s democratic theory provides adequate and convincing responses to the free rider problem, particularly in reference to public reasoning. What form might this response take? The first thing to note is that Bentham himself provides grounds from which we can infer that free-riding may not be a significant problem for his account. He believed that there was strong opposition of interests between the governing and the governed; and relative harmony within the governed, who composed the majority of the community. The harmony of interests among the governed could reduce the number of participants needed to produce a “sufficient” security against misrule by calling out bad practice. After all, fellow citizens who have the same interests as me may presumably be trusted more to call out abuses in roughly the same ways.

Similarly, the accusation of incoherence also does not apply to Bentham’s account. Contemporary theorists note that free-riding is less of a problem if participation in the provision of a public good itself confers benefits on the participant. Although some contemporary writers, like Gutmann & Thompson (2004) see Bentham’s views as non-participatory, he did believe that participation in pre-voting discussion created wider—including moral—benefits for participants. For example, we have seen that “Political Tactics” casts participation in public debate as a process that enhances the character of those involved; as well as making sound opinions more common (Bentham, 1999:30-31). It is likely that some of these benefits could only be achieved first-hand. Thus, while Bentham was sometimes content for citizens to rely on authority of others when formulating political opinions, he also describes a form of active participation and debate which should permeate all classes of society. The equivalent for public officials was “active aptitude”: the tendency to turn up and participate (Schofield, 2006:272).

Turning up was itself valuable. Some of the benefits of participation could only be secured first-hand. Only individuals themselves had the experience and knowledge needed to ensure that their interests were reflected in decision making. Bentham makes an argument of this sort when he distinguishes between public and private decisions. He believed that citizens themselves were the best judge of certain questions. These questions required their direct input because those individuals' responses to particular sanctions were too unpredictable for legislators account for in aggregate. This was the epistemic basis for Bentham's own version of a "best judge principle". For every individual, "the quantity of pleasure... which a man is liable to experience upon application of an exciting cause, will depend... [on his] sensibility" (Bentham, 1970:52). On matters like these, therefore, only individuals themselves had the understanding of their own preferences needed to make public reasoning work. For this reason, the success of governance and legislation depended on public participation. Without this, officials could never acquire a sufficient understanding of the sensibility (preferences) and behaviours of affected individuals.

On this account, each individual has an incentive to participate. Failing to do so leads to collective decisions that ignore their own preferences (sensibilities). This is an attractive potential solution to a significant potential problem, but it does not quite fit with Bentham's own account. Bentham did distinguish between "public" and "private" matters on the basis of how predictable individuals' response to particular sanctions could be from the legislator's perspective. He was also very clear, however, that the POT was to focus on public matters. These were common to all people; not just the (private) conduct of some individuals. Ironically, this implies that there may be no free-rider problem in relation to discussions of private conduct. But it does nothing to address the same issue in reference to the public questions that Bentham wished the POT to focus on.

Nonetheless, free-riding is not a fatal problem for the hybrid view. Bentham expected the "habit of reasoning" cited in "Political Tactics" to benefit other aspects of participants' lives. There is a distinct similarity between public reasoning in the POT and the type of instrumental reasoning that Bentham thought was key to individuals maximising their own welfare more generally. Maximising personal utility depended on having secure expectations on critical questions, and using these to make life plans in pursuit of one's own welfare. It is reasonable to suppose, therefore, that a public process through which citizens develop their ability to reason effectively should also enhance their ability to make effective life plans: to maximise their own utility. The personal benefits of first-hand participation are thus twofold. They are created through the security assured for one's own plans but *also* through an enhanced capacity to make and scrutinise those plans through effective instrumental reasoning. Whilst security against misrule might conceivably be enjoyed by a free-riding non-participant, the enhanced capacity to develop one's own life

plans and goals cannot be realised without participation. Benefits of these kinds are likely to be learned through active participation; not uncritical acquiescence and, or free-riding.

A habit of reasoning, honed through democratic participation, enhances the “utilitarian character” of citizens. It arms them with the ability to identify and scrutinise the connections between interests, information, language, and goals. In large part, maximising personal welfare depends on these understanding these connections and ideas: in one’s own life, as well as political life. This link can be seen in Bentham’s writing on education. The kind of public reasoning implied in “Political Tactics” continues themes from Bentham’s writing on education. In “Chrestomathia (1983a), Bentham sought to reform education to equip individuals to understand and pursue their utilitarian interests better. Public reasoning in a democracy complements that kind of education. It helps citizens to practice the critical faculties needed to make publicity effective as a security against misrule. Those are the same faculties required to understand and maximise utilitarian interests on a personal level. At the individual level, maximising utility takes account of personal circumstances and preferences. At the legislative level, it concerns more general trends. Planning one’s own life was thus a specific example of wider utilitarian reasoning. In IPML, Bentham describes the distinction between these two forms of decision-making. The difference between political and individual reasoning is one of scale, and available data: they are extensions of one another, rather than of a different types of thinking:

“It is plain, that of individuals the legislator can know nothing: concerning those points of conduct which depend upon the particular circumstances of each individual, it is plain, therefore, that he can determine nothing to advantage. It is only with respect to those broad lines of conduct in which all persons... may be in a way to engage, that he can have any preference for interfering” (Bentham, in Rosen, 2003:93)

Public reasoning can therefore be said to create better “private” reasoning. It uses the same skills; it thus confers some benefits which can only be achieved through participation, and not through free-riding.

To summarise, the free rider problem relies on the possibility that individuals may enjoy the benefits of a shared or collective good without having to contribute towards its creation. I do not think that this is a telling objection to the hybrid view. Bentham was clear that a major benefit of participation in democratic discussion and reasoning was its ability to create shared goods which benefit all, like security. He also suggested that there were additional reasons to participate. These include the capacity of public reasoning to enhance the character of those involved, and the personal benefits that these enhancements could

confer. Bentham's idea of self-interest also included the appreciation of benefits to others. He thought that people could be moved to act by sympathy. That sentiment could also reduce the potential for free-riding. The status of the hybrid reading as a self-interest-based view of public reasoning, therefore, does not preclude the participation of individuals for sympathetic reasons. Nonetheless, the scale of scrutiny required did mean that some members of the public would have to rely on the work of others. This "committee" of the POT creates some additional challenges for the hybrid reading.

The second objection that we shall examine relates to how the POT should function, and its susceptibility to sinister interest. That idea, as we have seen, was at the heart of Bentham's democratic theory. It explained why he came to believe that an enlightened legislator could not be trusted to maximise public welfare without a system of democratic accountability to regulate his conduct. Sinister interest also influenced the form that Bentham believed democratic government should take. His objection to a constitutional separation of powers, for instance, was motivated by a fear that granting constitutional vetoes to small groups within political elites would allow them to be exploited for sinister ends (Rosen, 2003: 232).

Whilst free-riding is not a fatal problem for Bentham's view of public reasoning, I do not believe that Bentham adequately accounted for how sinister interest should be mitigated in the POT. There is a particular issue here. Bentham did not consider sufficiently the danger that editors and journalists who helped run the POT could themselves be heavily influenced by the interests of an influential minority, rather than the majority. This is a weakness of his account. He did not acknowledge, in sufficient detail, that the supply of news would itself need public scrutiny and regulation to perform the role ascribed to it. Similarly, I believe that the demand for news may often reflect factors other than the epistemic merits of its content; for example entertainment value. This means that unregulated competition cannot be relied on to provide news sources with an incentive to scrutinise public officials, as Bentham wished them to.

These challenges are surmountable. It may be that relatively uncontentious regulatory intervention can overcome them. In a modern context, such interventions could include: independent regulation; a public code of conduct for news sources; public service obligations like those imposed on some UK broadcasters; or public provision of some content to mitigate concerns around commercial bias. The problem for the hybrid reading is not that solutions like these are difficult to formulate; it is that Bentham himself did not seem to tackle the problem with sufficient application.

It is clear that Bentham did not expect sinister interest to be anything like as much of a problem within the POT as it had been in public office. Public opinion was to be different from the closed groups of elites that had made decisions in the past. The POT was to be open to any person with a relevant interest, and would hold officials to account through public scrutiny. The structure of interests in society also meant that public opinion—by weight of numbers—was less liable to sinister ends. Bentham believed that there was a broad harmony of interests among the general population, and a huge disharmony of interests between the population and their governing elite (Bentham, 1989:68-9).

His belief in the opposition of interests between classes thus meant that he could expect most members of the POT to have much in common when considering how political power should be used. They all shared an interest in preventing abuses of power, and assuring security for their own life plans. Consequently, “ordinary” members of the POT would have little incentive to misdirect the Tribunal for sinister ends. They were thus different from members of the governing elite who had previously controlled public decisions. Unlike those officials, POT members would be likely to have particular interests which coincided with the public interest. Individual POT members would also have far less power than corrupt officials had done in the past. As one of a very large group, each individual Tribunal member would have a small share of control over the decisions. These characteristics contributed to what Bentham saw as the POT’s “incorruptibility”. As Niesen (2011) argues, this was part of the reason why Bentham believed that dynamic, unconstrained public opinion was necessary to the accountability of public representatives. Public opinion was—by its nature—less susceptible to sinister manipulation than private discussion within small groups of influential persons.

That said, the Public Opinion Tribunal was not simply an abstract aspiration. It was intended to operate in real societies with real people. Bentham spelt out some detail regarding the structure that it would take. In the Constitutional Code, for example, he explained that:

“persons considered as members of this tribunal are an indeterminate portion of the whole number of those of whom the community in question is composed. Those by whom actual cognizance is taken of the matter in question in the first instance may be considered as a Committee: those who in consequence of the opinions expressed by this same Committee, but without taking actual and particular cognizance of the circumstances of the case, join with them at different times in the same opinions, affections, wishes, designs and endeavours, constitute the body at large of which the smaller body abovementioned is the Committee. Of the Members of

this larger body, the number may be of any magnitude not exceeding the sum total of the adequately adult members of the community.”

(Bentham, 1990:121)

Whilst the POT was open to all, it would also need a committee, which took “actual cognizance” of the relevant question, to function in practice. This committee is an important component of how the POT could function: in effect, it could lead public opinion. We may therefore question how members of the committee could be expected to be immune to sinister influences. Bentham suggested that the committee could take different forms in different places. Where possible, newspapers would provide a major contribution: they would unearth and publicise evidence of misrule by public officials and print related opinions. Members of the public, in their capacity as POT members, could then exercise their judgement and express their views through the sanctions available to them, including the social and political sanctions (Bentham, 1990:124). There are some important tacit assumptions in this account. Optimistically, we might expect—as Bentham did—members of the public to respond rationally to evidence of misrule when presented with it, and punish wrongdoing. In order to do so, however, they require timely and accurate notification of officials’ malpractice, among other things. For that, they would rely heavily on the Tribunal’s committee and its members. Bentham’s view seems to assume that those committee members would perform their role with aptitude and probity; in the public interest. Without some specific governance or regulatory measures in place, however, we do not have grounds to make that assumption.

To understand why, consider Bentham’s analysis of sinister interests among public officials. As we have seen, that emerged from his earlier notions of probity, and the later concept of official aptitude. His argument was not that officials were susceptible to sinister influence because they had a particular or unusual weaknesses of character. They were not in possession of a uniquely or peculiarly bad character, which was absent elsewhere in the population. Officials were subject to the same motivations and behaviours as other citizens. What made them different was their position, which demanded that their actions should serve the whole of the community. This gave them, unlike the majority in society, sufficient power and influence to pursue their own interests at the expense of the community at large.

Consequently, officials were constantly subject to diverging public and private interests.³⁴ It was this divergence that made publicity for public officials essential, in order to bring their private and public interests closer together: to induce an artificial identification. Doing so would lessen the temptation to abuse public office which they, uniquely, were exposed to. Given the importance of that argument in Bentham's political theory, we might ask: "why should members of the POT committee be treated differently from public officials?" Committee members do not, like public officials, directly determine public policy. Consequently, we may hold that the danger of sinister influence among them is reduced, because they lack direct control over political outcomes through legislation.

Nonetheless, Bentham's principles of public reasoning depend on committee members as key figures. Without them, the public would not have essential information it needed to hold public officials to account: it could be unaware of wrongdoing. Committee members would effectively have the power to provide cover for malpractice in government; to undermine or augment political careers and policies. Their judgements and decisions could have a significant effect on political outcomes. Consider, for example, how important the mere timing of a critical publication could be to how the political sanction is used (e.g. publishing critical content before or after a general election). Without relevant information *before* they vote, electors are powerless to punish abuses at the ballot box. Given the importance of editorial decisions like these, is it not likely that committee members will, like public officials, be subject to considerable incentives and temptations associated with such an important a position? Will they not, too, be subject to diverging private and public interests, which demand artificial identification?

We have seen that Rosen alludes to this important issue in his comprehensive analysis of the "Constitutional Code" (1983:27-35). Bentham never quite explains how a middle class journalist or editor—with likely incentives which might often diverge from those of the majority of citizens—could be relied on to speak for the public interest. Of course, it follows from Bentham's other assumptions about motivation that these individuals would require a high degree of moral and intellectual aptitude. Like the public officials that Bentham had criticised, they would need to resist sinister influence as well as exercising effective judgement. He thus expected some from the middle class to play an important role as journalists and editors of newspapers to put public opinion into operation. The problem is that while Bentham provided extensive explanation of the mechanisms needed to assure the aptitude of public officials, he provided no equivalent for POT committee members. The kind of publicity that Bentham insisted on for intra-parliamentary procedures might produce an identification of interests on behalf of public officials, but it is not quite so clear how those actually involved in the POT itself might also be

³⁴ See Chapter 3 in the Book of Fallacies, quoted above on pp. 123-4

encouraged to display moral aptitude needed to perform their role. It could be said that Bentham's account is, therefore, dangerously reliant on a relatively small number of journalists and editors; but provides no sufficient account of how their conduct should be governed.

We could defend Bentham against this criticism: perhaps the POT itself provides an accountability mechanism for the conduct of members of its committee? Without some exogenous source of judgements necessary to operating that accountability, however, this response would suggest infinite regress. After all, some agent will also be needed to hold accountable those who scrutinise the conduct of POT committee members. Similarly, any attempt to claim that POT committee members have a different sort of character from others; that they are—by nature—less susceptible to sinister influence, would run counter to Bentham's fundamental analysis of human behaviour.

The contrast between the POT committee and other public officials is telling. Without the right institutions in place, Bentham thought it would be naïve to simply expect all public officials to serve the public interest; for example, through a sense of benevolence. In creating a Tribunal which fundamentally relies on members of the press, however, he seems to make an analogous mistake. On Bentham's logic, the composition of the POT calls for an artificial identification of interests between key members of the Committee, and the public at large; but no explicit mechanism is described to achieve that identification. There is, therefore, a “governance gap” in the hybrid reading of Bentham on public reasoning.

What responses to this objection might we consider from Bentham's work? Well, it is worth remembering that the POT does not always need to operate through established newspapers; although Bentham seemed to wish it to do so. While the “Constitutional Code” gives journalists and editors the pivotal role described above, Bentham did consider alternatives. For instance, Kaino has observed that Bentham wished the POT to operate in countries without a developed press (2008:21). He acknowledged, for example, that in Tripoli, “free Newspapers [were] the matchless instruments of notification” but were not present in that nation (Bentham 1990: 129 in Kaino, 2008:21). In places like this, Bentham considered the role that other institutions could play in the absence of a developed free press. For Tripoli, this meant attempting “to use the two universities, the 14 judicatories and the 3,000 mosques for the purpose of notification” rather than the press (Kaino, 2008:21). On one level, this illustrates Bentham's pragmatism. Whilst insistent that a free press was the best route toward using the POT to provide a security against misrule; he also wanted to harness public opinion in territories where that was not possible. On another level, it also demonstrates some of the weaknesses in Bentham's account. He was elsewhere, of course, highly critical of religious

authority and its propensity to benefit those who ran religious institutions. Bentham suggested that sinister interest was a systemic problem throughout the Church of England, for example. The idea that religious authorities can be trusted to notify the public of misrule—without an artificial incentive to do so—is therefore at odds with Bentham’s other views. For the hybrid reading of Bentham’s to remain consistent, therefore, it is important that the means by which citizens are notified of misrule—whether that be a free press, universities or mosques—are themselves governed in a way which insulates against sinister influence. A failure to do so would demand that we assume, by fiat, that members of the Committee of the POT are less susceptible to sinister action than others. The absence of a practical account of how the POT is governed is thus a weakness in the hybrid account; despite Bentham’s pragmatism about the role that alternatives to a free press could play.

Characteristically, Bentham foresaw some of these problems. He acknowledged the tendency of some newspapers to become agents of political parties, and of readers to select publications that reinforced their existing views (Bentham, 1990:47). He also believed that readers would naturally take an interest in content which affected them (Ibid: 47). Consequently, he expected newspapers that covered a variety of content to gain more readers than biased and narrow polemics. It is possible that the relative inattention Bentham paid to the governance of the POT thus reflects a wider assumption: that sinister arguments will be rejected in a competitive marketplace of ideas. One could argue that competition between newspapers should give them incentives to scrutinise government. It should reward effective scrutiny with increased sales, and punish poor journalism with commercial decline. After all, if citizens have such clear need for the scrutiny of a free press, should they not seek it out?

Unfortunately, I do not think that experience of the demand for, or supply of, news content justifies such optimism. In the modern world, we know that decisions to pay attention to, or consume, some media content over available alternatives reflect a range of factors. As we might expect, the entertainment value of content is—and has long been—a significant factor, for many people. The breadth of media publications that Bentham predicted is no guarantee that readers will pay attention to content of national importance—including notifications of misrule (Cutler, 1999:349)—rather than stories of more superficial interest, but less political importance. Consider recent experience in the United Kingdom. The UK broadcasting regulator OFCOM has found that rapidly growing sources like online social media are used by a significant minority of consumers to learn about celebrity news rather than items which traditionally lead broadcast and print media (OFCEM, 2015:7).

This could be seen as the social sanction at work. After all, readers approve or disapprove of persons and their conduct, as well as government policies. It is difficult to claim that such media usage empowers publicity as a security against misrule by public officials, however. In particular, there is no clear correlation between the significance of news to the public interest and the degree of scrutiny and attention it receives from consumers of new media. Realities like these undermine the degree to which we can expect a *laissez-faire* approach to ensure that the press, and wider media, play the public-spirited role Bentham wished them to. If citizens do not pay attention when an instance of misrule is exposed, they will not be able to secure themselves against its effects by preventing a reoccurrence.

The recent supply of news has provided some important examples of “public interest journalism” exposing alleged malpractice by public officials. One obvious example is the revelation of MPs’ expense claims before the 2010 General Election. Eggers & Fisher (2011) have found evidence that this coverage did provoke voters to punish those misbehaving MPs who had been exposed by the media. Bentham might well see this as an example of the Public Opinion Tribunal at work; provoking punishment via the political sanction. The picture is not always this simple, however. A look at the current UK media landscape should serve to emphasise that the supply of news is complex. It is often heavily influenced by a variety of commercial factors. Bentham would doubtless be unsurprised to see commercial newspapers trying to maximise sales, rather than selflessly providing political scrutiny.

In 2014, over 70% of UK national newspaper circulation was controlled by just three companies according to the Media Reform Coalition (MRC, 2014:1). Each of these companies will therefore be likely to have significant influence on the way political debate takes place. Each will, by its nature, also have a particular set of commercial and reputational incentives. Some of these companies are controlled by organisations with a wider portfolio of investments and interests; or small groups of wealthy individuals. There is no reason to suppose that those individuals or organisations should be immune to the kinds of sinister interests that Bentham thought could blight public decision-making. We cannot rely on their altruism. We cannot rely on competition to assure that the press provides effective scrutiny, because there is no evidence that political scrutiny is a consistently effective commercial strategy. My point here is not that those incentives are unmanageable, or that the barriers to an effective press cannot be overcome. Potential conflicts of interest within the press do, however, need to be acknowledged and addressed. An account of public reasoning which relies so heavily on the press to scrutinise the conduct of public officials, should also address the incentives to which POT committee members are subject; including those who run and work for newspapers. Just as we might think it naive to assume that a middle class journalist could speak for the public interest in a way that public officials were apparently incapable of, it

is also problematic to downplay the commercial incentives to which newspaper proprietors are inevitably, and regularly, subject.

There are several potential responses to this issue. Bentham believed that exposing the deliberations and decisions of public officials to appropriate publicity and scrutiny could overcome, or at least control, their tendency to favour sinister interests. There is no reason why an analogous regulatory system for press ownership should be impossible. Bentham was concerned about balance in newspapers, and suggested that publications could alternate between editors of opposite political views. This was intended to assure something close balance in their coverage. Whilst the issue of sinister interests in the organs of the Public Opinion Tribunal does not seem like an irresolvable one, therefore, it is one which demands greater consideration than Bentham himself afforded, in the “Constitutional Code”. This is perhaps a consequence of how Bentham thought about the POT. He considered it a protection against the harm that corrupt or incompetent governance could create. In doing so, he paid great attention to the misinformation and manipulation of which governing elites and public officials could make use. In comparison, he paid relatively little attention to the parallel—though lesser—risk of corruption or incompetence among those charged with making the POT a day-to-day reality. Clearly, the dangers related to sinister influence within the press were of less concern than those within government. Nonetheless, they require a fuller answer than Bentham afforded them.

Let us turn to another objection; this time in regard to Bentham’s expectations of what public opinion would achieve. Just as Bentham believed that a free press could hold government to account, he also thought that public opinion would, over time, become more and more enlightened. It was partly for this reason that the Tribunal could be trusted to represent the public interest. I believe that there is an unexplained gap between Bentham’s optimism for the future of public opinion, and his awareness of the vulnerabilities and weaknesses of the individual judgments that formed it. Schofield has observed the grounds for this optimism was Bentham’s belief that suppression of the truth took more time and energy than its revelation: “It is the characteristic of error to possess only accidental existence, which may terminate in a moment, whilst truth is indestructible” (Bentham, in Schofield, 2006:267). This meant that laws motivated by sinister interest could not last forever; unlike those founded on the happiness of the community. Eventually, the errors on which they were based would be exposed. At that time, the truth—which did reflect the public interest—would still be available for citizens and representatives to discover.

Thus, while in the past there had often been a majority in parliament to pass laws which harmed the public interest, often in pursuit of sinister interests, Bentham also thought that these laws could not continue indefinitely. They would be quickly overturned when their fallacious logical foundations were exposed. Without the protection of a self-serving parliamentary majority with an incentive to ignore the truth, public opinion was sufficient in “a single day...to unravel the web of corruption, and devote the corruptionists, if not to punishment under forms of law, at any rate to universal indignation and abhorrence, with a certainty of never more being reappointed” (Bowring, iii 456 in *Ibid*). Political actions based on sinister interests were therefore vulnerable in a way that those with a proper utilitarian justification could not be. They could survive only under the cover of a particular set of particular interests; and would fall apart under scrutiny. This could lead us to conclude, as Bentham seems to at points, that the progress of public opinion is inevitable.

Taken literally, it is difficult to reconcile this picture –of public opinion’s seemingly inevitable march towards truth—with the subtlety of Bentham’s analysis of deception and sinister interest. Clearly, in other works, Bentham demonstrates a sophisticated understanding of the influence of powerful minorities over public decisions, and the terms of debate. In the “Deontology”, for instance, he describes something similar to what would now be understood as “selection bias”. Judgment could be swayed by pecuniary interests, as had been the case in the jury system.³⁵ He wished to mitigate that problem through publicity. He was aware of just how easily ordinary members of the public could be led to have an inaccurate conception of their own interests, and provided systematic analysis of how this had been exploited by minorities with putative authority. As we have seen, Bentham also had radical views about the capacity of language to provoke inaccurate conceptions of interests, and the need for its radical reform. These risks demanded a critical and active public to guard against them.

Given all of these factors, it is difficult to see how we could be confident that the deceptions and delusions built up over decades or centuries could be brushed aside quickly; or with any certainty. We know that Bentham thought in great depth about how politicians should be given incentives to serve the public interest. It would be naïve, however, to suppose that those reforms could eradicate the effects of decades of deception and corruption in a single day. Of course, the quotation above may just be an example of Bentham using a more figurative mode of writing. Nonetheless, it creates a misleading impression. Bentham’s analysis of deception and sinister interest suggests that unwinding the pervasive influence of powerful elites is likely to be difficult and uncertain; not inevitable or quick. This optimism thus creates an inconsistency. It is also difficult to square with contemporary research on deliberation. Indeed,

³⁵ See the “Essay on the Art of Packing”, 1821.

theorists like Fishkin (2008) have shown that a good deal of time, effort, information, and facilitation is often necessary to safeguard public opinion against the kind of ill-informed conclusions that all democratic theorists wish to avoid. Without this necessary work, there is a constant danger that the conclusions of deliberation could err in any number of ways; including selective use of evidence, prejudice and behavioural bias.

Bentham himself suggests that producing non-manipulated decisions and opinions requires careful planning, effort, and co-ordination. Texts like “Political Tactics” attend to this planning in some detail. There, as elsewhere, Bentham’s account suggests that a failure to provide the right structure and process for decision-making could easily undermine the value of pre-voting discussion. I find those arguments compelling: they acknowledge some of the many pitfalls which can befall discussion, and the need to overcome them in to produce justified decisions. Indeed, that line of argument follows from Bentham’s critique of delusion and corruption. It is necessary to attend to the procedural conditions and details of discussion *because* human judgement is vulnerable to corruption and delusion.

The optimism cited above, however, does not follow this line of work. Given the complexities of the incentives involved in the generation of public opinion in contemporary democracies—and the efforts needed to promote effective deliberation—I do not think we have strong reasons to agree with Bentham that the onward march of public opinion toward greater enlightenment is a certainty. Clearly, there is a sense in which public opinion, as Bentham defines it, is immune to the corruption and error which had beset public officials. Those individuals were placed in situations where their particular interests consistently diverged from the public interest. Public opinion was different. It reflected views which were common to large numbers of people in the community; not a small group with particular circumstances and preferences. Despite this, the opposition between Bentham’s account of deception, and his hopes for the future of public opinion is too binary to reflect these nuances. When considering the problems which had encumbered attempts at reform in the past, he was aware of significant challenges. These included individuals’ capacity to be deceived, the danger of fallacies and delusions, and the sinister intent of a governing few to exploit them. In contrast, when looking to the future, Bentham seems to expect public opinion to progress in a way which belies the challenges he had identified (Bentham, 1983b:36).

It would be reassuring to believe that the truth is somehow “indestructible”, and that centuries of deception can be quickly unwound. Bentham’s own analysis of language and sinister interest, however, run counter to this. They suggest that progress may be won slowly, through concerted clarification of

language and repeated presentation of reasons; not quick revelation. Of course, Bentham did not discard his awareness of the faults which could hold back public opinion when he gave it a central place in his democratic theory. Significant among these was of “sinister begotten prejudice”. This caused a mass of citizens to support policies directed towards sinister interests rather than the public good (Rosen, 2003: 243). Nonetheless, it is difficult to reconcile ideas like sinister begotten prejudice with some of the more optimistic claims Bentham made about the inevitability of the progress of public opinion. As a critical force in society, Bentham wanted public opinion to speak for the public interest rather than merely reflecting mass opinion (Ibid). His suggestion that it could quickly or faultlessly, eradicate errors of the past, is not however convincing. Rather, his epistemology suggests that constant work will be needed to ensure progress; particularly to assure that citizens have access to the information and opportunities they need to exercise their own judgement without corruptive or delusory influence.

Let us turn to another objection. In “On Liberty”, Mill famously insisted that individuals should not be protected against things which offend them but cause no actual harm (Scarre, 2007:117). Despite this, he also suggested that there are many acts which “being directly injurious to agents themselves, ought not to be legally interdicted but which, if done publicly, are a violation of good manners...[thus] may rightly be prohibited” (Mill in Ibid:116). This raises the question: how should utilitarians respond to activities like these? In particular, how—if at all—should the preferences that some citizens hold in regard to these activities be considered?

More recently, Dworkin (1978) has presented that question as a critique of utilitarianism; based on “external preferences”. He gives the example of a society where some citizens have prejudiced views against homosexuality. Where “external” preferences are expressed regarding the behaviour of others in this way, Dworkin contends that utilitarians may respond in one of two ways. The unpalatable results that aggregating bigoted preferences might produce could be avoided by simply excluding them from the preferences which count according to a utilitarian calculus. Of course, the problem with this approach is that such an exclusion might well be arbitrary. Dworkin argues that there are no clear utilitarian grounds on which the fulfilment of external preferences can be said to be less valuable than that of other, more “palatable” preferences.

Alternatively, utilitarians may include external preferences on a level footing with other preferences in their calculus. In the kind of society Dworkin describes, this would have obvious counterintuitive results. The repression of minorities could be deemed desirable if a sufficiently large number of citizens had

negative external preferences towards them. Dworkin thus suggests that contemporary utilitarians are left with two unsatisfactory options. They can either achieve consistency towards all preferences and accept unpalatable results; or they may exclude certain external preferences by fiat. The latter approach is implied by Mill's suggestion that offence should not constitute harm. The former approach is suggested by the quotation above on "violations of good manners".

How might this critique apply to the hybrid account? Of course, Bentham was optimistic that public opinion would progress towards conformity with the principle of utility. He espoused a liberal response to what he called "sexual irregularities"; including homosexuality, bestiality and necrophilia. Whilst important, Bentham's views on those substantive questions do not answer the challenge presented. Would Bentham himself have excluded or accommodated external preferences? This is an interesting question for utilitarianism in general; but this challenge simply does not apply to the POT as it is presented in the first volume of the Constitutional Code (1983b). Bentham is clear that the POT should consider public rather than private matters. He defined these as questions which affected the interests of all in society rather than a particular group. The POT would not therefore consider private matters such as the desirability of different sexual preferences.

This is, however, a limited response. It implies that the POT should not pass judgement on the desirability of different preferences which are pursued in private. It says nothing, however, about other important questions which relate to the role of sexual preferences in the public realm. For instance, debates on same-sex marriage relate to the public recognition of relationships rather than merely their private exercise. It is conceivable that public acts like these could be considered "offences against decency" on the definition suggested by Mill in Chapter 5 of "On Liberty". Here, it is worth emphasising an important difference between Bentham and Mill. We have seen that Bentham believed that notions of taste had been exploited by ruling elites to manipulate citizens' behaviour. He therefore understood taste in this sceptical light: it had the potential to deceive individuals and impede progress. This strand of Bentham's thought complements his emphasis on the secret ballot, and the idea that public reasoning must be a critical process directed toward the self-interest of participants. It also complements Bentham's quantitative hedonism.

Mill took a different approach. He famously suggested that utility could differ in quality, as well as quantity. Consequently, he was more willing to accept that the preferences held by some citizens could be qualitatively superior to those held by others. The defensibility of that claim has been debated at length by

Mill scholars (West, 2004; Riley, 2010). It is not our immediate focus. Unlike Bentham's argument, however, a qualitative utilitarian view does leave citizens open to accusations that they have "the wrong sort" of preferences. Bentham's arguments on taste are both interesting and compelling. They suggest that the very notion of "superior" preferences is itself susceptible to exploitation by elites. I think that it is an advantage of Bentham's account that it seeks to explain the value of pre-voting discussion while retaining a quantitative utilitarian approach. In other respects, Bentham's response is, however, limited. The scope of the POT being limited to public matters addresses the potential problem of citizens having "nosy preferences" about each other's lives. It does little, however, to address related contemporary questions. These include the role of same-sex relationships in public institutions like marriage.

More broadly, some of Bentham's other assumptions are also difficult to apply to a contemporary context. It might be objected that the hybrid reading is dependent on a questionable assertion that individuals' self-interest can be reliably aggregated to reach the public interest. Whilst Bentham did make this assertion in texts, most famously in IPML, I do not think that this is essential to the hybrid view of public reasoning. Rather, he believed that there was little need for members of the general public to consider public interest, rather than self-interest. This was because the division between two classes in society meant that there was significant harmony between the interests of members of the Democratical Section. Consequently, Bentham provides reasons beyond his aggregative definition of the public interest to suggest that self-interested reasoning was required from the general public in regard to political decisions. Bentham's assumptions about class are, however, more problematic. Whilst it might be true that this is a powerful way to understand the division of interests in society, the picture he paints of a simple division between an opulent and dominant ruling few and a powerless many seems too simple to accommodate contemporary developments. For instance, within the Democratical Section there are, today, very significant differences in circumstances. These might include distinctions between those who rent or own property; are in manual or office-based professions; live in one region of a nation rather than another.

Recent social movements like "Occupy" have drawn a similar distinction between the richest 1% and remaining 99% of global population. Even if we do grant that this is a fundamental division of interests in society, it is difficult to imagine the kind of harmony that Bentham anticipates within the "Democratical" 99% of the population. Bentham's critique was perhaps well-suited to analysing the barriers to fundamental democratic reform that he encountered. But it is difficult to envisage how it might apply to some everyday decisions that governments encounter today, on which there are many different varieties of opinion across multiple classes.

One way to counter this objection might be to limit the scope of Benthamite public reasoning to those issues on which there is most likely to be a fundamental opposition of interests between a very small, powerful elite and the mass population. Bentham himself did not define a scope in this way. He believed that considering the pleasures and pains which would be created by an action was universally relevant. Nonetheless, we might reason that there are some issues to which the fundamental opposition Bentham supposes apply more readily. Most particularly, these could include the rules which directly regulate how political power is used: constitutional rules that regulate the actions of public officials. Bentham was clear that these kinds of rules should be subject to the influence of public opinion on an ongoing basis. No constitutional authors were infallible. Conditions in which constitutional procedures translate into effects on citizens' welfare inevitably change over time in a way which is not always predictable. For these reasons, public reasoning might understandably be limited in scope to the kinds of constitutional essentials on which Rawls later focused. This would, however, be a departure from the principles Bentham himself espoused.

Challenges to Benthamites

Let us now consider how a broader hybrid Benthamite view might contribute to the contemporary debate. A view of this sort would follow the broad principles set out in the hybrid reading. It may however provide conclusions that Bentham himself did not express. How does a hybrid Benthamite compare to contemporary alternatives? In Chapter 1, I sketched out five dimensions over which contemporary accounts tend to vary. These refer to the: focus, transformative power, status, and output attributed to pre-voting discussion. To understand the answer to this question, let us briefly recap the terrain that contemporary views inhabit.

Prominent among these is Habermas, who advocates an ideal form of deliberation. In this situation, participants move beyond strategic action, aimed at success, to communicative action: the cooperative search for mutual understanding. Doing so involves realising a type of discursive equality, which gives every participant an equal and fair input into deliberation. Its aim is a consensus which is justifiable in light of the generalizable interests of all relevant parties (Habermas, 1990:89, 1996: 305-306, 308, 2001: 34). As I have argued in response to Crimmins (2011), it would be a mistake to underplay the important differences between Bentham and Habermasian deliberation. Principal among these is the orientation of debate; which must transcend self-interest on the Habermasian model, but genuinely reflect it according to Bentham's views on secret ballot and the POT. The hybrid Benthamite view has a focus on self-interest rather than the public interest. This view suggests that the Public Opinion Tribunal, thus public opinion, is directed primarily towards censure of public officials. This task—spotting bad practice by officials—

requires individuals to understand that their behaviour is “bad for me”. It does not require them to consider whether it is “bad for society”, or unjustifiable according to some wider standard.

Nonetheless, there are some interesting similarities between hybrid Benthamite and Habermasian deliberative views. Bentham wished to eradicate the influence of will on will; whether direct (corruption), or through the understanding (delusion). He did not want to eradicate the influence of understanding on understanding, which we might term persuasion. In this respect, the rejection of voters’ independence of one another before voting in “Political Tactics” is consistent with Bentham’s writing on corruption and delusion. At the broadest level, the hybrid Benthamite enterprise resembles that of Habermas. The aim is to allow persuasion based on legitimate reasons rather than the threat of force or misunderstanding. It wishes participants to learn from one another before voting. But this is a different form of learning: the intent is to clarify understanding of existing interests rather than to transform them. Whilst Habermasian deliberative democrats espouse the transformative power of pre-voting discussion, the hybrid Benthamite view does not. This complements principles behind Bentham’s rejection of the public ballot. In regard to transformative power, the benefits ascribed to public reasoning, for example in “Political Tactics”, are understood in terms of correcting individuals’ conception of interests; not preference change. This complements Bentham’s belief that sinister interests were a leading threat to decision-making and, that their negation was a driving force behind the need for democratic reform.

The hybrid Benthamite account thus casts pre-voting discussion as something with important instrumental benefits. It allows citizens to practice their ability to reason instrumentally and critically. It helps protect against sinister abuses of power, and aims to give individuals a more secure basis on which to plan their own lives, and maximize their own welfare. This is different from many contemporary deliberative accounts. D’Agostino, (1996:57), for example, describes pre-voting discussion in terms of its ability to justify the exercise of political power. The hybrid Benthamite view considers the instrumental benefits of discussion, but does not write public reasoning into a wider justification of the state. Its role, through public aptitude, is to promote official aptitude, which benefits the community.

Finally, contemporary views vary in regard to the outputs they expect to be created by discussion. There is little reason to suppose that Bentham required the kind of “stronger deliberative” consensus espoused by contemporary theorists like Habermas. Individuals’ interests may not overlap in a consensus, particularly not for the same reasons. As a consequence, the hybrid Benthamite view is content to allow persistent

disagreement; so long as this is based on an accurate reflection of relevant interests rather than delusory or corrupted discourse.

In Chapter 1, I outlined some contemporary theories of public reasoning. The hybrid view shares significant characteristics with “weak deliberative” accounts. It values pre-voting discussion. Unlike many views, however, its focus is self-interest, and it does not expect anything like full consensus. What does this tell us about the contemporary debate? The hybrid Benthamite view illustrates how significant the common ground between aggregative and deliberative views is. Clearly, this construal of Bentham sits neither with “rational choice” aggregative theories which give pre-voting discussion no significant role, nor with the “stronger deliberative” views of Habermasians. It demonstrates the preference transformation and public orientation are not essential to providing an account of deliberation. It also demonstrates that aggregative accounts need not ignore the conditions of preference formation, as some deliberative theorists suggest. Seeing where a hybrid Benthamite view fits in relation to contemporary alternatives should also help us to identify some of its strengths and weaknesses. To do this, I will consider three challenges, and potential responses to them. These are: the accusation that it reinforces existing power structures; does not understand moral conflict, and cannot conceptualise manipulation.

I believe that a hybrid Benthamite view of public reasoning is well positioned to counter two of these challenges, but less well placed to respond to the way in which contemporary theorists seek to understand moral conflict. Consider the first objection. Gutmann & Thompson (2004) have been among the leading proponents of deliberative democracy in contemporary political theory. In previous chapters, I have challenged the way in which they have sometimes referred to Bentham as an archetypally aggregative democrat. References like those are often difficult to reconcile with the democratic arguments that Bentham himself proposed.

Some aspects of the deliberative critique of aggregative theory proposed by Gutmann & Thompson do raise more difficult challenges for the hybrid Benthamite view. Consider, for instance “Why Deliberative Democracy?”. This text suggests that aggregative democratic theories accept existing power structures in a way which is difficult to justify. This would be a particularly troubling accusation for a Benthamite account of public reasoning. After all, the later Bentham fundamentally believed that democratic participation was an essential source of security against sinister abuses made possible by the existing distribution of power in society.

Gutmann & Thompson claim that it is an essential feature of aggregative democratic approaches to take “existing or minimally corrected preferences as given, as the base line for collective decisions” (Gutmann & Thompson, 2004:16). They deem this acceptance of existing preferences to be fundamentally conservative. Those preferences are formed and informed in relation to a background shaped by the way society has been; not how it could, or should, be. Consequently, by accepting the satisfaction of existing preferences as an appropriate criterion of right, aggregative theories “accept and may even reinforce existing distributions of power in society” (Ibid).

This claim aligns with some arguments we have encountered before. There are, for example, similarities between this deliberative view and some feminist critiques of utilitarianism. One such critique was examined in Chapter 1. Like Gutmann & Thompson, Annas (1977) also suggests that existing preferences are an inappropriate baseline for collective decisions because they unduly reflect existing power structures. In particular, Annas’ critique of Millian utilitarianism associates those power structures with patriarchy. On her view, Millian utilitarianism was susceptible to undue conservatism. Assessing outcomes in reference to existing preferences renders utilitarians unable to step back and consider the preferences that citizens should, or would, have; given proper deliberative conditions. In contrast, deliberative democrats like Gutmann & Thompson claim that the appropriate conditions of deliberation help citizens to move beyond initial preferences, which may unthinkingly reflect patriarchal or other iniquitous power distributions. The implication is that decisions made in an aggregative way—without deliberation—allow unfair treatment of particular groups by the state and civil society to continue in perpetuity. Utilitarianism would thus supposedly have allowed and perpetuated the subjugation of women who expressed an apparent preference for their own denial of basic property and democratic rights. The expressed preferences of those women, alongside those of men in the community, would be sufficient to justify this state of affairs.

So while deliberative democracy corrects individuals’ misconceptions of their interests, aggregative theories are destined to amplify them to the level of collective decisions. The accusation of undue conservatism also has much in common a “conventional view” of Bentham, outlined in Chapter 2. That interpretation suggests that Bentham was content to aggregate, rather than correct, “raw” preferences. Contemporary theorists often imply a number of different conceptions of “raw preferences”. I outlined three of these in Chapter 2; prominent among these is the view that preferences are “raw” if they have not been formed in appropriate deliberative conditions.

As I hope has begun to be clear from the arguments we have examined, I do not believe that conservatism of this sort is a justified accusation against the hybrid Benthamite view. Bentham himself acknowledged individuals' capacity to hold erroneous conceptions of their own interests. Often, implicitly and explicitly, the implication of this critique is that a critical form of public opinion is required to correct these misconceptions. Public reasoning is a process which helps to operate and form this kind of public opinion. It attempts to correct the delusions which cause individuals to misunderstand their own interests. The fact that correcting these misconceptions does not amount to preference change does not make the hybrid Benthamite response any less valid; unless of course, we assume from the start that a deliberative view is the beginning and end of what public reasoning should be. As suggested above, deliberative theorists have not yet provided grounds for that conclusion. Gutmann & Thompson are right to suggest that preferences which have not been duly considered may reflect current pervasive power structures. The hybrid Benthamite view I have described accepts that premise, and the idea that public reasoning is needed to prevent it. It rejects the idea, however, that full deliberative preference change is required for that task. Rather, a form of public reasoning which focuses on clarifying citizens' self-interest, and the language used to describe factors on which it depends, also provides a plausible response to this challenge.

More broadly, the idea that Benthamite utilitarianism allows this kind of domination has, of course, been around for a long time. It is closely related to a well-known critique of Bentham's democratic theory: that his form of utilitarianism creates opportunity for a "tyranny of the majority". Mill famously used that phrase in his introduction to "Utilitarianism". It represents an idea that constitutional checks, like the separation of powers and a written constitution enshrining guaranteed individual rights, are needed to protect the security of minority groups. In many ways, however, Bentham was not a simple majoritarian. He had mixed opinions about the importance of majority rule in representative democracy. His rejection of the separation of powers was not based on the idea that the will of the majority should not be checked, but on a fear that allowing a minority veto would serve sinister interests (Rosen, 2003:232).

Rosen (2003: 232-244) has also shown that Bentham's objections to the oppression of groups did not depend on whether that group constituted a minority or majority of the total population. In this sense, Bentham was not simply content to allow the domination or exploitation of minorities in pursuit of the interests of a majority. He objected to what he saw as oppression in many different territories, whether the majority or minority was subject to it. Bentham thus cited the oppression of the Turks under the Greeks; the Helots under the Spartans; the Protestants in France under Catholic rule; the Catholics under Protestant rule in Ireland and "in the Anglo-American United States the Blacks under the Whites" (Bentham, in *Ibid*: 233).

In each case, Bentham believed that oppression was harmful to both the oppressed and the oppressors, because it undermined security; in which everyone had a personal stake. It was only on the basis of that security that individuals could make plans for their lives to maximise utility (Kelly, 1990:73). Conversely, arbitrary rules that picked out individuals or groups for ill treatment made conditions less predictable and secure for everybody, even if they were intended to serve the community.

The critically important role that security plays should certainly ameliorate concerns that Bentham's account of public reasoning could allow minority groups to be oppressed. As Rosen notes, this is a point on which Bentham disagreed with conclusions later reached by some Twentieth Century utilitarians, like Smart:

“For both Bentham and Mill, the disappointment of settled expectations (such as the sudden reduction of an individual from citizen to slave) creates such pain and is so fundamental an injustice that it can never be justified on grounds of the augmentation of happiness elsewhere”
(Rosen, 2003: 236)

Finally, the hybrid Benthamite account may be challenged in relation to moral conflict. Although it encourages citizens to question existing power structures, and encourages belief in a form of security held by all citizens, it also does not approach moral conflict in the same way that many contemporary theorists do. It is a strength of the hybrid Benthamite account that it does not expect the kind of full consensus demanded by Habermasian deliberation. Ironically, however, in doing so it also fails to acknowledge the kind of legitimate fundamental moral disagreement which is important to contemporary public reason(ing). It is outside the scope of this present work to decide between fundamentally different ethical theories. Nonetheless, we should note how essentially different Bentham's ethics were from those of contemporary public reasoning theorists. There are consequently some significant barriers to incorporating a Benthamite conception of public reasoning into the contemporary debate. The hybrid view is rooted in Bentham's writing on language and metaphysics and motivation. This has the benefit of meaning that we can adduce a view of public reasoning which works with Bentham's wider corpus, rather than contradicting his other views. This also means that those who endorse a different view of language or metaphysics may not be able to accept the conclusions of Bentham's view. For instance, many contemporary theories do not see knowledge relevant to fundamental questions as grounded in sensory experience of physical entities.

Additionally, many debates on public reasoning focus on the rules for considering fundamental moral disagreements on questions like abortion. The principles that Bentham espoused suggest that these might best be understood in terms of the pleasures and pains at stake, with individual citizens considering and clarifying their own interests in reference to information they collect through discussion. Rather than focusing on abstract moral rights, therefore, individuals should understand questions like these in terms of their own interests, and those of people directly affected. Bentham might well dismiss rights-based moral views on abortion or marriage as unjustified deontology: the projection of personal prejudices onto the canvass of others' behaviour. Consider Rawls' argument that arguments regarding "constitutional essentials and matters of basic justice" should reflect reasons which are justifiable to all reasonable persons (Rawls, 2005:214). This precludes arguments based on a particular comprehensive moral view which are not shared by other reasonable persons being grounds for arguments about voting on a question like abortion. For Rawls, this standard applies both to public officials carrying out their duties, and to members of the public when voting and deliberating on questions within the scope of public reason. In contrast, the principles Bentham suggested differentiate between these two audiences.

Extrapolating from Bentham's writing, we might conclude that there should be no a priori limits on the reasons which public officials should consider. Rather, they should be expected to pursue their duties with impartiality between the interests of all in the community, as suggested by the oath of office Bentham outlined in the "Constitutional Code". Bentham's conception of class entailed that different rules applied to the general public. They were expected to clarify the interests and information relevant to a particular decision and to engage on the basis of securing their own interests rather than in pursuit of an impartial regard for the interests of the whole community. A Benthamite citizen might therefore approach a question like abortion in this way; by considering how a change to the regulation around abortion would affect their own interests, including those of their family and friends. They might consider, for example, how the welfare of themselves and their children would be affected by a liberalisation of term limits. In doing so, they would be likely to question accepted notions of rights and abstract principles to test whether they relate to their own interests, or merely the personal, and often sinister, tastes of others.

For some this might show the limitations of a Benthamite view. Whilst contemporary theorists often attempt to conceptualise and accept implacable moral disagreement on questions like these, the Benthamite approach seeks their resolution solely by reference to the interests concerned. The starting point of Bentham's writing is thus inimical to much contemporary writing about public reason. Whilst Bentham's premises are unlike Rawlsian views, they do have more in common with epistemic democrats like Estlund (2008) and Talisse (2009). These thinkers conceive of the value of democratic participation in terms of the accuracy of the knowledge it produces. In line with these accounts, we might try to

understand the accuracy of beliefs and knowledge which feed into the democratic process in more general terms, rather than the explicitly physicalist and utilitarian terms proposed by Bentham. This would require something beyond an entirely subjectivist understanding of preferences. We have seen, however, that there is good reason to believe that Bentham himself was comfortable with that principle, given his extensive writing on delusion and corruption.

Let us now return to the question of manipulation, and a challenge for aggregative theories presented by Fishkin:

“...the definition of manipulation turns in part on the alternative of good conditions and good information... as a benchmark for comparison... These good conditions are in act, a good part of what I mean by deliberation” (Fishkin, 2011:33)

A theme throughout this thesis has been the avoidance of manipulated outcomes from pre-voting discussion. We have seen that Bentham was very aware of the ability of elites to manipulate individuals’ judgments, yet also extremely optimistic about the ability of public opinion to overcome that problem. Just as some critics allege that a utilitarian view of public reasoning is unduly conservative, others—for related reasons—suppose that aggregative theorists like Bentham cannot sufficiently explain the problem of manipulation in a democracy.

Fishkin has suggested that any democratic view needs to provide some account of deliberation in order to understand what manipulation of this sort is. In his view, manipulation is the flip-side of deliberation. It is what happens when individuals form preferences without the correct deliberative conditions. On this logic, it would be difficult to understand what manipulated agreement even is unless one can also define what, by contrast, effective deliberation would be. The importance of preference formation should, therefore, cause us to demand that “people get good information, have access to arguments on competing sides, and have the chance to weigh the merits of those arguments”. (Fishkin, 2011:37).

Those conditions are uncontroversial, but they are also very broad. Few of the theorists we have encountered would deny the importance of these factors as appropriate conditions for decision making before voting. That said, some aggregative views, and the conventional reading of Bentham, might deny the need to prescribe deliberative conditions for preference formation. Instead, they take preferences as fixed inputs for aggregation.

Since manipulation of outcomes (or at least the part Fishkin draws attention to) takes place during preference formation, a non-deliberative view (on the definition he implies) can, by definition, have nothing to say about how to define or manage this problem. It is an implication of Fishkin's argument that some account of the appropriate conditions for preference formation must be central to any view of how citizens should reason together before voting on questions which are important to the community.

The participatory and hybrid views of Bentham provide significant detail about the appropriate conditions for preference formation. The conventional view, in contrast, provides very little detail, and instead focuses on defining appropriate decision making mechanisms for dealing with established preferences. It is a weakness, therefore, of the conventional view that it says comparatively little about the appropriate conditions under which citizens should consider how to vote before voting: broadly, it is a non-deliberative view on Fishkin's definition. The intent to "have something to say about what constitutes a reasonable... context for the formation of preferences" is a central feature of contemporary deliberative democracy (Christiano, 2002:32). Often, contemporary democrats provide an account of preference formation in concert with a move beyond voting on mere self-interest. This reflects an expectation that people should vote according to something beyond self-interest.

With regard to the information provided to voters, there are two distinct levels of manipulation which might occur. First, there might be malign influence on the way in which voters form their preferences, which causes them to reflect something other than an informed and considered view of relevant information. Second, it is possible that incorrect information may cause individuals to draw incorrect conclusions *after* their preferences have been formed. The implication of Fishkin's argument is that a non-deliberative view may be able to define and counter manipulation on the second of these levels. Since it assumes that individuals have appropriate information on which to make choices, however, it cannot have enough to say about the former, deeper level of manipulation.

One prominent response to that challenge is provided by deliberationists who outline discursive conditions that enable or provoke the type of learning which is required of citizens in a relevant forum. Habermas, for example, holds that the terms on which citizens interact with one another in an ideal speech situation precludes manipulation by definition. To this end, he famously refers to the "force of the better

argument”.³⁶ This is an exemplar of appropriate influence. It appeals to participants’ sense of reason, and their capacity to make decisions in an informed and autonomous way. The conditions which preclude manipulation thus also demand that citizens should move beyond the pursuit of mere self-interest when they deliberate together.

Rawls, in contrast, argues that citizens hold special duties towards one another when they deliberate or vote on questions of fundamental importance (Rawls, 2005:214). This places a duty on participants to appeal to reasons which are justifiable from within other reasonable persons’ reasonable comprehensive doctrines rather than, for example, to act strategically in an attempt to manipulate them. D’Agostino has proposed a purely procedural theory of public reasoning in an attempt to address what he calls “reflexivity” the idea that any conception of public reason must itself be somehow justified through public reasoning (1996). Rather than stipulating the substantive answer to this question, therefore, he outlines parameters for deliberation by a constitutional convention of individuals who represent real concrete citizens. Any reasonable individual will thus have reason to accept the parameters of their representatives’ deliberation since she will “see in these rules a reasonable basis for collective decision” (D’Agostino, 1996: 161). This procedural approach thus defers the task of determining how citizens should reason together to delegates. Delegates to the convention are meant to better reflect a rational consideration of the interests of concrete individuals: they do the deliberation that citizens would do under ideal conditions. Here too, there is a move away from mere self-interest. The convention requires: a “public” (not self-interested) orientation; some agreed exemplars of public reason; a diversity of conceptions of public reason; and “reasonable” rules of order for deliberation (1996:161).

The hybrid view of Bentham falls under this, admittedly very broad, family of approaches. It is “deliberative” in the wide sense Fishkin implies. It seeks to define some conditions under which participants should deliberate with one another. It explains the ways in which it is appropriate, and inappropriate for them to influence one another before voting.

Many non-deliberative views of democracy are also concerned with the information that citizens have access to before voting. It is a typical assumption of social choice theory that individuals express their preferences in conditions of appropriate information. Some cases of potential manipulation, can be simply understood in these terms. Deliberately providing voters with misleading or incomplete information just before they vote can lead them to draw errant conclusions from their existing preferences. Manipulation

³⁶ Habermas (1996), *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, Cambridge, MA: MIT Press, Chapter 7.

of the process through which citizens actually form their preferences may be more difficult for some non-deliberative views to conceptualise, however. Consider, for instance, social choice theory, which is often considered a paradigm case of non-deliberative approaches. It is a starting assumption of most views of this kind that the preferences of participants, which need not be self-interested, are fixed. Rather than transforming through a political process of deliberation like that described and tested by Fishkin, agents are supposed to be endowed with fixed preferences (Elster, 1986:105).

So, social choice theorists may be able to identify and explain manipulation which occurs at the point where citizens consider how to vote based on their underlying preferences. The effects of false technical beliefs, induced through poor information have been modelled for exactly this reason (Austen-Smith, 1992). When it comes to manipulation before this, when citizens are actually forming their preferences, however, social choice may have something of a blind spot. Any account which assumes these preferences as fixed and attends to how they should be aggregated rather than formed or revised is likely to be less equipped to satisfactorily explain what is troubling about individuals being manipulated at this prior stage.

This is a problem for the conventional view of Benthamite public reasoning. As we have seen, that interpretation focuses on aggregating citizens' preferences through voting and says little about how those preferences should be formed. It attends in great detail to how democratic institutions should operate, and rules by which representatives should be held to account, but relatively little to the discursive conditions in which citizens should form their views on important matters. It shares a 'blind spot' for manipulation at this earlier stage with some accounts in the social choice tradition. A hybrid Benthamite view is somewhat different. It does provide an account of the information and conditions that citizens need to make appropriate decisions, but not along the deliberative lines commonly proposed in contemporary debates.

Similarly, many deliberative views ask citizens to consider more than self-interest when deciding how to vote, because their conception of quality deliberation embodies notions of public-spiritedness, or impartiality. This does not, however, mean that deliberative views must in any way rule out arguments which present claims based on self-interest playing a role in deliberation. What emerges from this complex picture is that the challenge implied for non-deliberative theories by Fishkin, in his contribution to "Manipulating Democracy" (2011) at least, is not as conclusive as it appears. The conditions he puts forward are extremely broad (some broad conditions of appropriate information, access to arguments on competing sides, and the chance to weigh those arguments). In reality they may be satisfied by a variety of theories; including those customarily referred to as both "deliberative" and "aggregative".

Elsewhere, Fishkin builds these conditions into a more comprehensive picture of deliberation. The crucial question for us is whether the hybrid Benthamite account meets these conditions. I believe it does. Bentham wrote extensively about the level of information and clarity of information needed for individuals to understand their own interests. The POT demonstrates his concern with citizens being provided with balanced information and arguments on which to determine their political views. His writing on sinister interests, deception, delusion, and the need to root out fallacies all suggest a determination to furnish individual citizens with the information and conditions they needed to reach an accurate rather than deluded conception of their own interests. Of course, contemporary theorists might question various aspects of these views. It would, however, be unfair to suppose that Bentham did not provide the conceptual resources needed to define what manipulation is.

Thus, while Fishkin himself builds his own particular account of deliberation, there is nothing in the critique cited above which suggests that public spirited, Habermasian deliberation is required to understand what manipulation is. Rather, Fishkin sets the definition of deliberation at a low level: there should be some account of the appropriate information, balance of arguments on either side, and a chance to weigh those arguments, which citizens need to reach a conclusion before voting. The hybrid Benthamite account meets that test. It proposes that individuals require the chance, through operation of the Public Opinion Tribunal, to gather evidence, exercise the political and moral sanctions, and clarify the understanding of their own interests before voting. This view does not resemble contemporary deliberation. Nonetheless, it does offer a view of the legitimate conditions under which individuals should consider their preferences before voting. These are the conditions needed for citizens to understand their own interests, free from delusion and fallacious logic. In defining these conditions, the hybrid Benthamite view also defines their opposite: manipulated outcomes.

So, whilst the conventional view of Bentham has similarities with some examples of contemporary social choice theory, and the participatory view resembles some examples of contemporary deliberative democratic theories, there is a third alternative. The hybrid view of Bentham could, like the conventional view, hold that individuals should vote according to their self-interest at the ballot box. Like the participatory view, it could also acknowledge that it is necessary to stipulate appropriate conditions for preference formation; rather than merely aggregating expressed preferences and ignoring the conditions under which they were formed.

This reflects the fact that there is real common ground between aggregative and deliberative democracy. Benthamites cannot be deliberative in the sense that Habermasians are. They do not espouse preference change through a process of deliberation. Nonetheless, the hybrid view does prescribe some conditions under which individuals should consider their preferences before voting. It is not content to unthinkingly aggregate whatever expressed preferences pop up from a society by chance. Unlike some of the examples of contemporary deliberative democracy mentioned above, however, it would ask citizens to consider self-interest before voting and rely on the aggregation of votes at the ballot box, rather than the transformative power of deliberation, to produce decisions which reflect the interest of the whole community.

The three broad Benthamite options in response to the challenge identified by Fishkin thus complement three broad responses put forward by contemporary democrats:

- Conventional Benthamite (self-interested)
- Participatory Benthamite (non- self-interested participation)
- Hybrid Benthamite (self-interested participation)

We have seen that the many putatively deliberative accounts of democracy call for individual voters to vote according to the public interest and for the prior process of discourse between citizens to reflect conditions necessary to their doing so effectively. What is critical about Fishkin's own definition of deliberation, however, is not that it is focused in some way on the public interest. Rather, his argument turns on the conditions of preference formation; not orientation. It is, therefore, a warning against blind aggregation taking place without attention being paid to deliberative conditions before voting. So, while it seems implicit in much writing on deliberation that participants should aim to advance something more than their self-interest, this need not be the case.

Instead, what enables such views to reduce scope for manipulation is their ability to account for appropriate conditions of preference formation and, by extension, to rule out inappropriate conditions which allow citizens to be manipulated. Whilst it could be seen as a criticism of the conventional Benthamite view of public reasoning, therefore, the accusation does not affect the participatory or hybrid views to the same degree.

Having examined some potential objections to the hybrid Benthamite view, let us now consider some of its potential advantages. Below, I argue that a self-interested view of public reasoning has some important advantages over alternatives: false consensus, salience, and exclusion.

False consensus

False consensus may be a significant problem for some deliberative views, and less of a problem for others. While there can undoubtedly be significant benefits to well-conducted deliberation, deliberative democrats also acknowledge that common problems encountered by citizens can undermine its value. Important among these is false consensus; the idea that putative agreement can be reached between participants which does not fairly reflect their own sincere and committed consideration of relevant facts and reasons.

False consensus may happen for a number of different reasons, for example because “some communication [is] systematically distorted by power” (Young, 2001: 685). Deliberators may therefore unknowingly reflect the influence of powerful interests in the way they communicate with one another and reach conclusions which are not genuinely justified. False consensus thus makes deliberation more of an act of theatre than a genuine exercise of reflection.

The potential problem of manipulation thus applies most to those “stronger” deliberative views, which expect full consensus, and least to those “weaker deliberative”, “social choice” aggregative and “rational choice” aggregative views for which this type of full agreement, reached through productive deliberation between citizens, is inessential to good decision-making.

This expectation of agreement, or consensus, between deliberating citizens is particularly important to the “stronger deliberative” views we have examined. For views of this kind, consensus plays an important role in explaining the normative value of pre-voting discussion. While useful in that respect, however, the expectation also has the capacity to make manipulated consensus more likely. Citizens may feel particular pressure to conform to an established or majority-held view when they know that deliberations must reach agreement at a specified time. Less influential deliberators may be bullied into agreement as the need to reach consensus becomes more urgent. Under such conditions, a tendency to promote conformity may undermine deliberation (Przeworski, 1998:147). Deliberative democrats often acknowledge this risk (Fishkin, 2009:86).

The consequence of these kinds of social pressure in deliberation, therefore, can be to deny citizens the input they should have. It can also enhance the control of elites over outcomes. Fishkin's own research suggests that those with the ability to use social pressure to shape deliberative outcomes are often those who benefit from the status quo (Ibid).

False consensus is a particular danger in "stronger deliberative" theories because they define an ideal of deliberation which should result in consensus when conducted properly. The view that full consensus is the appropriate result of ideal deliberation is likely to create expectations amongst participants. They may feel, for example, that failures to reach agreement are somehow the result of some form of "deliberative pathology, such as the lingering taint of self-interest" (Waldron, 1999:111). "Weaker deliberative" theorists go some way to assuaging this concern because they accept that deliberation, even perfectly conducted, will often not result in agreement. This acknowledgement means that an expectation that lingering disagreements are some kind of deliberative pathology or fault of those involved should be less common. Similarly, "social choice" aggregative and "rational choice" aggregative views are also unlikely to promote false consensus through the expectation of agreement because they readily acknowledge the importance of aggregation as a closing mechanism for many real world decisions.

By placing a greater emphasis on counting than talking, aggregative theories thus attempt to give each citizen an equal input via elections rather than through ideal discursive conditions (Waldron, 1999:110-16). Counting votes thus aims to ensure a formal equal say for each citizen, and has the advantage of not relying on less demonstrable deliberative norms being present before voting to ensure that talking is not manipulated. Of course, talking could still be manipulated in such a situation but this will have a less damaging effect on a democratic theory the less that it relies on deliberation to change preferences. An advantage of emphasising counting over talking in this way is that it is likely to be less open to the kinds of social conditions that might make dissenters believe that their views have not been adequately heard:³⁷

Aggregation "allows those on the losing side to hold on to their integrity" because their views have been treated seriously, and with as much respect as others' but have been outweighed". (Bellamy, 2009:106)

³⁷ Of course, deliberationists, too often advocate counting votes, but the extensive role they afford deliberation before voting nevertheless provides more possible opportunities for elite manipulation

By placing more emphasis on counting, and less on talking, therefore, we make dissenters less likely to believe that they only lost votes among peers because discursive conditions were manipulated. This advantage comes at a price, however. Aggregationists eschew “epistemic moralised” deliberation: the idea that discussion before voting has a particular moral status beyond its instrumental benefits. As a result, they cannot claim the supposed advantages of that kind of public deliberation. Primary among these claimed advantages is citizens’ increased regard for the public interest. According to some deliberative theses, orientating deliberation towards the public interest creates the expectation that arguments should be justified in more “public spirited” terms. Deliberation thus has an important benefit: the expectation of “public-spirited” deliberation controls elite power because justifying arguments during public deliberation gives incentives for officials to appear impartial i.e. appeal to the public interest.

The hybrid Benthamite approach sits between these two groups. Like aggregative theories, it does not value or expect full consensus from participants, so should be less susceptible to false consensus. Like some deliberative theories, however, it also acknowledges the importance of describing appropriate conditions for the formation of preferences. In doing so, it goes some way toward explaining the value of pre-voting discussion. Importantly, the hybrid Benthamite view is also different from many accounts of deliberation in that it expects citizens to consider their own interests before voting. I believe that a self-interest-orientated debate should reduce expectations of consensus.

Theorists of deliberation often draw a direct link between expectation of consensus and the danger of false consensus. Where consensus is required, or overwhelmingly expected, participants are more likely to feel that they have to agree to a collective decision that they do not believe is truly justified. Fishkin, for example, expresses a particular concern that the expectation of agreement can distort deliberative outcomes; that “the necessity of a consensus will require something closer to false consensus—like a jury verdict reached under great pressure for a decision... despite sincere misgivings” (Fishkin, 2009:86). Some legal and political scholars, sceptical about the benefits of deliberation, have suggested that the effects of distortions like these critically undermine the learning that deliberating together produces. Posner argues, for example, that deliberative democratic theory is likely to load the terms of debate to favour particular outcomes (Posner, 2003:130-140).

The kind of pressure which Fishkin alludes to is a well-known danger of jury deliberations where a unanimous verdict is required. This scenario was famously represented in the 1957 film “Twelve Angry Men”. The film depicts a jury required to reach a unanimous verdict on the guilt of a young Spanish-

American man accused of murdering his father. Deliberations begin in a febrile environment, with the majority of jurors taking the initial view that the defendant is guilty. The knowledge that a unanimous verdict is required, and that a dominant view prevails within the jury places those jurors who dissent from the majority view under immense pressure to conform. It can lead, as in this fictional case, legitimate questions to be dismissed out-of-hand:

Juror number 7: “Well, what's there to talk about? Eleven men in here think he's guilty. No one had to think about it twice except you”

In cases like these, deliberation may harness an unthinking conformity to expressed views of other participants rather than a genuine opinion reached on the basis of relevant evidence. As deliberative theorists acknowledge, the risk of false consensus is related to an expectation that participants should all agree with one another once deliberation has been properly conducted. A jury deliberating under the requirement of unanimity is one example of this kind of expectation. A minority of dissenters is seen as an inconvenient obstacle to a wider group which needs to reach a collective decision. Expectations of this sort make it more likely that legitimate views will go unconsidered, that legitimate concerns will not be voiced and that false consensus will result on important questions.

Of course, there are some important differences between jury deliberation and reasoning in the wider public. In the wider public, individuals may rely on the media and other general sources of information rather than a small group of individuals. The secret ballot may also insulate voters from social pressure exerted by their peers, so reduce the risk of false consensus. If the danger of false consensus is related to the expectation of agreement, then other factors are also likely to be important. In particular, the object of deliberation may well affect whether participants and observers expect agreement and whether they believe it to be valuable.

Remember that the hybrid Benthamite view is not “aggregative” in the sense that it neglects the importance of pre-voting discussion. Nor is it deliberative in the sense proposed by Habermas, which implies that full consensus on generalizable interests should result from discourse. It defines some conditions of appropriate information and reflection before voting. It does so, however, to allow participants to collect information and clarify and understand their own interests rather than to transform their preferences towards the public interest. The object of pre-voting discussion is thus clarifying self-interest. I believe that this has some important implications for the problem of false consensus.

Jurors, or voters, subject to a high risk of false consensus usually deliberate on the same question. As a result, dissenting minorities must directly contradict the views of a prevailing majority in order to freely express their opinion. Dissent requires the accusation that the majority is mistaken. Deliberation regarding a shared question, like the guilt or innocence of an individual, or whether a proposal advances or hinders the national interest therefore heightens the risk of false consensus because it demands more of minorities who wish to dissent against a prevailing view.

The picture would be different if discussion were directed towards individuals' self-interest rather than a single common object, like the public interest. As Goodin has noted, most Western societies are willing to accept the notion that individuals should usually be regarded as the best available judge of their own interests. Thus, "Usually (not always) each individual is probably (not necessarily) the best judge of his own interests" (Goodin, 1995: 127). Therefore, if public reasoning is directed toward self-interest, the risk of false consensus changes significantly. The best judge principle implies that dissenting minorities cannot so readily be told that their preferences are wrong, any more than one consumer can tell another that they are wrong to like raspberry more than chocolate ice cream. This is particularly the case if we adopt, as Bentham did, a quantitative form of utilitarianism, which rejects the notion that some preferences are superior to others.

Participants in hybrid Benthamite public reasoning are not told that they must reach agreement; or that they must all consider the same question. Rather, they are encouraged to consider information in relation to a matter on which they are the best judge: their own interests and preferences. Following discussions, they then vote in secret, since this is "the only security for genuineness of suffrage" (Bowring, iii, 599). Focussing on self-interest means that each individual voter considers a slightly different question. On this matter he or she is likely to have more information and experience than others. In these conditions, it should be easier for minorities to dissent, and to register their legitimate concerns at the ballot box. Conversely, it should be more difficult for majorities to, deliberately or unwittingly, pressurise dissenters to conform.

In this regard, the hybrid Benthamite view, which emphasises voting in secret and individual reasoning on the basis of self-interest, is an extension of some contemporary deliberative views. The reasons cited by deliberative democrats like Fishkin for preferring a mode of deliberation which does not seek full consensus are also reasons to think that a self-interested mode of public reasoning will be less susceptible

to manipulation through false consensus than public interested alternatives. Individuals need the opportunity make their own mind up without excessive social pressure. This is aided by a self-interested conception of public reasoning and impeded by public-interested reasoning. Asking citizens to deliberate together on the public interest, or something approaching it, is more liable to generate an expectation of agreement than asking them to deliberate regarding their own interests; there may be as many correct answers to the latter as there are people participating. Deliberative views which ask citizens to consider what serves their own interests therefore hold the advantage that they are less likely to create the social pressure which provokes false consensus. The format of deliberation in juries may lend itself more to false consensus than that of deliberative polling with confidential questionnaires. So too, should deliberation on self-interest—by individuals who then vote in secret—be less likely to encourage social pressure which undermines the outputs of deliberation.

The hybrid Benthamite view thus sits alongside the forms of non-Habermasian deliberative democratic theory which try to mitigate the danger of false consensus and promote opportunities for legitimate dissent. Fishkin's (2009) "When the People Speak" (2009) is a prominent example of this. He acknowledges that it is a genuine problem for deliberation and could undermine the value of its outcomes. He looks at empirical evidence of where false consensus tends to distort the quality of deliberation, and finds that "most of the empirical literature supporting [these] distortions...comes from jury studies" rather than analyses of deliberative poll (DP) or other forums. He contends that there are critical differences between these two contexts. Those differences reflect the fact that Deliberative Polling was deliberately designed to reduce the danger of false consensus—by insulating individuals against undue social pressure:

"A jury arrives at an agreed verdict. The necessity for such agreement creates social pressure for consensus. By contrast a DP solicits opinion in confidential questionnaires and tries to avoid any pressure for consensus...A jury does not have moderated discussion. The jury foreman is a leader but not a moderator..." (Fishkin, 2009: 129-30)

Fishkin's mode of Deliberative Polling is thus designed to shield participants from the social pressure which can engender false consensus. It does so by: reducing expectation that full collective discussion should create full agreement; allowing individuals to express views in private; and facilitating discussion to try to allow more equal participation and influence. Bentham, of course, had similar concerns. These were central to his advocacy of the secret ballot. He wished to protect voters from local landowners' undue influence. In Fishkin's account of deliberation, confidential questionnaires play the same role as the secret ballot did for Bentham. Both allow individuals to register their views without the influence or judgement of peers. There are striking similarities between Bentham's objection to public voting, and

Fishkin's objection to public consultation. The former warns that ... "Expose him to the public eye... [and] You expose him at the same time to the eyes of some individual or individuals who have an interest of *their own opposite to the interest of the public* and who have it in their power to make you suffer if you do not give preference to theirs" (Bentham, 2002:431, italics added). The latter warns that "Public consultations that seek agreement... expose participants to social pressure to reach agreement. It is a far better guarantee to avoid consensus-seeking processes and to gather the opinions, before and after, in confidential questionnaires or secret ballots" (Fishkin, 2011: 38)

The hybrid Benthamite view accepts this concern, and extends it by defining the object of public reasoning as the clarification of self-interest, thus further reducing the danger of false consensus. Similarly, "weaker" deliberative theorists, like Gutmann & Thompson (2008) share the contention that deliberation should not be expected to produce full consensus. For them, deliberation is valuable even where consensus is impossible, because it clarifies positions and builds reciprocal relations between citizens. False consensus is, however, a particular danger in "stronger deliberative" theories because they define an ideal of deliberation which should result in consensus when properly conducted. These accounts—because they create an expectation of consensus—always seem more likely to create accusations that deliberators have somehow "failed" when disagreement persists. I believe that this is because they are more willing to evaluate and criticise the preferences that citizens form; or the conditions under which they are shaped. Accusations of that sort are likely to disadvantage those with dissenting views (Waldron, 1999:111).

Exclusion

Related to the danger of false consensus is that of exclusion. Public-interested modes of public reasoning often create expectations of the types of reasons and persons which should prevail in key discussions before voting. Like the expectation of consensus on the public interest, these demands risk excluding legitimate dissent.

One of the great advantages claimed for contemporary deliberation is that it encourages "we" rather than "I" thinking from citizens (Mansbridge, 1991: 7-8). What is potentially troubling about this contention is that encouraging "we" thinking is often also accompanied by a definition of exactly who "we" are. Public-interested views of deliberation commonly require that particular forms of argument are required in order to aid learning about the common interest: "Appeals to deliberation amount to demands for a certain type of discourse in democratic political settings; reasonable, foresighted, steady, and oriented to a common, not sectarian, problem" (Sanders, 1997: 356). Phillips draws on a similar principle to critique

some deliberative views (1995:155). It is, she contends, the precepts of these of these kinds of approaches which can reflect majority orthodoxies and thus stifle the ability of minorities to contribute legitimate views to public discourse.

One effect of appeals to the public interest can be to constrain the types of reasons to which individuals may appeal, thus invoking “connotations of rationality, reserve, cautiousness, quietude, community, selflessness and universalism” (Ibid: 348). These ideas might, in the abstract, be devoid of the influence of class, gender, power, or race. However, their implementation in the real world can be a very different story. There is therefore a danger in relying on the terms of public deliberation to reach the public interest: public-spirited parameters for interaction can so easily be usurped by existing social cleavages. It is for this reason that critics of deliberative democracy worry about how abstract deliberative principles can be put into action, and suggest that proponents of some deliberative theories “seem to live on another planet... they are devoid of race, class, and gender and all the benefits and liabilities associated... with these features” (Sanders, 1997: 353). Rawlsian public reasoning is an obvious example of this kind of approach (Rawls, 1997).

We have seen that Posner has pointed towards a similar danger. He suggests that perhaps “what motivates many deliberative democrats is... a desire to change specific political outcomes, which they believe they could do through argument, if only anyone could be persuaded to listen, because they are masters of argumentation” (Posner, 2004). So, the demand to orientate deliberation to the public interest creates standards for justification or expression of arguments. To different degrees, these have the potential to be exclusionary. The criteria established for encouraging citizens to preclude certain types of reasons from debate can be exploited by elites in support of their favoured outcomes; to effectively exclude some members of the community.

Public-interest-orientated theories often state that properly conducted deliberation excludes no one; but patterns of exclusion may in fact be insidious (Sanders, 1997: 353). Deliberative views often address this concern by requiring “equal access to decision-making assemblies or substantive equality in resources and power” as part of the ideal they strive to realise (Phillips, 1995:154). Without a more detailed practical account of how these conditions might be realised, however, real world deliberation where decisions are required within reasonable timeframes and from people grounded in existing social and cultural conditions, may give too little weight to some voices and too much to others.

Of course, well conducted deliberation is designed to prevent that happening. These principles are typically intended to prevent any substantive bias in debate. Facilitation should also ensure that seldom heard voices have the impact they deserve. Note, however, that none of these deliberative principles must—by definition—be directed at the public interest. Information, substantive balance, diversity, conscientiousness, and equal consideration all apply to individuals' considering their own interests as much as those considering the public interest (Fishkin, 2009:34). The hybrid Benthamite view is thus compatible with measures suggested by deliberative democrats to protect against false consensus and exclusion. It could even be seen as an extension of those measures. By accepting principles like these, and directing citizens to consider self-interest rather than the public interest, it may be possible to safeguard against manipulation through effective facilitating and appropriate conditions without exposing participants to the dangers of exclusion which some public-interested deliberative views create. The relevant comparison is not thus between a jury and deliberative polling, but between deliberative polling focused on the public interest or on self-interest; on an object shared between all members of the group, or different objects held by each as an individual.

In contrast, deliberation aimed at something beyond self-interest will be more susceptible to exclusion. As Phillips suggests, deliberation of this sort (which I have termed “stronger”) often endorses an idealised form of discussion. This encourages “oppressed groups” to “put their own partial interests aside—to address the shared concerns of all humanity, to think beyond their own interests and needs”. Of course, this is something that no self-interested view of deliberation would want to encourage. A move away from self-interest would not be necessary, for example, under the conventional Benthamite, or some social choice, approaches. Phillips' critique is that requiring oppressed groups to move towards consideration and expression of wider, even universal, interests can in reality “lock them into the very structures they are trying to dislodge” (Phillips, 1995:147). Thus, an acceptance of prejudice or bias into the assumptions which undergird deliberation itself protects that bias from being effectively challenged, because when “disregard based in prejudice goes unrecognised by both those who are subject to it and those who are prejudiced, prejudices cannot possibly be challenged” within a debate framed by those very assumptions (Sanders, 1997: 353). The influence of such exclusionary practices, placing less weight on the views of some than others, may also be insidious, thus difficult to identify and eradicate through deliberation. We have seen that one deliberative response to this, from Fishkin, puts forward principles to ensure that all have a reasonable input into collective decisions. Indeed, the hybrid Benthamite view complements those principles, and perhaps extends them.

Saliency

Finally, I contend that self-interested public reasoning is less liable to manipulated outcomes because it induces discussion on issues of higher saliency to individuals. These, in turn, are less likely to be the product of manipulation. Critics of deliberative democracy sometimes lament the high-minded, public spiritedness it requires of citizens, and argue that “People who vote on the basis of their self-interest are at least voting about something they know at first-hand, their own needs and preferences. Beware the high-minded voter” (Posner, 2003:113)

There is some research on deliberation which suggests that deliberation on higher saliency issues is less likely to produce significant preference change. I believe that self-interested deliberation could be understood as “high saliency” deliberation. It follows from this that self-interested deliberation is likely to have a less significant effect—in whichever direction—on participants’ expressed preferences. Now it is possible, of course, that changes provoked by deliberation could be either positive or negative; they could produce a better or worse reflection than preceding views of the interests and reasons at stake. A difference in saliency is therefore itself normatively inconclusive.

It is important, however, to remember that different accounts ask public reasoning to play different roles in the democratic process. The participatory Benthamite, and some contemporary deliberative views, for example, suggests that public reasoning helps to make votes cast at the ballot box a better reflection of the public interest. In contrast, the conventional Benthamite and some social choice views instead wish individual votes to reflect individual interests, and rely on the aggregation of votes to produce decisions which serve the public interest. This, in turn, means that a conventional Benthamite view could conceivably produce decisions which serve the public interest *without* the need for public-interested deliberation, if a subsequent aggregation of interests through voting can take place successfully. The hybrid Benthamite account shares that view.

Why might that be an advantage? A public-interested account needs the benefits of the changes to preferences created by deliberation in order to produce “good” (public interest-reflecting) decisions. To achieve those benefits, it must also accept the greater risk of manipulation that lower saliency deliberation implies. In contrast, a self-interested account does not rely on this change in preferences during deliberation. Thus, achieving good decisions through a public interested deliberative view implies more risk of manipulated outcomes than achieving good decisions through a self-interested deliberative view,

like the hybrid Benthamite account. Assuming the success of both views, the former requires lower salience deliberation to produce good outcomes; the latter assumes higher salience deliberation.

This is important. Various studies on deliberation show that preferences tend to change less when a highly salient issue is being considered. There is some evidence that deliberation produces less preference change for “high salience” issues (Farrar et al, 2010). List et al., for example, showed that “deliberation increases proximity to single-peakedness...at least on low- to moderate-salience issues where there has not been much prior deliberation and where proximity to single-peakedness is not already high” (List et al, 2013:90-91).

Studies like this generally measure salience by the amount of coverage an issue has received in relevant media. List, for example, distinguishes between high and low salience issues by reference to the amount of “public attention” received before deliberative polling:

“Electric utility policies in Texas and revenue sharing in New Haven had received little public attention...The Monarchy in Britain, the airport in New Haven, and the Constitutional Referendum in Australia had received incomparably more. These latter three were highly salient” (List et al, 2013:86)

There are fewer studies of explicitly self-interested deliberation. It is plausible to suppose, however, that viewing a question with regard to one’s self-interest renders it of higher salience for an individual than regarding it by reference to the public interest, particularly since familiarity or proximity to an issue is usually used as a proxy for salience. If this is the case, then it follows that, in line with the claims of deliberative democrats, we should expect a greater degree of preference change during public-interested deliberation than during self-interested deliberation. This concurs with the greater emphasis given to preference change by a variety of deliberative democrats than by non-deliberative theorists (Elster, 1986).

If it is true that a self-interested view of public reasoning promotes less change in preferences then it may be true that such a view also presents less opportunity for preferences to change for the worse as well as the better. The potential high salience of self-interested reasoning thus reduces the potential gains and losses associated with deliberation. If this is the case, it might be reasonable to believe, by extension, that deliberation on matters of self-interest rather than public interest should cause less change in stated preferences.

The three suggested advantages of self-interested over alternative views of deliberation are therefore contingent on the assumption that aggregation of votes can subsequently reach the public interest, as some aggregative democrats like Bentham suppose.

Conclusion

This thesis has attempted to consider two separate areas of debate in Bentham scholarship and political theory to discern what the latter might learn from the former. Chapter 1 explored two challenges for a Benthamite account of public reasoning: providing an account of effective discussion before voting, and reducing opportunities for manipulated outcomes from discussion. These demands flow from a division made in contemporary writing on public reasoning: between “aggregative” and “deliberative” democracy. The former is often accused of neglecting the role of pre-voting discussion, but could offer fewer opportunities for manipulated outcomes. The latter provides a rich account of pre-voting discussion but perhaps—in doing so—create more potential for outcomes to be manipulated.

Chapter 2 outlined a “simple” or “conventional” view of Bentham. This is often suggested by contemporary public reasoning theorists. The conventional view suggests that Bentham did little to provide an account of effective discussion before voting. Rather he was content to allow the aggregation of “raw” preferences to produce collective decisions.

In the third chapter, we examined more “participatory” views of Bentham. These go further to understand his views on the role of pre-voting discussion. In doing so, they respond better to the challenge of explaining the value of discussion. For different reasons, however, participatory views are often lacking as interpretations of Bentham’s work. Some depart from crucial themes in his wider corpus (Ben-Dor); while others focus on parliamentary rather than public deliberation.

Chapter 4 analysed some important principles in Bentham’s democratic theory. From these, it suggested that a “hybrid” view of Bentham was possible. Like the participatory view, this provides an explanation of the value of pre-voting discussion. Unlike the participatory view, however, it maintains a self-interested orientation to debate. This kind of participation fosters “public aptitude” from citizens. This is the quality required from them to nurture official aptitude from their political representatives.

Finally, in this chapter, I have considered some objections to the hybrid Benthamite view. Many of these can be answered from within the principles Bentham himself laid out; others are more problematic. Overall, I have tried to show that the kind of participation that Bentham wished to encourage from citizens was more than simply “aggregative” (i.e. reliant solely on the aggregation of votes). Rather, he wanted to see an active and critical engagement from citizens which could serve as a security against misrule by officials. The principles outlined therefore inhabit a territory between aggregative and deliberative contemporary theories. They also have some important advantages. The hybrid view explains why pre-voting discussion should be considered an important component of democratic procedure. It does so, however, by relying on self-interest rather than the public interest. As a result, it is potentially less susceptible to criticisms made of deliberative democracy: that it enables a paternalism about citizens’ preferences—and the conditions in which they are formed—which could facilitate their manipulation.

Like some aggregative theories, the hybrid view is focused on the self-interest of participants rather than the public interest, and reflects existing preferences rather than transforming them. Like some deliberative theories, the hybrid view provides an account of how citizens should interact with one another and reflect before voting. The hybrid view therefore illustrates that there is far more common ground between aggregative and deliberative democratic theory than proponents of each view often suppose. This interpretation thus clarifies our understanding of Bentham in relation to contemporary writing on public reasoning. It also strengthens the view that aggregative and deliberative democracy should be seen as compatible, rather than competing types of approach.

References:

- Alexander, Larry. 1985 "Pursuing the Good-Indirectly." *Ethics* 95:315-32.
- Annas, Julia. 1977. "Mill and the Subjection of Women." *Philosophy* 52:179-94.
- Andrew, Christopher. 1977. "The British Secret Service and Anglo-Soviet Relations in the 1920s Part I: From the Trade Negotiations to the Zinoviev Letter." *The Historical Journal* 20:673-706.
- Arneson, Richard. 2004. "Democracy Is Not Intrinsically Just." In *Justice and Democracy: Essays for Brian Barry*, edited by Keith Dowding, Robert E. Goodin and Carol Pateman. 40-68. Cambridge: Cambridge University Press.
- Arrow, Kenneth J. 1963. *Social Choice and Individual Values*. New York: Wiley.
- Audi, Robert. 1991. "Religious Commitment and Secular Reason: A Reply to Professor Weithman." *Philosophy & Public Affairs* 20: 66-76.
- Austen-Smith, David. 1992. "Strategic Models of Talk in Political Decision Making." *International Political Science Review* 13: 45-58.
- Ball, Terrence. 1993. "Was Bentham a Feminist?" In *Bentham, Critical Assessments, Volume 1*, edited by Bikhu C. Parekh. London: Routledge.
- Barber, Benjamin R. 2003. *Strong Democracy: Participatory Politics for a New Age*. Berkeley: University of California Press.
- Bellamy, Richard. 2009. "The Republic of Reasons: Public Reasoning, Depoliticisation and Non Domination" In *Legal Republicanism: National and International Perspectives*, edited by Samantha Besson and José Luis Martí, 102-20. Oxford: Oxford University Press.
- Ben-Dor, Oren. 2000. *Constitutional Limits and the Public Sphere: A Critical Study of Bentham's Constitutionalism*. Oxford, Hart Publishing.
- Benhabib, Seyla. 1996. "Toward a deliberative model of democratic legitimacy." In *Democracy and difference: Contesting the boundaries of the political*, edited by Seyla Benhabib, 67–94. Princeton: Princeton University Press.
- Bentham, Jeremy. 1818a. *Plan of Parliamentary Reform in the form of a Catechism, with reasons for each article*. London: T.J. Wooler. Online:
https://books.google.co.uk/books/about/Plan_of_Parliamentary_Reform_in_the_Form.html?id=iU-VBKIEz5UC&redir_esc=y

Bentham, Jeremy. 1818b. *Church of Englandism and its Catechism Explained*. London: Effingham Wilson. Online:

<https://ia800303.us.archive.org/1/items/churchofengland00bentgoog/churchofengland00bentgoog.pdf>

Bentham, Jeremy. 1821a. *The Elements of the Art of Packing as Applied to Special Juries, Particularly in Cases of Libel Law*. London: Effingham Wilson. Online:

https://books.google.co.uk/books/about/The_Elements_of_the_Art_of_Packing.html?id=FM0DAAAQAAJ.

Bentham, Jeremy. 1821b. *On the Liberty of the Press, and Public Discussion*. London: William Hone.

Online: <https://books.google.co.uk/books?id=ICJXAAAACAAJ>.

Bentham, Jeremy. 1822. *An Analysis of the Influence of Natural Religion on the Temporal Happiness of Mankind*. London: Carlile. Online: <https://archive.org/details/analysisinfluen00bentgoog>.

Bentham, Jeremy. 1830. *The Rationale of Reward*. London: John and H.L. Hunt. Online:

https://books.google.co.uk/books?id=L5Q7AAAAYAAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.

Bentham, Jeremy. 1843. *The Collected Works of Jeremy Bentham*. Edited by John Bowring. Online:

http://oll.libertyfund.org/?option=com_staticxt&staticfile=show.php%3Ftitle=1920.

Bentham, Jeremy. 1970a. *The Collected Works of Jeremy Bentham: Introduction to the Principles of Morals and Legislation*. Edited by Herbert L.A. Hart. London: Athlone Press.

Bentham, Jeremy. 1970b. *Of Laws in General*. Edited by Herbert L.A. Hart. London: Athlone Press.

Bentham, Jeremy. 1977. *The Collected Works of Jeremy Bentham: A Comment on the Commentaries and a Fragment on Government*. Edited by Herbert L. A. Hart and James H. Burns. London: Athlone Press.

Bentham, Jeremy. 1981. *The Correspondence of Jeremy Bentham. Volume 4: October 1788 To December 1793*. Edited by Alexander Taylor Milne. London: Athlone Press.

Bentham, Jeremy. 1983a. *Chrestomathia*. Edited by M.J. Smith and W.H. Burton. Oxford: Clarendon Press.

Bentham, Jeremy. 1983b. *The Collected Works of Jeremy Bentham: Constitutional Code: Volume 1*. Edited by Fred Rosen and James H. Burns. Oxford: Clarendon Press.

Bentham, Jeremy. 1983c. *Deontology Together with a Table of the Springs of Action and the Article on Utilitarianism*. Edited by Amnon Goldworth. Oxford: Clarendon.

Bentham, Jeremy. 1989. "Constitutional Code Rationale." In *The Collected Works of Jeremy Bentham: First Principles Preparatory to Constitutional Code*, edited by Philip Schofield. Oxford: Clarendon Press

Bentham, Jeremy. 1990. *The Collected Works of Jeremy Bentham: Securities against Misrule and Other Constitutional Writings for Tripoli and Greece*. Edited by P Schofield. Oxford: Clarendon.

Bentham, Jeremy. 1993, *The Collected Works of Jeremy Bentham: Official Aptitude Maximized, Expense Minimized*. Edited by Philip Schofield. Oxford: Clarendon Press.

Bentham, Jeremy. 1995. *The Panopticon Writings*. Edited by Miran Božovič. London: Verso Books.

Bentham, Jeremy. 1997. *De L'ontologie Et Autres Textes Sur Les Fictions*. Edited by Philip Schofield, Jean Pierre Cléro and Christian Laval. Paris: Éditions du Seuil

Bentham, Jeremy. 1998a. *The Collected Works of Jeremy Bentham: Legislator of the World: Writings on Codification, Law and Education*. Edited by Philip Schofield and Jonathan Harris. Oxford: Clarendon Press.

Bentham, Jeremy. 1999. *The Collected Works of Jeremy Bentham: Political Tactics*. Edited by Michael James, Cyprian Blamires and Catherine Pease-Watkin. Oxford: Clarendon Press.

Bentham, Jeremy. 2000. *The Collected Works of Jeremy Bentham: The Correspondence of Jeremy Bentham: Volume 11, January 1822 to June 1824*. Edited by Catherine Fuller. Oxford: Clarendon Press.

Bentham, Jeremy. 2001. *The Collected Works of Jeremy Bentham: Writings on the Poor Laws: Volume 1*. Oxford: Clarendon Press.

Bentham, Jeremy. 2002a. "Necessity of an Omnipotent Legislature." In *The Collected Works of Jeremy Bentham: Rights, Representation and Reform – Nonsense upon Stilts and Other Writings on the French Revolution*, edited by Philip Schofield, Catherine Pease-Watkin and Cyprian Blamires. Oxford: Clarendon Press.

Bentham, Jeremy 2002b. *The Collected Works of Jeremy Bentham: Rights, Representation and Reform – Nonsense upon Stilts and Other Writings on the French Revolution*. Edited by Philip Schofield, Catherine Pease-Watkin and Cyprian Blamires. Oxford: Clarendon Press.

Bentham, Jeremy. 2014. *The Collected Works of Jeremy Bentham: Of Sexual Irregularities, and Other Writings on Sexual Morality*. Oxford: OUP.

The Behavioural Insights Team, 2014. EAST: Four Simple Ways to Apply Behavioural Insights. <http://www.behaviouralinsights.co.uk/publications/east-four-simple-ways-to-apply-behavioural-insights/>

Binmore, Kenneth. 2011. *Rational Decisions*. Princeton: Princeton University Press.

- Bohman, James. 1998. "Survey Article: The Coming Age of Deliberative Democracy." *Journal of Political Philosophy* 6: 400-25.
- Bonner, John. 1995. *Economic Efficiency and Social Justice: The Development of Utilitarian Ideas in Economics from Bentham to Edgeworth*. Aldershot: Elgar.
- Bozzo-Rey, Malik. 2010. "Publicity and Nomography: Bentham's School of Legislation." *Journal of Bentham Studies* 12: 1-20. Online: [http://discovery.ucl.ac.uk/1323708/1/012%20Bozzo-Rey%202010 .pdf](http://discovery.ucl.ac.uk/1323708/1/012%20Bozzo-Rey%202010.pdf)
- Brink, David. 1992. "Mill's Deliberative Utilitarianism." *Philosophy & Public Affairs* 21: 67–103.
- Brunon-Ernst, Anne. 2012. *Utilitarian Biopolitics: Bentham, Foucault and Modern Power*. London: Pickering & Chatto.
- Budge, Ian. 1993. "Jeremy Bentham: A Re-Evaluation in the Context of Empirical Social Science in Bentham." In *Bentham: Critical Assessments, Volume 3*, edited by Bhikhu C. Parekh. London: Routledge.
- Chambers, Simone. 2004. "Behind Closed Doors: Publicity, Secrecy, and the Quality of Deliberation." *The Journal of Political Philosophy* 12: 389–410.
- Choi James, David Laibson, Brigitte Madrian, and Andrew Metrick. 2003. "Optimal Defaults." *American Economic Review Papers and Proceedings*. 93:180-185.
- Cohen, Joshua. 1989. "Deliberative democracy and democratic legitimacy." In *The Good Polity: Normative Analysis of the State*, edited by Alan Hamlin and Philip Pettit, 17–34. Oxford: Blackwell.
- Cohen, Joshua. 1998. "Democracy and Liberty." In *Deliberative Democracy*, edited by Jon Elster, 185-231. Cambridge: Cambridge University Press.
- Coleman, Jules. 1989. "Rationality and the justification of democracy." In *Politics and Process: New Essays in Democratic Thought*, edited by Geoffrey Brennan and Loren Lomasky. 194-220. Cambridge: Cambridge University Press.
- Condorcet, Marie Jean Antoine de Caritat. 1785. *Essai sur l'application de l'analyse a la probabilité des decisions rendues a la pluralité des voix*, Paris: L'Imprimerie Royale.
- Converse, Philip E. 2006. "The nature of belief systems in mass publics." *Critical Review: A Journal of Politics and Society*, 18: 1-74.
- Crimmins, James. 2006. "Mill on Bentham." Paper presented at the John Stuart Mill Bicentennial Conference 1806-2006, University College, London 5-7 April 2006.

Crimmins, James. 2011. *Utilitarian Philosophy and Politics: Bentham's Later Years*. London: Continuum International Publishing Group.

Cutler, Fred. 1999. "Jeremy Bentham and the Public Opinion Tribunal." *Public Opinion Quarterly* 63: 321-46.

D'Agostino, Fred. 1996. *Free Public Reason: Making it up as we go*. Oxford and New York: Oxford University Press.

De Champs, Emmanuelle. 2008. *La De 'Ontologie Politique' Ou La Pense 'E Constitutionnelle De Jeremy Bentham*. Paris: Librairie Droz.

De Champs, Emmanuelle. 2014. "Utility, Morality, and Reform: Bentham and Eighteenth-Century Continental Jurisprudence." In *Bentham's Theory of Law and Public Opinion*, edited by Xiaobo Zhai and Michael Quinn, 184-207. Cambridge: Cambridge University Press.

Department for Work and Pensions (DWP), 2010, *Making automatic enrolment work: A review for the Department for Work and Pensions*. Online:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214585/cp-oct10-full-document.pdf

Dietrich, Franz, and Christian List. 2007. "Strategy-proof judgment aggregation." *Economics and Philosophy* 23: 269-300.

Diggs, Bernard J. 1973. "The common good as reason for political action." *Ethics* 83: 283-293.

Dinwiddy, John. 1982. "Bentham on private Ethics and the Principle of Utility." *Revue Internationale de Philosophie* 36: 278-300.

Dinwiddy, John. 2003. "Adjudication under Bentham's Pannomion." In *Bentham. Selected Writings by John Dinwiddy*, edited by William Twining. Stanford: Stanford University Press.

Downs, Anthony. 1957. "An economic theory of Political Action in a Democracy." *Journal of Political Economy* 65: 135-150.

Druckman, James N., and Kjersten R. Nelson 2003. "Framing and Deliberation: How Citizens' Conversations Limit Elite Influence." *American Journal of Political Science* 47: 729-745.

Dryzek John S. 2002. *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*. Oxford and New York: Oxford University Press.

Dryzek, John S., and Christian List. 2003. "Social Choice Theory and Deliberative Democracy: A Reconciliation." *British Journal of Political Science* 33: 1-28.

- Dworkin, Ronald. 1978. *Taking Rights Seriously*. London: Duckworth
- Edelman, Murray. 1974. "The Political Language of the Helping Professions." *Politics & Society* 4: 295-310.
- Eggers, Andrew, and Alexander Fisher. 2011. "Electoral Accountability and the UK Parliamentary Expenses Scandal: Did Voters Punish Corrupt MPs?" *Political Science and Political Economy Working Paper*, London School of Economics 8.
- Elster, Jon. 1986. "The Market and the Forum: Three Varieties of Political Theory." In *Foundations of Social Choice Theory*, edited by Jon Elster and Aanund Hylland, 103-30. Cambridge: Cambridge University Press.
- Elster, Jon. 1998. "Introduction." In *Deliberative Democracy*, edited by Jon Elster, 1-18. Cambridge: Cambridge University Press.
- Elster, Jon. 2011. "Deliberation, Cycles, and Misrepresentation." Paper presented at the Epistemic Democracy in Practice Conference, Yale University, October 20.
- Estlund, David M. 1994. "Opinion leaders, Independence, and Condorcet's Jury Theorem." *Theory and Decision* 36: 131-162.
- Estlund, David M. 2008. *Democratic Authority: A Philosophical Framework*. Princeton and Oxford: Princeton University Press.
- Fearon, James D. 1998. "Deliberation as Discussion." In *Deliberative Democracy*, edited by Jon Elster, 44-69. Cambridge: Cambridge University Press
- Ferraro, Francesco. 2010 "Direct and Indirect Utilitarianism in Bentham's Theory of Adjudication." *Journal of Bentham Studies* 12: 1-24.
- Fishkin, James. 2009. *When the People Speak*. Oxford: Oxford University Press.
- Fishkin, James. 2011. "Manipulation and Democratic Theory." In *Manipulating Democracy*, edited by Wayne le Cheminant and John M. Parrish, 31-41. New York: Routledge.
- Frazer, Michael. 2007. "John Rawls: Between Two Enlightenments." *Political Theory* 35: 756-80.
- Freeman, Samuel. 2000. "Deliberative Democracy: A Sympathetic Comment." *Philosophy & Public Affairs* 29: 371-418.
- Freeman, Samuel. 2011. "Problems with Some Consequentialist Arguments for Moral Rights." Presented at the University College London Colloquium on Legal and Social Philosophy, London, 2 March.

- Frey, Bruno S., and Alois Stutzer. 2001. "Beyond Bentham—Measuring Procedural Utility." *ESifo Working Paper Series* 492. Online: http://www.cesifo-group.de/ifoHome/publications/working-papers/CESifoWP/CESifoWPdetails?wp_id=14560965
- Freyenhagen, Fabian. 2011. "Taking Reasonable Pluralism Seriously: An Internal Critique of Political Liberalism." *Politics, Philosophy & Economics* 10: 323-42.
- Frohock, Fred M. 1999. *Public Reason: Mediated Authority in the Liberal State*. Ithaca and London: Cornell University Press.
- Gambetta, Diego. 1998. "'Claro!': An Essay on Discursive Machismo." In *Deliberative Democracy*, edited by Jon Elster, 19-44. Cambridge: Cambridge University Press.
- Gargarella, Roberto. 1998. "Full Representation, Deliberation, and Impartiality." In *Deliberative Democracy*, edited by Jon Elster, 260-280. Cambridge: Cambridge University Press.
- Gaus, Gerald F. 2003. *Contemporary Theories of Liberalism*. London: Sage.
- Gaus, Gerald F. 2011. "A Tale of Two Sets: Public Reason in Equilibrium." *Public Affairs Quarterly* 25:305-325.
- Goldworth, A. 2002. "Bentham's Concept of Pleasure: Its Relation to Fictitious Terms." In *Bentham: Moral, Legal and Political Philosophy Volume 1*, edited by Gerald Postema. Ashgate: Dartmouth.
- Goodin, Robert. 1995. *Utilitarianism as a Public Philosophy*. Cambridge: Cambridge University Press.
- Goodin, Robert. 1996. "Institutionalising the public interest: The defence of deadlock and beyond." *American Political Science Review* 90: 331-343.
- Goodin, Robert. 2000. "Democratic Deliberation Within." *Philosophy & Public Affairs* 29: 81-108.
- Goodin, Robert. 2003. *Reflective Democracy*. Oxford: Oxford University Press.
- Goodin, Robert, and Christian List. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9: 207-306.
- Gunn, J.A.W. 1968. "Jeremy Bentham and the Public Interest." *Canadian Journal of Political Science* 1: 398-413.
- Gutmann, Amy. 1980 *Liberal Equality*, Cambridge: Cambridge University Press.
- Gutmann, Amy, and Dennis Thompson. 1996. *Democracy and Disagreement*. Harvard: Harvard University Press.

- Gutmann, Amy, and Dennis Thompson. 2004. *Why Deliberative Democracy?*. Princeton: Princeton University Press.
- Habermas, Jurgen. 1989. "Morality and ethical life: does Hegel's critique of Kant apply to discourse ethics?" *Northwestern University Law Review* 83: 38–53.
- Habermas, Jurgen. 1997. *Between Facts and Norms*. Translated by William Rehg. Cambridge: Polity Press.
- Habermas, Jurgen. 1999. "Between Facts and Norms: An Author's Reflections." *Denver University Law Review* 76: 937-942.
- Habermas, Jurgen. 2001. "'From Kant's 'Ideas' of Pure Reason to the 'Idealizing' Presuppositions of Communicative Action: Reflections on the Detranscendentalized 'Use of Reason'." In *Pluralism and the Pragmatic Turn: The Transformation of Critical Theory. Essays in honor of Thomas McCarthy*, edited by William Rehg and James Bohman. Cambridge, Massachusetts: MIT Press.
- Habermas, Jurgen. 2008. *Between Naturalism and Religion*. Edited by Ciaran Cronin. Cambridge: Polity.
- Halevy, Elie. 1952 *The Growth of Philosophic Radicalism*, translated by Mary Morris, corrected edn. London.
- Hardin, Russell. 2013. "The Free Rider Problem." In *The Stanford Encyclopaedia of Philosophy* (Spring 2013 Edition), edited by Edward N. Zalta. Online: <http://plato.stanford.edu/archives/spr2013/entries/free-rider>.
- Harrison, Ross. 1983. *Bentham: The Arguments of the Philosophers*. London: Routledge & Kegan Paul.
- Harrison, Ross. 1995. "Jeremy Bentham." In *The Oxford Companion to Philosophy*, edited by Ted Honderich, 85-88 Oxford: OUP.
- Harrison, Ross. 2003. Review of *Constitutional Limits and the Public Sphere: A Critical Study of Bentham's Constitutionalism*, by Oren Ben-Dor. *Utilitas* 15: 255-257.
- Her Majesty's Government (HMG), 2013. *Automatic enrolment opt out rates: findings from research with large employers*. London: TSO. Online: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/227039/opt-out-research-largeemployers-ad_hoc.pdf.
- Hume, L. J. 1978 "Revisionism in Bentham Studies." *The Bentham Newsletter* 1: 3-20.
- James, Michael R. 1981. "Public Interest and Majority Rule in Bentham's Democratic Theory." *Political Theory* 9: 49-64.

- James, Michael R. 2015. "Constituency Deliberation." *Political Research Quarterly* 68: 552-563.
- Johnson, James. 1998. "Arguing for Deliberation: Some Skeptical Considerations." In *Deliberative Democracy*, edited by Jon Elster, 19-44. Cambridge: Cambridge University Press.
- Kaino, Michihiro. 2008 "Bentham's Concept of Security in a Global Context: The Pannomion and the Public Opinion Tribunal as a Universal Plan." *Journal of Bentham Studies* 10: 1-29.
- Keen, Steve. 2001 *Debunking Economics*. London: Zed Books
- Kelly, Paul J. 1990 *Utilitarianism and Distributive Justice*, Oxford: Clarendon.
- Kim, Oliver, and Mark Walker. 1984 "The free rider problem: Experimental Evidence." *Public Choice* 43: 3-24.
- Klemp, Nathaniel J. 2011 "When rhetoric turns manipulative: disentangling persuasion and manipulation." In *Manipulating Democracy*, edited by Wayne le Cheminant and John M. Parrish, 59-86. New York: Routledge.
- Knight, Jack, and James Johnson. 1994. "Aggregation and Deliberation: On the Possibility of Democratic Legitimacy." *Political Theory* 22: 277-96.
- Kohn, M. 2000. "Language, power, and persuasion: Towards a critique of deliberative democracy." *Constellations* 7: 408-429.
- Larmore, Charles. 2003. "Public Reason." In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 368-93. Cambridge: Cambridge University Press.
- Lever, Annabelle. 2007. "Mill and the Secret Ballot: Beyond Coercion and Corruption." *Utilitas*, 19: 354-378.
- Lewis, William S. 2008. "War, Manipulation of Consent, and Deliberative Democracy." *The Journal of Speculative Philosophy* 22: 4.
- Lieberman, David. 1985. "Historiographical Review: From Bentham to Benthamism." *The Historical Journal* 28: 199-224.
- Lieberman, David. 2000. "Economy and Polity in Bentham's Science of Legislation." In *Economy, Polity and Society: British Intellectual History, 1750-1950*, edited by Richard Whatmore and Stefan Collini, 107-34. Cambridge: Cambridge University Press.

- Lieberman, David. 2014. "Bentham's Jurisprudence and Democratic Theory: An Alternative to Hart's Approach." In *Bentham's Theory of Law and Public Opinion*, edited by Xiaobo Zhai and Michael Quinn, 119-142. Cambridge: Cambridge University Press.
- List, Christian. 2006. "The Discursive Dilemma and Public Reason." *Ethics* 116: 362-402.
- List, Christian, and Philip Pettit. 2002. "Aggregating Sets of Judgments: An Impossibility Result." *Economics & Philosophy* 18: 89-110.
- Lister, Andrew. 2010. "Public Justification and the Limits of State Action." *Politics, Philosophy & Economics* 9: 151-75.
- Lovett, Frank, and Philip Pettit. 2009. "Neorepublicanism: A Normative and Institutional Research Program." *Annual Review of Political Science* 12: 11-29.
- Lyons, David. 1973. *In the Interests of the Governed*. Oxford: Oxford University Press.
- Lyons, David. 1994. *Rights, Welfare, and Mill's Moral Theory*. Oxford: Oxford University Press.
- Mack, Mary P. 1962. *Jeremy Bentham: An Odyssey of Ideas 1748-92*. London: Heinemann.
- Mackie, Gerry. 1998. "All Men Are Liars: Is Democracy Meaningless?" In *Deliberative Democracy*, edited by Jon Elster, 69-96. Cambridge: Cambridge University Press.
- Manin, Bernard. 1987. "On legitimacy and political deliberation." *Political Theory* 15: 338-368.
- Mansbridge, Jane. 2006. "Conflict and self-interest in deliberation." In *Deliberative democracy and its discontents*, edited by Samantha Besson, and Jose Luis Martí, 107-132. Aldershot: Ashgate.
- Mansbridge, Jane with James Bohman, Simone Chambers, David Estlund, Andreas Føllesdal, Archon Fung, Cristina Lafont, Bernard Manin and Jose Luis Martí. 2010. "The place of self-interest and the role of power in deliberative democracy." *The Journal of Political Philosophy* 18: 64-100.
- McCarthy, Thomas. 1994. "Kantian Constructivism and Reconstructivism: Rawls and Habermas in Dialogue." *Ethics* 105: 44-63.
- McKerlie, Dennis. 1988. "Egalitarianism and the Separateness of Persons." *Canadian Journal of Philosophy* 18: 205-25.
- Media Reform Coalition (MRC). 2014. *The elephant in the room: a survey of media ownership and plurality in the United Kingdom*. Online: <http://www.mediareform.org.uk/wp-content/uploads/2014/04/ElephantintheroomFinalfinal.pdf>.

Mill, John Stuart. 1838. "Essay on Bentham". *London and Westminster Review*, August. Reprinted 1962: *Utilitarianism; On Liberty; Essay on Bentham*. Edited by Mary Warnock, Cleveland: Meridian.

Mill, John Stuart 1951. "'Utilitarianism' and 'on Liberty' " In *The Collected Works of John Stuart Mill*. New York: E.P. Dutton.

Mill, John Stuart. 1970. *Subjection of Women*. Cambridge, Massachusetts and London: M.I.T. Press.

Miller, Nicholas R. 1983. "Pluralism and Social Choice." *The American Political Science Review* 77: 734-47.

Nagel, Thomas. 1987. "Moral Conflict and Political Legitimacy." *Philosophy & Public Affairs* 16: 215-40.

Nelson, William. 1989. "Evaluating the Institutions of Liberal Democracy." In *Politics and Process*, edited by Geoffrey Brennan and Loren E. Lomasky. Cambridge: Cambridge University Press.

Neumann, Michael. 1989. "Co-Ordination and Utility." *The Philosophical Quarterly* 39: 66-74.

Niemeyer, Simon. 2011. "The Emancipatory Effect of Deliberation: Empirical Lessons from Mini-Publics." *Politics & Society* 39: 103-140.

Niemeyer, Simon. 2004. "Deliberation in the Wilderness: Displacing Symbolic Politics." *Environmental Politics* 13: 347 – 372.

Niemi, Richard G., and Kent Jennings. 1991. "Issues and Inheritance in the Formation of Party Identification." *American Journal of Political Science* 35: 970-978.

Niesen, Peter. 2011. "Bentham and Dumont on Parliamentary Deliberation: Un petit Mappemonde de Chaos." Paper presented at the New Directions in Bentham Studies Conference, 9 December 2011, at the Bentham Project, University College London. (References are taken from this original version of the paper in English.). The text has also been published in French: Niesen, Peter. 2009. "Une petite mappemonde du chaos: La délibération parlementaire chez Bentham et Dumont," in *Bentham et la France: fortune et infortunes de l'utilitarisme*, edited by Emmanuelle de Champs and Jean-Pierre Cléro, 129-41. Oxford: SVEC.

Nussbaum, Martha. 2011. "Perfectionist Liberalism and Political Liberalism." *Philosophy & Public Affairs* 39: 3-45.

Ober, Josiah. 2011. "Weighted Expertise Aggregation: An Aristotelian middle way for epistemic democracy." *Princeton/Stanford Working Papers in Classics*. Online:

<https://www.uio.no/english/research/interfaculty-research-areas/democracy/news-and-events/events/seminars/2011/papers-yale-2011/Yale-Ober.pdf>

OFCOM (The Office of Communications), 2015. "News Consumption in the UK." Online: <http://stakeholders.ofcom.org.uk/market-data-research/other/tv-research/news-2015/>.

O'Flynn, Ian. 2010. "Deliberating about the Public Interest." *Res Publica* 16: 299-315.

Ogden, Charles K. 1931. "Bentham's Theory of Language." *Psyche* 12: 3-43.

Ogden, Charles K. 1932. "Bentham's Theory of Fictions." London: Paul, Trench, Trubner.

Ogden, Charles K. 1993. "Bentham's Philosophy of as-If." In *Bentham: Critical Assessments, Volume 3*, edited by Bhikhu C. Parekh. London: Routledge.

O'Neill, Onora. 1986. "The Public Use of Reason." *Political Theory* 14: 523-51.

O'Neill, Onora. 1997. "Political Liberalism and Public: A Critical Notice of John Rawls, Political Liberalism." *Philosophical Review* 106: 411-428.

Page, Benjamin I. and Robert Y. Shapiro. 1993. "The Rational Public and Democracy." In *Reconsidering the Democratic Public*, edited George E. Marcus and Russell L. Hanson. University Park, Pennsylvania: Pennsylvania State University Press.

Park, Joseph, H. 1931. "England's Controversy over the Secret Ballot." *Political Science Quarterly* 46: 51-86.

Pateman, Carol. 1970. *Participation and democratic theory*. Cambridge: Cambridge University Press.

Pease-Watkin, Catherine. 2003. "Bentham's Panopticon and Dumont's Panoptique." *Journal of Bentham Studies* 6: 1-8.

Pettit, Philip. 2004. "The common good." In *Justice and democracy: Essays for Brian Barry*, edited by Keith Dowding, Robert Goodin, and Carol Pateman, 150-169. Cambridge: Cambridge University Press.

Petracca, Mark, P. 1991. "The Rational Choice Approach to Politics: A Challenge to Democratic Theory." *The Review of Politics* 53: 289-319.

Phillips, Anne 1995. *The Politics of Presence*. Oxford: Oxford University Press.

Posner, Richard. 2003. *Law, Pragmatism and Democracy*. Cambridge, MA: Harvard University Press.

Posner, Richard, 2004. "Smooth Sailing." *Legal Affairs* January/ February. Online: https://www.legalaffairs.org/issues/January-February-2004/feature_posner_janfeb04.msp.

Postema, Gerald, J. 1983. "Facts, Fictions and the Law: Bentham on the Foundations of Evidence." *Archiv für Rechts- und Sozialphilosophie* 16: 37-64.

Postema, Gerald, J. 1986. *Bentham and the Common Law Tradition*. Oxford: Oxford University Press.

Postema, Gerald J. 2014. "The Soul of Justice: Bentham on Publicity, Law, and the Rule of Law." In *Bentham's Theory of Law and Public Opinion*, edited by Xiaobo Zhai and Michael Quinn, 40-62. Cambridge: Cambridge University Press.

Przeworski, Adam. 1998. "Deliberation and Ideological Domination." In *Deliberative Democracy*, edited by Jon Elster, 140-160. Cambridge: Cambridge University Press.

Quinn, Michael. 2008. "A Failure to Reconcile the Irreconcilable? Security, Subsistence and Equality in Bentham's Writings on the Civil Code and on the Poor Laws." *History of Political Thought* 29: 320-343.

Quinn, Michael. 2011. "Post-Modern Moments in the Application of Empirical Principles: Power, Knowledge and Discourse in the Thought of Jeremy Bentham Versus Michel Foucault." *Revue d'études benthamiennes* 8. Online: <https://etudes-benthamiennes.revues.org/245>

Quinn, Michael. 2014. "Popular Prejudices, Real Pains: What Is the Legislator to Do When the People Err in Assigning Mischief?" In *Bentham's Theory of Law and Public Opinion*, edited by Xiaobo Zhai and Michael Quinn, 63-89. Cambridge: Cambridge University Press.

Quong, Jonathan. 2004. "The Scope of Public Reason." *Political Studies* 52: 233-50.

Quong, Jonathan. 2008. "Three Disputes About Public Justification: Commentary on Gaus and Vallier." *Public Reason* 1-17. Online: <http://www.publicreason.net/wp-content/PPPS/Fall2008/JQuong1.pdf>

Rawls, John. 1982. "Social Unity and Primary Goods." In *Utilitarianism and Beyond*, edited by Amartya Sen, 159-85. Cambridge: Cambridge University Press.

Rawls, John. 1985. "Justice as Fairness: Political Not Metaphysical." *Philosophy and Public Affairs* 14: 223-51.

Rawls, John. 1997. "The Idea of Public Reason Revisited." *The University of Chicago Law Review* 64: 765-807.

Rawls, John. 1999. *A Theory of Justice*. Cambridge, Massachusetts: Harvard University Press.

Rawls, John. 2005. *Political Liberalism*. New York: Columbia.

Raz, Joseph. 1990. "Facing Diversity: The Case of Epistemic Abstinence." *Philosophy & Public Affairs* 19: 3-46.

- Reidy, David, A. 2000. "Rawls' Wide View of Public Reason: Not Wide Enough." *Res Publica* 6: 49-72.
- Richardson, Henry, S. 2003. *Democratic Autonomy: Public Reasoning About the Ends of Policy*. New York and Oxford: Oxford University Press.
- Ridge, Michael. 1998. "Hobbesian Public Reason." *Ethics* 108: 538-68.
- Riker, William, 1982 *Liberalism against Populism*. San Francisco: W.H. Freeman.
- Riley, Jonathan. 1990. "Utilitarian Ethics and Democratic Government." *Ethics* 100: 335-348.
- Riley, Jonathan. 2010 "Justice as higher pleasure." In *John Stuart Mill: Thought and Influence- The Saint of Rationalism*, edited by Georgios Varouxakis and Paul Kelly, 99-129. London and New York: Routledge.
- Rosen, Fred. 1983. *Jeremy Bentham and Representative Democracy*. Oxford: Oxford University Press.
- Rosen, Fred. 1997 "Utilitarianism and the Punishment of Innocent: The Origins of a False Doctrine." *Utilitas* 9: 23-37.
- Rosen, Fred 2003 *Classical Utilitarianism from Hume to Mill*. London and New York: Routledge.
- Rosen, Fred. 2014. "Introduction." In *Bentham's Theory of Law and Public Opinion*, edited by Xiaobo Zhai and Michael Quinn, 1-6. Cambridge: Cambridge University Press.
- Scarre, Geoffrey. 2007. *Mill's 'On Liberty', A Reader's Guide*. New York: Continuum.
- Schofield, Philip. 1999. "Political and Religious Radicalism in the Thought of Jeremy Bentham ." *History of Political Thought* 20: 272-91.
- Schofield, Philip. 2006. *Utility and Democracy*. Oxford: Oxford University Press.
- Schofield, Philip. 2010. "Jeremy Bentham and H.L.A. Hart's "Utilitarian Tradition in Jurisprudence."" *Jurisprudence* 1: 147-67.
- Schofield, Philip. 2014. "Jeremy Bentham on Taste, Sex, and Religion" In *Bentham's Theory of Law and Public Opinion*, edited by Xiaobo Zhai and Michael Quinn, 90-118. Cambridge: Cambridge University Press.
- Schofield, Philip. 2014. "A Defence of Jeremy Bentham's Critique of Natural Rights." In *Bentham's Theory of Law and Public Opinion*, edited by Xiaobo Zhai and Michael Quinn, 208-230. Cambridge: Cambridge University Press.

- Schwartzberg, Melissa. 2007a. *Democracy and Legal Change*. New York and Cambridge: Cambridge University Press.
- Schwartzberg, Melissa. 2007b. "Jeremy Bentham on Fallibility and Infallibility." *Journal of the History of Ideas* 68: 562-86.
- Schwartzman, Micah. 2004. "The Completeness of Public Reason." *Politics, Philosophy & Economics* 3: 191-220.
- Schwartzman, Micah. 2010. "The Sincerity of Public Reason." *Journal of Political Philosophy* 19: 375-98.
- Sears, David, O. 2001. "The Role of Affect in Symbolic Politics." In *Citizens and Politics: Perspectives from Political Psychology*, edited by James H. Kuklinski, 14-40. New York: Cambridge University Press.
- Shafe, James, R. 2014. "Challenges for a revised view of Bentham on public reasoning." *Revue d'études benthamiennes*, 13. Online: <https://etudes-benthamiennes.revues.org/761>
- Shapiro, Robert. 1998. "Public Opinion, Elites, and Democracy." *Critical Review* 12: 501-28.
- Sidgwick, Henry. 1874. *The Methods of Ethics*. London: MacMillan.
- Sigot, Nathalie. 2009. "Bentham's Way to Democracy." *History of European Ideas* 35: 112-15.
- Stears, Marc. 2010. *Demanding Democracy: American Radicals in Search of a New Politics*. Princeton and Oxford: Princeton University Press.
- Stokes, Susan. 1998. "Pathologies of Deliberation." In *Deliberative Democracy*, edited by Jon Elster, 123-39. Cambridge: Cambridge University Press..
- Sunstein, Cass, R. 2003. *Why Societies Need Dissent*. Cambridge, Massachusetts and London: Harvard University Press.
- Sunstein, Cass, R. 2011. "Empirically informed Regulation." *The University of Chicago Law Review* 78: 1349-1429.
- Sunstein, Cass, R.. 2015. "Fifty Shades of Manipulation." *Journal of Behavioral Marketing*, Forthcoming; subsequently published in amended form: Sunstein, Cass, R. 2016. "Fifty Shades of Manipulation." *Journal of Marketing Behavior* 1:213-244.
- Talisse, Robert. 2009. *Democracy and Moral Conflict*. Cambridge: Cambridge University Press.
- Taylor, A.J.P. 1972. *Beaverbrook*. London: Hamilton

Thaler, Richard, .H., and Cass .R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth and Happiness*. London: Penguin.

Turnbull, Nick. 2008. "Dewey's Philosophy of Questioning: Science, Practical Reason and Democracy." *History of the Human Sciences* 21: 49-75.

Von Rautenfeld, Hans. 2004. "Charitable Interpretations: Emerson, Rawls, and Cavell on the Use of Public Reason." *Political Theory* 32: 61-84.

Waldron, Jeremy. 1987. "Theoretical Foundations of Liberalism." *The Philosophical Quarterly* 37: 127-150.

Waldron, Jeremy. 1999. *Law and Disagreement*. Oxford: Oxford University Press.

Wall, Steven. 2006. "Debate: Democracy, Authority, and Publicity." *The Journal of Political Philosophy* 14: 85-100.

West, Henry, R. 2004. *An Introduction to Mill's Utilitarian Ethics*. Cambridge: Cambridge University Press.

Wolff, Jo. 1993. "Hume, Bentham, and the Social Contract." *Utilitas*, 5: 87-90.

Young, Iris, Marion. 1990. *Justice and the politics of difference*. Princeton: Princeton University Press.