# Path Dependence in Deliberation over Preference Rankings

April 1, 2020


**Abstract**

In this paper we study the so-called anchoring effect, a form of path-dependence that occurs in group deliberation when the opinion of those who speak early, regardless of their perceived or real reliability, carry substantially more weight than the opinion of those who speak later. We show, first, that anchoring can occur even among fully rational agents deliberating on preference rankings, hence complementing the results obtained in [26]. We then pitch anchoring against one well-known positive aspect of deliberation, namely that it helps avoiding irrational group preferences through the creation of single-peaked profiles. We show that although deliberation performs well in eliminating irrational group preferences, in a significant number of cases this positive effect is off-set by the emergence of path-dependence. We do this in a minimal, computational model of deliberation where participants repeatedly exchange their views on a given matter and update their own accordingly. We finally briefly discuss

the consequences of this finding on the question of collective and individual responsibility.

Deliberative approaches to collective decision making have pointed out a number of advantages of group deliberation. It should allow for a more widespread access to relevant evidence, arguments and information, and as such it should increase the possibility of more informed decision by the participants, and reduces the chance of voter manipulation [9, 18, 19, 34]. Others argue that the process of sharing arguments, as is the case in deliberation, allows for a more reliable scrutiny of arguments and will thus make it more likely for the decision makers to recognize shortcomings, fallacies and mistakes [8, 20, 35, 12]. Epistemic efficiency of group deliberation has also been observed in theoretical models [27]. Not only can deliberation help reach these positive outcomes, it has also been argued to be an effective mean of avoiding problematic features of voting, like Condorcet cycles or the doctrinal paradox [32, 22, 45]. Dryzek and List [14] even claim that deliberation provides incentives not to misrepresent one's preference, and so avoids one important form of strategising.

The literature in social psychology and formal epistemology is, however, more pessimistic regarding the possible outcomes of deliberation. Authors there have been prompt to point out that the improved performance of "real" deliberation is highly contingent on how it is implemented: Deliberation can lead to polarization [5] or to spurious unanimity due to pluralistic ignorance [46] or so-called hidden profiles [47]. It can also give unfair advantage to those with good argumentative skills, and be prone to various form of path-dependency coming from strategic considerations or more fundamentally the way participants take the opinions of others into account [26, 7].

This paper focuses on the latter type of bias, namely a type of path-dependency that results in an unjustifiably high weight to the opinions expressed earlier in the course of deliberation. This is an instance of a more general phenomenon named *anchoring* that has been extensively studied in social psychology and collective decision making, [49, 40, 6]. According to [49], the anchoring effect is the disproportionate tendency of decision makers to be biased toward an initially presented information.

Instances of anchoring are resilient and hard to avoid, and equally difficult to explain. Following Tversky and Kahneman [49], anchoring effect has been studied extensively and illustrated in numerous instances in a variety of domains [3, 6, 17, 37, 38, 39, 41, 42, 43]. There are several proposals in the literature for the underlying mechanisms that produce anchoring. Tversky and Kahneman [49] suggest that anchoring occur as people fail to make correct adjustments to an initially presented value. Another proposal for the emergence of anchoring, that has empirical support, is based on confirmatory hypothesis testing [6, 41]. This proposal suggests that the effect emerges as the initially presented information make the decision makers prone to consider it as a plausible hypothesis and seek, focus on, and remember the evidence that confirms it rather than those that reject it. See [24] for an overview of studies on different instances of anchoring as well as an overview of different proposals for the emergence of anchoring.

These proposals explain anchoring in terms of the decision makers' capacity to process the evidence correctly. Other instances of anchoring, particularly in deliberative decision making, may result from the specific deliberation procedure that is adopted. Such cases of anchoring can also emerge for groups whose members are individually rational in weighting the evidence they receive and adopt a rationally justified strategy for the process of adjustment. Hartmann and Rafiee Rad [26] provided some insight into the underlying mechanisms that can result in the emergence of anchoring in such groups. They showed that fully rational (Bayesian) agents with no shortcomings in processing the evidence can fall prey to anchoring. As such anchoring can be also seen as a structural bias of deliberation, not necessarily the result of some failure of the individual participants. This observation, however, only bears on exchange and updates of probabilistic credences, and thus leave open the case where the goal of deliberation is the exchange of individual preferences. This is the case we cover in this paper, and we show that fully rational agents can fall prey to anchoring in that case too.

A proper philosophical assessment of anchoring requires, however, weighing it against purported positive aspects of deliberation. In other words, whether anchoring is philosophically consequential depends on the positive aspect it is pitched against. Anchoring, for instance, introduces an obvious possibility for strategising, not in terms of misrepresentation of one's preferences but in terms of timing of interventions. To the extend that the order in which the participant speak might not reflect their expertise or relative importance, anchoring also threatens truth-tracking or the formation of better-informed judgments. By the same token, it might give an unfair advantage to some group members on the basis of their positioning within the deliberation process.

A full philosophical assessment of anchoring goes beyond the scope of a single paper, but as a first step here we investigate the relation between anchoring and the elimination of cyclic and intransitive group preferences. In particular we ask to what extent is the elimination of cycles affected by anchoring. We show that although deliberation eliminates intransitive profiles to a large extent, in a significant number of cases these positive effects of deliberation are off-set by the emergence of path-dependence. We do so using a computational, minimalistic model of social influence. In this model the participants repeatedly exchange their opinions and update their preferences accordingly, following a rational preference update mechanisms.

As we will argue in Section 3, how this positive and negative effects fair with respect to each other, and with respect to evaluation of deliberation as a process of collective decision making depends on the context of the deliberation and the way that path-dependence correlates to the agents expertise.

# 1 The Model

## 1.1 Informal Description

The model we use in this paper has been developed in [48] to study the elimination of voting cycles and creation of single-peaked preferences in deliberation. Here we use it to measure anchoring.

The model belongs to the category of what List [31] calls models of "deliberation as preference transformation". It focuses exclusively on how the participants change their preferences when learning those of others. In other words, this is a model of social influence that abstracts away from many features of a richer understandings of deliberation. It does not captures the exchange of reasons and arguments, nor does it explicitly restrict the participants to express views that they take to be universalizable or expressions of the common good [2]. Nothing prevent, however, to interpret the preferences of the participants in terms of generalizable interests or of their subjective views of the common good. The model also leaves out strategic considerations cf. [30]. In our model, the participants do not have specific agendas, nor try to maintain their status or reputation, or even to get to what they take to be the correct view.

Deliberation, in this model, proceeds as follows. The participants enter deliberation by holding certain preferences, represented by either weak or strict rankings, over a over a given set of alternatives. One participants then publicly and sincerely announce thier full preference ranking in a fixed order. The other participants then update their ranking accordingly, using a distance-minimization rule. They do so *after each individual announcement*, as opposed to wait until everyone has expressed their opinions. That is, they update their preference rankings by minimizing a given distance measure between the preference ranking just announced and their old ranking. We allow the participant to be somewhat biased towards their own preferences, i.e. to give more weight to their own opinion than the opinion of others.

We use and compare three well-known measures between orderings — the Kemeny–Snell [29], the Cook–Seiford [11] and the Duddy–Piggins [15] distances — so that our results do not hinge on the particulars of one of them. Our agents deliberate for a fixed number of rounds, after which we check the measure anchoring and proximity to single-peakedness in the resulting preference profile. Anchoring is measured as follows: we first take the average distance between the ranking of each group member at the end of the deliberation to each of the rankings submitted by the agents at the start, and on the basis of that calculate the relative frequency of cases where the resulting profile is closest to the initial ranking of the first speaker compared to other group members. Proximity to single-peakedness, on the other hand, is calculated in the same way as [44, 32]: it is the relative size of the largest sub-group that is single-peaked with respect to one ranking.

## 1.2 Formal details

A group $N$ of $n$ individuals enters deliberation, each holding a preference ordering over a given set of alternatives, $a_1, \ldots, a_j$. For computational reasons, the results reported here cover only the case of three alternatives. Let $\mathcal{R}$ be the set of all possible rankings over the set of three alternatives. With strict preference rankings over three alternatives, there are 6 different rankings in $\mathcal{R}$, and 13 if we allow for weak rankings as well.

A *round* of sharing consists of $n$ steps, one for each participant. At step $i$ $(1 \leq i \leq n)$ participant $i$ announces her ranking and the rest of the group update their opinion accordingly. Each participant announces her preferences exactly once per round. This process continues for a fixed number of rounds.

Let $R = \langle r_1, \ldots, r_n \rangle$ be a preference profile. After the $i$-th member speaks, the rankings are updated to $\langle r'_1, \ldots, r'_n \rangle$ where $r'_j$ is the ranking for which

$$\sqrt{\text{bias}_i d(r_i, r'_j)^2 + \text{bias}_j d(r_j, r'_j)^2}$$

is minimal, for a given distance measure $d$. In the strict case, the $r'_j$ are picked from among the domain of strict rankings, and analogously in the weak case.

The parameters $\text{bias}_j$ and $\text{bias}_i$ represent, respectively, the biases each agent has towards her own and others' preferences. This allows us to capture different degrees to which the participants are open to changing their mind upon learning the preferences of others. We take these two parameters to be in the [0,1] interval, with $bias_j = (1 - bias_i)$. When $bias_j = bias_i$ then $j$ views $i$ as an equal peer. She assigns as much weight to his opinion as to her own. If $bias_j > bias_i$ then $j$ is biased towards her own opinion, with stronger biases for larger differences between $\text{bias}_j$ and $\text{bias}_i$. If $bias_j = 1$ then $j$ does not take $i$'s opinion into account at all. We make the admittedly strong simplifying assumption that all agents are biased in the same way towards themselves and all others. So we keep the values $\text{bias}_j$ and $\text{bias}_i$ constant for all $i$ and $j$ in the formula above.

For the distance parameter $d$ we use the Kemeny–Snell [28, 29], the Cook–Seiford [10], and the Duddy–Piggins [15] measures. Their precise definitions and examples are in the Appendix. These measures, although very close to one another, differ mostly in the underlying notion of "betweenness" that they use.

Deliberation continues for a fixed number $w$ of rounds, after which we check for the emergence of anchoring by observing whether or not the final profile is closest to the initial ranking submitted by the first speaker. We do this for both the weak and strict preferences.

When the domain is restricted to strict preferences, the agents choose between six possible ordering of the alternatives. For small groups ($n \leq 6$), then, the probability that the final profile is closest to that of the first speaker should on average be around $1/n$. For larger groups ( $n > 6$) this probability should on average be 1/6 or approximately 16%. The emergence of anchoring is marked by probabilities higher than what is expected on average.

We then calculate the probability of anchoring as the proportion of cases where the final profile is anchored in a sample of 1000 simulations. For every simulation we also calculate the proximity to single-peakedness [32], as the largest portion of the final profile that is single-peaked with respect to some structuring dimension.

This gives an indication of how much the deliberation procedure has lead the group towards single-peaked preferences. We then calculate the average proximity to single-peakedness over the 1000 simulations.

# 2 Results

We start by looking at deliberation with strict rankings only. The restriction is intended for both the input and the output of the deliberation procedure. Then we move to weak rankings, i.e., allowing for indifference. So, when we talk about deliberation with strict rankings we mean deliberation in which the agents start with a strict ranking and through out the deliberation they change their opinion by moving between strict rankings. Consequently the output profiles will also be linear orders. For the case of weak rankings, on the other hand, the starting profile, intermediate steps and the final profile range through preorders.

## 2.1 Deliberation with Strict Rankings Only

Our first results show the probability of anchoring in the deliberation (with each of the three distance measures) for increasing group sizes. Figure 1 shows the relative frequency of anchoring in a group of deliberating agents with a low bias towards their own opinion. Recall that the low bias indicates that the participants are relatively open to change their opinion in light of what they learn from their fellow group members. The graph shows the relative frequency of cases where the final profile after the deliberation is closest to the initial ranking submitted by the first speaker, as the size of the group increases. The results are calculated from 1000 simulations for each data point.

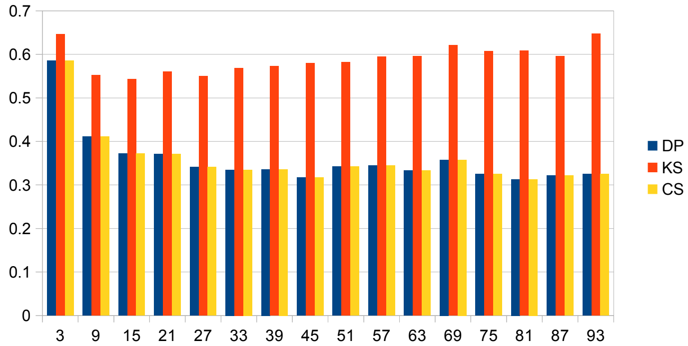


Figure 1: Relative frequency of anchoring for a group of agents with a bias of 0.6 towards themselves, as the number of participants increases.

Our simulations show a high frequency of anchoring. For the Kemeny-Snell measure, the final profiles is closest to the initial preferences of the first speaker up to about 4 times more than what is expected (16%). For Duddy-Piggins and Cook-Seiford, anchoring is at least twice of what is on average expected. The frequency of anchoring, however, seems to be more or less stable over increasing group sizes.

Figure 2 shows the same frequency as in Figure 1 but for a group of participants with higher biases towards their own opinion. As expected, increasing agents' bias towards their own opinion reduces their move towards the opinions announced by their fellow agents. This naturally decreases the frequency of anchoring.

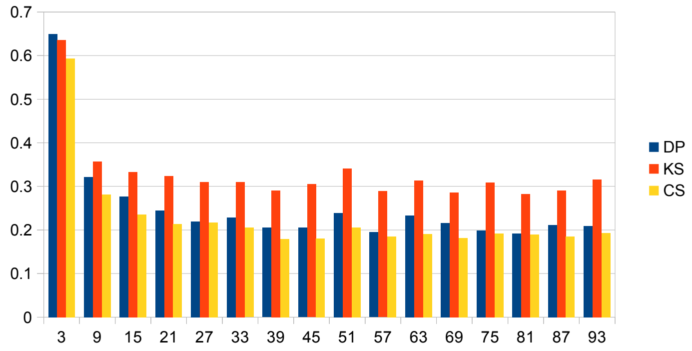



Figure 2: Probability of anchoring for a group of agents with a bias of 0.8 towards themselves, as the number of participants increases.

As Figure 2 shows, however, even for higher biases the anchoring effect still persists though with a lesser (but still noticeable) intensity. For Kemeny-Snell measure this is around twice as high as what is expected on average and for Duddy-Piggins and Cook-Seiford we see an increase of at least 40% of what was expected on average.

Comparing Figure 1 and Figure 2 suggests a correlation between the probability of anchoring and the agents' openness to change their mind. The next two plots cash out this correlation in more detail. Figure 3 shows the frequency of anchoring for a group of size 33 as a function of agents' bias towards their own opinion. In simpler terms, Figure 3 depicts the frequency of anchoring as a function of how open the deliberating agents are to changing their mind.

For all three measures the frequency of anchoring seems to be stable for biases up to 0.7 and then decreases for higher biases. Figure 4 shows the same probability as Figure 3 but for a larger group. As also implied by Figures 1 and Figure 2 increasing the group size does not seems to play any significant role in frequency

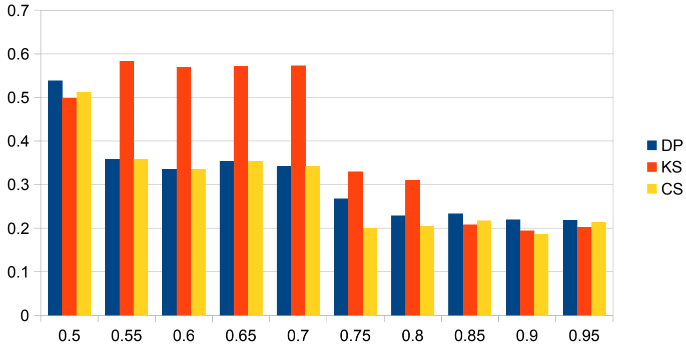



Figure 3: Probability of anchoring for a group of size 33, as the agents bias towards themselves increases.

of anchoring.

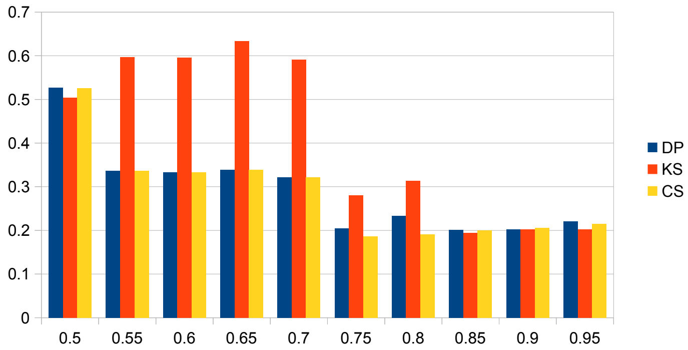



Figure 4: Probability of anchoring for a group of size 63, as the agents bias towards themselves increases.

## 2.2 Deliberation with Weak Preferences

We now move to weak preference rankings. Here for small groups ($n < 13$) the frequency that the final profile is closest to the initial preference of the first speaker is on average expected to be approximately $1/n$. For larger groups ($n > 13$) this probability is on average expected to be around $1/13$ or approximately 8%. As before the anchoring emerges when we observe probabilities higher than what is expected on average.

Figure 5 shows the frequency of anchoring in a group of agents who are, to a large extent open to change (low bias, analogous to Figure 1). The graph shows the frequency of anchoring as the size of the group increases. As before the results are calculated from 1000 simulations (for each data point).

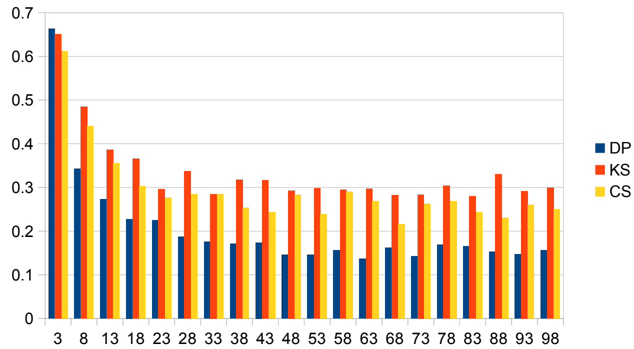



Figure 5: Probability of anchoring for a group of agents with bias of 0.6 towards themselves, as the number of participants increases, weak preferences.

As the graph shows anchoring in deliberation with non-strict rankings follow a similar patterns as for strict preferences. Namely that deliberation is heavily anchored by the first speaker in smaller groups and the probability of anchoring decreases and becomes more or less stable over the group size as the size of the group increases.

Figure 6 shows the frequency of anchoring as the function of agents openness to change, for a group of size 33 with weak preferences. Here anchoring seems to follow a decreasing pattern as the bias increases.

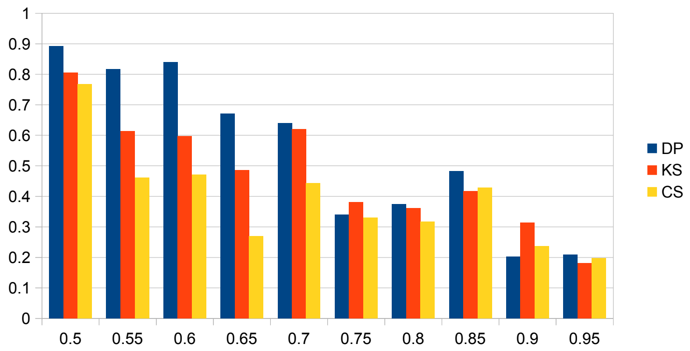
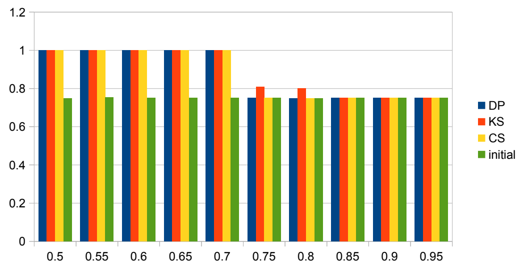


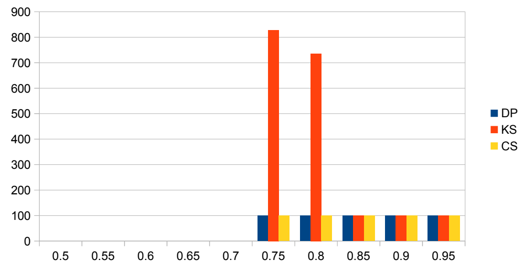

Figure 6: Probability of anchoring for a group of 33 as the bias increases, weak preferences.

## 2.3 Anchoring vs Single-peakedness

We now study how anchoring in deliberation interacts with the creation of single-peaked preferences, and hence the avoidance of irrational group preferences or more generally of Arrowian impossibility results. Deliberation, in our model, indeed create single-peaked profiles, up to a bias of 0.7, and in doing so eliminates all cyclic and intransitive profiles. Figure 7-(a) shows the increase in proximity to single-peakedness, with strict preferences, for a group of size 33 as the agents bias towards their own opinion increases and Figure 7-(b) shows the percentage of cyclic profiles still present after the deliberation.




(a) Proximity to Single-Peakedness, Group of size 33 (strict preferences)




(b) Percentage of the cyclic profiles still present after deliberation, Group size 33 (strict preferences)

Figure 7:

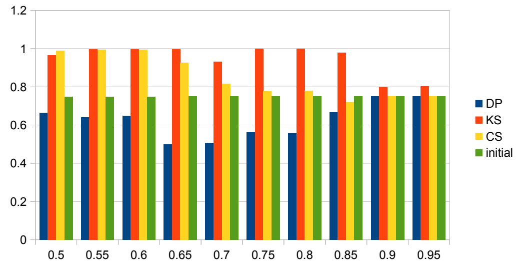
As the results suggest, for strict rankings up to a bias of about 0.7 deliberation creates single-peakedness in every instance and hence all cyclic or irrational profiles are eliminated up to this bias. For higher biases the agents become more resistant to changing their opinion. Interestingly, Kemeny-Snell measure seems to create cyclic profiles for biases from 0.7 to 0.85.[1]

The corresponding notion of single-peakedness for weak preferences is that of single-plateuness (see Appendix for details). Similar to the case of strict preferences, creation of single-plateau profiles ensures the elimination of intransitive profiles and thus avoids Arrowian impossibility results.

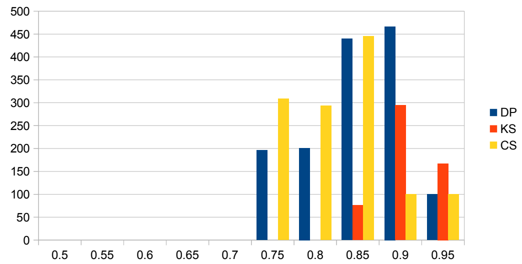
As Figure 8(a) shows deliberation with weak preferences, fosters creation of single-plateau preferences. For up to a bias 0.85 Kemeny-Snell creates single-plateau preferences in every instance. For Cook-Seiford we see significant increase in proximity to single-plateauness up to a bias of 0.7. For higher biases, the agents hardly revise their opinion and as such the proximity to single-plateauness remains

[1]Notice that for strict preferences Kemeny-Snell measure creates cycles for biases between 0.75 and 0.85 and for weak preferences Duddy-Piggins creates intransitive profiles for biases between 0.8 and 0.9. This observation is discussed in more details in [48].

unchanged for both measures. Duddy-Piggins, however, seems to systematically decrease single-plateauness in the profile up to a bias of 0.9. Nevertheless, as Figure 8(b) shows all three measures completely eliminate intransitive profiles up a bias of 0.7. For biases higher than 0.75 for Duddy-Piggins and Cook-Seiford and 0.9 for Kemeny-Snell, deliberation creates intransitive profiles.




(a) Proximity to Single-Plateauness, Group of size 33 (weak preferences)

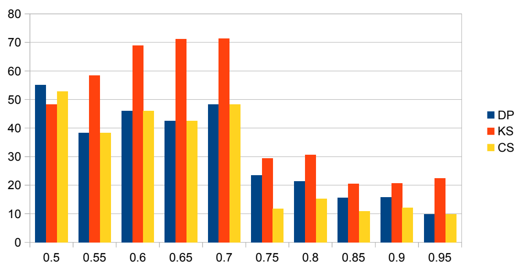



(b) Percentage of the intransitive profiles still present after deliberation, Group size 33 (weak preferences)

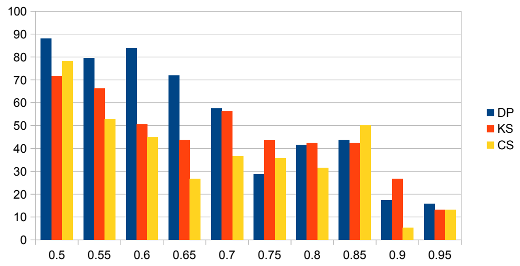
Figure 8:

To investigate how creation of single-peakedness and elimination of cyclic and intransitive profiles are affected by the emergence of anchoring, we next plot the percentage of cyclic and intransitive profiles for which the final profile is anchored by the first speaker. Figure 9 shows this for a group of size 33. Recall that on average this is expected to be approximately 16% and 8% (i.e. 1/6 and 1/13) for strict and weak preferences respectively and a probability higher than this shows the emergence of anchoring.

Recall from Figure 7 that up to a bias of 0.7 deliberation significantly increases the proximity to single-peakedness and by doing so completely eliminates cyclic profiles. As Figure 9(a) shows, however, a large percentage of the initially cyclic




(a) Percentage of the cyclic profiles that end up anchored, Group size 33 (strict preferences)




(b) Percentage of the intransitive profiles that end up anchored, weak preferences

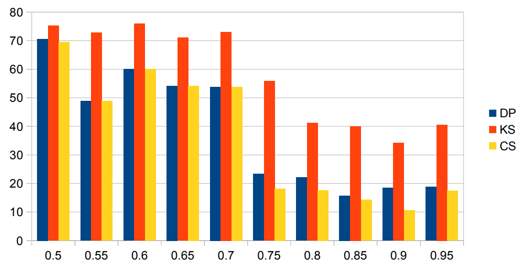
Figure 9:

profiles profiles that end up single-peaked are also anchored by the first speaker (up to 70 % for Kemeny-Snell and around 40-50 % for Duddy-Piggins and Cook-Seiford).

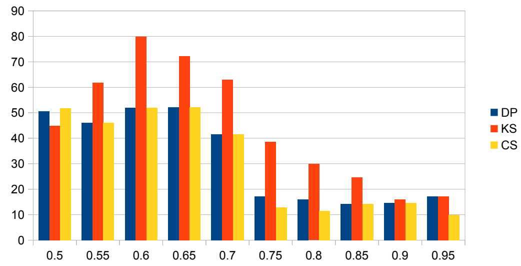
Similarly Figure 9(b) shows that for weak preferences and for biases up to 0.7, a considerable percentage of cases where deliberation has been successful in eliminating intransitive profiles, do suffer from anchoring (around 60-85 % for Duddy-Piggins, 40-70% for Kemeny-Snell and around 25-75 % for Cook-Seiford). It is worth noting that for low biases, Kemeny-Snell has a stronger off-setting effect for strict preferences while Duddy-Piggins and Cook-Seiford appear to have stronger off-setting effect for weak preferences.

Figure 10 shows the percentage of cyclic strict profiles for which the final (non-cyclic) profile is anchored, as a function of agents bias towards their own opinion for small, medium and large sized groups.

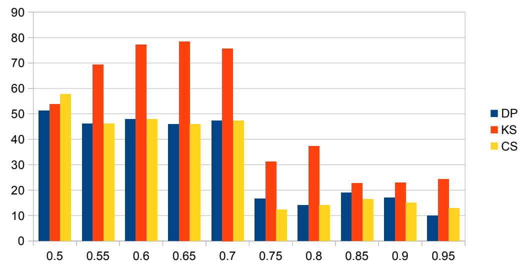
These imply a large overlap between the cases where deliberation proves beneficial by eliminating cyclic and intransitive profiles and the cases where it results in path-dependence and the emergence of anchoring. As Figure 10 suggests, the observation of Figure 9 for the overlap between elimination of cycles and emergence of anchoring persists over different group sizes.




(a) Percentage of the cyclic profiles that end up anchored, strict preferences, Groups size 9




(b) Percentage of the intransitive profiles that end up anchored, strict preferences, Group size 39




(c) Percentage of the intransitive profiles that end up anchored, strict preferences, Group Size 73

Figure 10:

# 3 Discussion

## 3.1 Anchoring as a structural bias in deliberation

Anchoring effect, though easily observable, has proved difficult to explain [40]. There are several proposals in the literature for the underlying causes of anchoring but consensus is yet to be reached. One proposal [49] explains it as the result of decision makers' failure to correctly adjust for the initial information that they receive while others explain it as the result of confirmatory hypothesis testing [6, 41].

These proposals, however, root the emergence of anchoring in the agents capacity to correctly process the information. As such the burden for such undesirable consequences is laid upon the deliberative agents rather than the deliberation procedure itself. As we argue now, the results presented above show that anchoring is not necessarily the participant's fault: fully rational agents with no cognitive or computational limitations can fall prey to that bias.

Clearly, the participants in our model have no cognitive or computational limitations when it come to revising their preferences in view of those of others. They always correctly and completely make the calculations required by minimizing the given distances between rankings. There is, for instance, no random "noise" or shocks in the model, which would capture mistakes or unexpected changes in the participants' preferences.

More importantly, the participants follow what can be seen as rational preference change policies. To see this, recall first that these policies are captured by the minimization rules, given either the Kemeny-Snell, the Duddy-Piggins or the Cook-Seiford distance measure. Now, in the process of deliberation, each participant repeatedly faces an aggregation problem. They should decide how to aggregate their own opinion with the opinion that was just expressed by another group member. Minimizing distance according to these three measures can be viewed as a rational approach to solve this aggregation problem. These preference change policies indeed satisfy all classical Arrowian postulates (for all the three measures), except for the somewhat controversial (see, e.g., [28]) axiom of

Independence of Irrelevant Alternatives (IIA). Furthermore, at the axiomatic level none appear noticeably more coherent than the others. Indeed, axiomatizations of each of these measures are almost identical. Kemeny–Snell and Duddy–Piggins are very similar and differ from each other only in the way they treat logical redundancies. The only difference between Cook–Seiford and Kemeny–Snell or Duddy–Piggins is in the way that notion of "betweenness" is defined for these measures (See the appendix for more details). Although this can lead to difference in some update scenarios, these differences are not normatively significant.

Distance minimization, in our model, is of course tilted by the bias parameters, which allow the participants to be more or less open to changing their view in light of the opinions expressed by others. This allows to capture a wide range of attitudes towards opinions of others: from considering them as equal peers and giving their opinions the same weights as one's own, to completely ignoring them. One can interpret the bias parameters in at least two ways. First, it can be given a non-ideal interpretation, in terms of cognitive inertia [1], that is a form of resistance to opinion change. This parameter can, however, also be seen as capturing either comparative expertise, similar to [21], or "mutual respect" [36]. The former represents to what extent each agent consider themself more of an "expert" in comparison with the others. Here we have assumed that the bias is the same for all agents. This means that in our model the agents consider themselves at least as expert, on average, as any other. This interpretation, however, assumes the question at hand is such that the participants think that they can compare one another in terms of individual expertise. If instead we interpret the biases as mutual respect, we do not need to presuppose any form of correctness of the attitudes or even expertise on the issue. This interpretation seems better suited for our case where the agents deliberate over preferences, and is in any case compatible with viewing the participants as fully rational deliberators.

Together with the results in [26], who work with deliberation on probabilistic credences, the results presented above thus show that anchoring can emerge from *structural bias* in deliberation, rather than the stemming from shortcomings on the participants' part. That is, it is a bias that can emerge as a result of the structural properties of the deliberation process itself. This highlights the importance of how the deliberation process is set up. The order in which the agents share their views with their group members, the way that the information is propagated through the group, the way that the agents update their opinion in light of what they learn from others, can all result in emergence of biases such as anchoring, even for the "best of us", i.e. fully rational and idealized participants. A better understanding of these effects is crucial in designing and implementing deliberation procedures that avoids these biases, or at least minimizes the chance of their occurrence as far as possible.

## 3.2 Modelling Choices and Representation of Deliberation

Of course, as acknowledged in Section 2, the model we use here is very minimal. It leaves out many aspects usually associated with rational deliberation, most importantly the process of exchanging reasons for and against certain judgments. For many deliberative theorists, starting already in [25], this aspect is central to genuine group deliberation. As such, one could argue that our model is rather one of the process of social influence that might, or might not, accompany the "true" deliberation.

We do not see this as a shortcoming or a limitation, because the model is consistent with richer understandings of rational deliberation. As we have argued above both the bias parameter and the process of preference change in our model can be given a normative or rational interpretation. These, in turn, are in line with the so-call reflexive and social aspects of rational deliberation identified by [13]. So it seems plausible to assume that a comprehensive and substantial model of deliberation such as those envisioned by deliberative theorists would result in a process of opinion change similar to what we study in our model.[2]

One aspect of such more substantial accounts of deliberation which might seem to bear on the investigation of the anchoring bias is that the process of exchanging reasons and arguments can help better identifying the expertise in the groups and thus weighting the opinion of the group members in a more reliable fashion. The process of identifying individual expertise, however, comes with its own associated biases such as social influence, confirmation bias, etc, that can hinder any positive input from the argumentation process. Hence a complete model of rational deliberation that wishes to accommodate the argumentation process needs to also account for these biases. Nevertheless, this can be investigated in our model (in a simplified way) by calibrating the bias values accordingly and possibly applying different bias for different group members. Furthermore, although a better assessment of others' expertise can result in more justified weighting of their opinions, as long as the resulting opinion dynamics can be seen as one coming from

[2]One could, furthermore, envision cases of massive deliberation processes in which the direct and fully public exchange of arguments and reasons is unfeasible or even impossible. In many real instances of deliberation, at least in relatively large groups, the arguments and exchange of reasons is carried out within smaller subgroups that might or might not end up in a consensus. The exchange of arguments between subgroups is however rather minimal and only the aggregated opinion of the subgroups is reported to rest of the group. See e.g. [23]

repeated distance minimization, our model shows that emergence of the anchoring effect in the deliberation is inevitable.

## 3.3 Anchoring and Coherent Aggregation

We have observed, in our model, two phenomena that emerge as a result in deliberation. First is the positive consequence that deliberation fosters coherent aggregation by increasing proximity to single-peakedness, and by the same token almost completely avoids irrational group preferences. Cyclic group preferences are problematic to the extend that they rule out the possibility of social choice. If the group preference is cyclic then there is no most preferred elements in the ordering. Intransitive (but not necessarily cyclic) group preferences, on the other hand, are problematic to because they prevent the construction of a full social welfare ordering. Although they might allow for social choices, the resulting group preferences remain incoherent. These are serious issues for the aggregation of preference rankings, and so the elimination of such profiles speak strongly in favour of pre-voting deliberation.

Our results confirm this positive effect, but also shows that it often come together with a negative, structural bias, namely path dependence through anchoring. How do these positive and negative aspects interact? Anchoring does not prevent coherent aggregation, nor does it reduces the increases in proximity to single-peakedness that can be achieved through deliberation. In a large proportion of cases the two effects go hand in hand. So as far as one focuses on avoiding incoherent group preferences, anchoring does not appear to be a *direct* concern.

Anchoring, however, can *indirectly* affect our overall, positive assessment of deliberation and in the end outweigh the sheer avoidance of incoherent group preferences. Here it is important to distinguish two different contexts of collective decision making: those in which there is what Estlund [21] calls a "procedure-independent" standard for evaluating the outcome of deliberation and voting, and those in which there is no such procedure-independent standard. In simpler terms, cases where there is a correct answer to the question under deliberation, and those in which there is not clearly one.

For cases where there are procedure-independent standards, the expertise of the individual agents becomes relevant. Assuming that this expertise is recognizable, anchoring can diminish the epistemic value of the deliberative outcome if the order in which the participants express their opinion does not correlate with their expertise. Agents with no or little expertise can be assigned an unjustifiably high weight in the deliberation simply as a result of their positioning in the process, and the opinion of those with considerably higher expertise can be overshadowed and receive much less weight in the final verdict. In these scenarios, although deliberation creates coherent preference profiles that can be correctly aggregated, the final profile can be anchored in such a way that causes the final verdict to diverge from the correct answer and forfeit the epistemic advantages that are expected from deliberation. It is important to notice, however, that if the order in which opinions are expressed does correctly correlate with the expertise of the agents, anchoring can actually increase the epistemic value of the deliberation outcome. In such cases, the evaluation of anchoring as a positive or negative phenomenon depends on how one would evaluates the epistemic vs procedural characteristics of the decision making process.

On the other hand, if there is no procedure-independent standard, the advantages of deliberation are focused on its procedural characteristics, such as fairness and equal participation. In these cases path-dependence seems more problematic and the negative consequences arising from it weight stronger against the positive effect of eliminating incoherent profiles. Although deliberation leads to coherent, rational preference profiles, the anchoring in the final profile seem to introduce a form of unfair advantage to those speaking early. This also opens the door to possibility of strategising by allowing one to increase or decrease the weight of some opinion in the final verdict by changing the position of the speaker within the group.

# 4 Conclusion

Recall that, on the one hand, deliberation has been praised by political theorists both from an epistemic and a procedural perspectives. In particular, for decisions involving preference rankings, deliberation seems to offer a natural way to circumvent incoherent group preferences and, more generally, Arrow's impossibility results. On the other hand, social psychologists seem to be more pessimistic and more cautious about the benefits of deliberation. In particular it has been showed that real deliberations can bring about also many instances of biases and undesirable consequences on both the epistemic and procedural views. Many of such biases, however, have been associated with, or explained in terms of cognitive limitations of the agents involved in deliberation. Thus providing a general assessment of deliberation requires an more in depth analysis that clarifies what the positive and negative effects of deliberation hinge on and how different biases can emerge in the course of deliberation.

Here we focused on one instance of the well-known anchoring effect that will result in path-dependence of the outcome of the deliberation. In particular anchoring emerges when the opinion that are expressed earlier receive a higher weight in the final decision. We challenged the standard view that associates this bias with individual failures, and showed that anchoring can emerge merely as a result of the updating mechanism and thus the specific implementation of the deliberative process, even for fully rational agents with no cognitive shortcoming.

Next we investigated how the emergence of these negative effects correlates with positive consequences of deliberation. In particular we looked at one of the most important benefits of deliberation for preference rankings, namely the elimination of cyclic and intransitive profiles. Using our computational model we showed that a considerable percentage of the cases where deliberation seems beneficial by eliminating cyclic or intransitive profiles the resulting preferences are effected by anchoring. This does not, as such, threaten the coherence of the post-deliberation group preference. Anchoring does, however, bring in a number of negative consequences which, we have argued, can outweigh the sheer effect of avoiding voting cycles and intransitive group preferences.

Randomising the order in which the deliberating agents announce their opinion seems an immediate remedy for the problem of anchoring. Indeed, randomising the order prevents agents from strategising by trying to share their preference earlier. This however does not change the fact that the first speaker receives a higher weight in deciding the final profile. In other words randomising the order in which opinions are expressed avoids the problem of strategising but deliberation still falls prey to unfairness by giving some group members, now chosen at random, an unjustifiably high weight in the final outcome.

Taken all together, these findings have important consequences for our understanding of individual and group responsibility in collective decision-making processes. On the one hand, the possibility of coherent aggregation through the creation of single-peaked preferences facilitates the attribution of responsibility to the deliberating collective, and its redistribution to the individual members [33]. On the other hand, anchoring shows that even if, as in our model, the participants are guarantied equal *participation* in the decision making process, i.e. they all speak equally often, this does not translate in equal *influence* on the outcome. This is so even for fully idealized and rational agents that are not bound by any power or hierarchical relations. This inequality in influence only results from a structural bias that puts more weight on those that speak earlier, even if the order of speakers is not under the control of any group member or of an external planner like, as the case in our model, where it is randomly assigned at the beginning of deliberation. Such cases are natural, and as such show, in the same line as [16], that so-called responsibility voids can be more wide-spread than previously assumed [4].

# References


[1] A. Allport and G. Wylie. Task switching, stimulus–response bindings, and negative priming. *Control of cognitive processes: Attention and performance XVIII*, pages 35–70, 2000.

[2] J. Bohman. *Deliberative democracy: Essays on reason and politics.* MIT press, 1997.

[3] J. Bohman. *Judgement in Managerial Decision Making.* UK: Wiley, 2002.

[4] M. Braham and M. VanHees. Responsibility voids. *The Philosophical Quarterly*, 61(242):6–15, 2011.

[5] A. Bramson, P. Grim, D. J. Singer, W. J. Berger, G. Sack, S. Fisher, C. Flocken, and B. Holman. Understanding polarization: Meanings, measures, and model evaluation. *Philosophy of science*, 84(1):115–159, 2017.

[6] G. Chapman and E. Johnson. Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes*, 79:1–39, 1999.

[7] H. Chung and J. Duggan. A formal theory of democratic deliberation. *American Political Science Review*, 114(1):14–35, 2020.

[8] J. Cohen. Deliberation and democratic legitimacy. *A. Hamlin & P. Pettit (eds.): The Good Polity: Normative Analysis of the State.*, pages 17–34, 1989.

[9] J. Cohen. The economic basis of deliberative democracy. *Social Philosophy and Policy*, 6(2):25–50, 1989.

[10] W. D. Cook. Distance-based and ad hoc consensus models in ordinal preference ranking. *European Journal of Operational Research*, 172(2):369–385, 2006.

[11] W. D. Cook and L. M. Seiford. Priority ranking and consensus formation. *Management Science*, 24(16):1721–1732, 1978.

[12] N. C.S. *The Constitution of Deliberative Democracy.* ew Haven: Yale University Press, 1996.

[13] J. S. Dryzek and C. List. Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science*, 33(1):1–28, 2003.

[14] J. S. . C. L. Dryzek. Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science*, 33(1):1–28, 2002.

[15] C. Duddy and A. Piggins. A measure of distance between judgment sets. *Social Choice and Welfare*, 39(4):855–867, 2012.

[16] H. Duijf. Responsibility voids and cooperation. *Philosophy of the social sciences*, 48(4):434–460, 2018.

[17] N. Epley and T. Gilovich. Putting adjustment back into the anchoring and adjustment heuristic: differential processing of self-generated and experimenter-provided anchors. *Psychological Science*, 12:391–396, 2001.

[18] D. Estlund. Who's afraid of deliberative democracy? on the strategic/deliberative dichotomy in recent constitutional jurisprudence. *Texas Law Review*, 71:1437–1477, 1993.

[19] D. Estlund. pinion leaders, independence, and condorcet's jury theorem. *Theory and Decision*, 36(2):131–162, 1994.

[20] D. Estlund. Beyond fairness and deliberation: The epistemic dimension of democratic authority. *W. Bohman and J. Rehg (edt): Deliberative Democracy: Essays on Reason and Politics*, 36(2):173–204, 1997.

[21] D. Estlund. *Beyond fairness and deliberation: The epistemic dimension of democratic authority*, volume 32. MIT Press Cambridge, MA, 1997.

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

[23] N. Fay, S. Garrod, and J. Carletta. Group discussion as interactive dialogue or as serial monologue: The influence of group size. *Psychological science*, 11(6):481–486, 2000.

[24] A. Furnhama and H. C. Boob. A literature review of the anchoring effect. *JThe Journal of Socio-Economics*, 40(1):35–42, 1997.

[25] J. Habermas. *The theory of communicative action, Volume 1.* 1984.

[26] S. Hartmann and S. Rafiee Rad. Anchoring in deliberation. *Erkenntnis*, 2018.

[27] S. Hartmann and S. Rafiee Rad. Voting, deliberation and truth. *Synthese*, 195(3):1273–1293, 2018.

[28] J. Kemeny. Mathematics without numbers. *Daedalus*, 88(4):577–591, 1959.

[29] J. Kemeny and L. Snell. Preference ranking: An axiomatic approach. In *Mathematical Models in the Social Sciences*. Hafner, 1962.

[30] D. Landa and A. Meirowitz. Game theory, information, and deliberative democracy. *American Journal of Political Science*, 53(2):427–444, 2009.

[31] C. List. Democratic deliberation and social choice: A review. 2017.

[32] C. List, R. C. Luskin, J. S. Fishkin, and I. McLean. Deliberation, single-peakedness, and the possibility of meaningful democracy: Evidence from deliberative polls. *The Journal of Politics*, 75(1):80–95, 2012.

[33] C. List, P. Pettit, et al. *Group agency: The possibility, design, and status of corporate agents*. Oxford University Press, 2011.

[34] B. Manin. On legitimacy and political deliberation. *Political Theory*, 15(3):338–368, 1987.

[35] J. Marti. The epistemic conception of deliberative democracy defended. *S. Besson & J. Marti (eds.): Deliberative Democracy and Its Discontents*, pages 27–56, 2006.

[36] C. Martini, J. Sprenger, and M. Colyvan. Resolving disagreement through mutual respect. *Erkenntnis*, 78(4):881–898, 2013.

[37] T. McElroy and K. Dowd. Susceptibility to anchoring effects: how opennessto-experience influences responses to anchoring cues. *Judgment and Decision Making*, 2:48–53, 2007.

[38] E. B. Mussweiler, T. and F. Strack. Anchoring effect. *R. Pohl (ed.): Cognitive Illusions: A Handbook of Fallacies and Biases in Thinking, Judgement, and Memory*, pages 183–196, 2004.

[39] T. Mussweiler and F. Strack. Thinking the unthinkable: the effects of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology*, 19:67–91, 1989.

[40] T. Mussweiler and F. Strack. Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3):437–446, 1997.

[41] T. Mussweiler and F. Strack. Hypothesis-consistent testing and semantic priming in the anchoring paradigm: a selective accessibility model. *Journal of Experimental Social Psychology*, 35:136–164, 1999.

[42] T. Mussweiler and F. Strack. The semantics of anchoring. *Organizational Behavior and Human Decision Processes*, 86:234–255, 2001.

[43] T. Mussweiler and F. Strack. Subliminal anchoring: judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes*, 98:133–143, 2005.

[44] R. G. Niemi. Majority decision-making with partial unidimensionality. *American Political Science Review*, 63(2):488–497, 1969.

[45] P. Pettit. Deliberative democracy and the discursive dilemma. *Philosophical Issues*, 11:268 – 299, 02 2010.

[46] D. A. Prentice and D. T. Miller. Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64(2):243, 1993.

[47] G. Stasser and W. Titus. Hidden profiles: A brief history. *Psychological Inquiry*, 14(3-4):304–313, 2003.

[48] . reference removed for blind review. .

[49] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1974.

# Appendix: technical definitions

**Note for reviewers.** This section is taken verbatim from [48]. Depending on the Journal's editorial policy, it could be instead refereed to as an online Appendix.

**Weak and strict preference rankings:** Let $R_i$ be a preference pre-order or ranking for an agent $i$ over a set of alternatives $\mathcal{A}$. A preference profile $R$ is a set of preference rankings, one for each agent $i$. When $aR_ia'$, we say that $i$ weakly prefers $a$ over $a'$. When it is furthermore not the case that $a'R_ia$ then we say that $i$ strictly prefers $a$ over $a'$, and write $a >_i a'$. Indifference is defined as usual. We call a ranking $R_i$ strict whenever for all $a, a'$, either $a >_i a'$ or $a >_i a'$. Otherwise we call it weak.

**Voting cycles:** Given a profile $R$ of rankings, an alternative $a$ is a Condorcet winner whenever it wins a majority on any pairwise comparison: for each alternative $a' \neq a$, we have that a majority of agents strictly prefers $a$ over $a'$. Pairwise majority yields cyclic social preferences, i.e., a voting cycle, whenever there are three alternatives $a_1$, $a_2$, and $a_3$ such that there is a majority of agents that strictly prefers $a_1$ over $a_2$, a (possibly different) majority that strictly prefers $a_2$ over $a_3$, and one that strictly prefers $a_3$ over $a_1$.

**Transitive Social Preference:** A profile $R$ of rankings is said to produce a transitive social preference, if the pairwise-majority-comparison relation $M_R$ associated with it is transitive, i.e., for all alternatives $a, b$, and $c$, if $a\, M_R\, b$ and $b\, M_R\, c$ then $a\, M_R\, c$.

**Single-peaked and single-plateau profiles:** Let $\succ$ be a strict ordering of the alternatives in $\mathcal{A}$. We say that a profile $R$ is strictly single-peaked relative to $\succ$ whenever for each agent $i$ there is an alternative $a$ such that for all $a'$, $a' \succ a$ implies $a' <_i a$ and $a \succ a'$ implies $a' <_i a$. A profile $R$ is single-plateau relative to $\succ$ whenever for each agent $i$ and triple of alternatives $a, b, c$, such that $a \succ b \succ c$ or $c \succ b \succ a$, it is not the case that $aR_ib$ and $cR_ia$. A profile $R$ is weakly single-peaked relative to $\succ$ whenever for each agent $i$ there is an alternative $a$ such that for all $a'$, $a' \succ a$ implies $a'R_ia$ and $a \succ a'$ implies $aR_ia'$. Strict single-peakedness implies single-plateauness, which in turns implies weak single-peakedness, but none of these implication goes the other way around. When $R$ is single-peaked, weakly or strictly, or single-plateau, relative to $\succ$ we say that the latter is a structuring dimension for $R$.

**Arrowian postulates:** Let $f$ be the aggregation function and $R$ a profile of individual preferences. Then $f$ satisfies *rationality* if for all $R$, $f(R)$ is a complete pre-order. It satisfies *weak pareto* if $xf(R)y$ but not $yf(R)x$ whenever $x <_i y$ for every agent $i$. *Independence* mens that for any profiles $R$ and $R'$, if $xf(R)y$ but not $xf(R')y$ then there must be an agent $i$ such that $xR_iy$ but not $xR'_iy$. *Non-dictatorship* means that there is no $i$ such that for all profiles $R$, $f(R) = R_i$.

## Distance measures

The Kemeny–Snell distance between rankings is defined as follows. First construct an *agenda* containing, for each pair of alternatives $a_i, a_j$ $i \neq j$, the propositions $(a_i, a_j)$ and $\neg(a_i, a_j)$. For a ranking $r$ define the *judgement set* $J_r$ as follows: if $a_i$ is weakly preferred to $a_j$ according to $r$ then put $(a_i, a_j) \in J_r$, otherwise put $\neg(a_i, a_j) \in J_r$. To illustrate this, consider the ranking $r_1 = (a_1, a_2, a_3)$ over three alternatives, meaning that $a_1$ is strictly preferred to $a_2$, which is in turn strictly preferred to $a_3$. The corresponding judgement set $J_{r_1}$ is

$$J_{r_1} = \{(a_1, a_2), \neg(a_2, a_1), (a_1, a_3), \neg(a_3, a_1), (a_2, a_3), \neg(a_3, a_2)\}$$

If, instead, we take $r_2 = (a_3, a_2, a_1)$, then $J_{r_2}$ is:

$$J_{r_2} = \{\neg(a_1, a_2), (a_2, a_1), \neg(a_1, a_3), (a_3, a_1), \neg(a_2, a_3), (a_3, a_2)\}.$$

The Kemeny–Snell distance between any two rankings $r_1$ and $r_2$ is defined as the Hamming distance between $J_{r_1}$ and $J_{r_2}$, i.e., the number of binary changes that one has to make to transform $r_1$ into $r_2$. In our example, the distance $d(r_1, r_2)$ between $r_1$ and $r_2$ is 6.

Kemeny and Snell [29] characterize their measure uniquely by a set of intuitive axioms. The first axiom ensures that the measure is mathematically a distance measure. That is, for all rankings $r_1, r_2$, and $r_3$,

A1.1. $d(r_1, r_2) \geq 0$,

A1.2. $d(r_1, r_2) = d(r_2, r_1)$,

A1.3. $d(r_1, r_3) \geq d(r_1, r_2) + d(r_2, r_3)$, with equality holding if and only if $r_2$ is between $r_1$ and $r_3$.

The next axiom requires the distance to be invariant under the relabelling of the alternatives. This ensures that the distance does not depend on the specific alternatives that we are ranking but only on the way they are ranked.

A2. If $r_1'$ and $r_2'$ result from applying the same permutation of objects to $r_1$ and $r_2$, then

$$d(r_1', r_2') = d(r_1, r_2).$$

Next it is required that if two rankings agree in the beginning and/or at the end, then the distance should only depend on the middle segment where they differ.

A3. If two rankings $r_1$ and $r_2$ agree except for a set $S$ of $k$ elements, which is a segment in both $r_1$ and $r_2$, then $d(r_1, r_2)$ may be computed as if these $k$ objects were the only objects being ranked.

Their last axiom sets a unit of measurement for their distance.

A4. The minimum positive distance is 1.

Crucial for this axiomatization is the notion of *betweenness*. The betweenness relation for rankings in the Kemeny–Snell axiomatization is defined in terms of betweenness of the corresponding judgement sets. We say that the judgement set $J$ is *between* judgement sets $J_1$ and $J_2$ if $J_1, J_2$, and $J$ are distinct and, on each proposition, $J$ agrees with $J_1$ or with $J_2$ (or both).

The *Duddy-Piggins* distance measure uses the same representations for agendas and rankings. Now construct the graphs of all possible judgement sets, with an edge between $J_1$ and $J_2$ if and only if there is no judgement sets between them. Betweenness is defined in the same way as for the Kemeny–Snell distance. The Duddy–Piggins distance between $r_1$ and $r_2$ is the length of the shortest path between $J_{r_1}$ and $J_{r_2}$ on this graph. To look at an example again, start, as before, with $r_1 = (a_1, a_2, a_3)$ and $r_2 = (a_3, a_2, a_1)$ with the corresponding $J_{r_1}$ and $J_{r_2}$. The shortest path between $r_1$ and $r_2$ is through the judgement sets

$$J_{r_1} - - - - J_1 - - - - J_2 - - - - J_3 - - - - J_{r_2},$$

defined as follows:

$$J_1 = \{(a_1, a_2), \neg(a_2, a_1), (a_1, a_3), \neg(a_3, a_1), (a_2, a_3), (a_3, a_2)\},$$

$$J_2 = \{(a_1, a_2), (a_2, a_1), (a_1, a_3), (a_3, a_1), (a_2, a_3), (a_3, a_2)\},$$

$$J_3 = \{(a_1, a_2), (a_2, a_1), \neg(a_1, a_3), (a_3, a_1), \neg(a_2, a_3), (a_3, a_2)\},$$

which gives a Duddy–Piggins distance of 4 between $r_1$ and $r_2$.[3]

Thus the Duddy–Piggins distance between two rankings $r_1$ and $r_2$ also reflects the number of steps required to move from ranking $r_1$ to $r_2$. The difference between the Kemeny–Snell and Duddy–Piggins measures can be seen in what changes are permitted at each step. For the Kemeny–Snell measure, only a single binary change is allowed at each step. Notice that this can result in an inconsistent judgement set, which then requires another binary change to become consistent. For example, consider the ranking $r = (a, \{b, c\})$, where alternatives $b$ and $c$ are ranked equal and strictly below the alternative $a$. The corresponding judgement set will then

[3]This example shows indirectly the kind of "double-counting", mentioned in [15], that results from taking the Hamming distance between $J_{r_1}$ and $J_{r_2}$. From the fact that $(a_1, a_2)$ and $(a_2, a_3)$, and their negations, are in $J_{r_1}$ and $J_{r_1}$, respectively, we know by transitivity and completeness that $(a_1, a_3)$ and its negation must be in $J_{r_1}$ and $J_{r_2}$, respectively. The Duddy–Piggins measure ignores this third step, to arrive at a distance of 4, while Kemeny–Snell includes it, giving 6.

be

$$J_r = \{(a,b), \neg(b,a), (a,c), \neg(c,a), (b,c), (c,b)\}.$$

A binary change represented by replacing $\neg(b,a)$ with $(b,a)$ will result in an inconsistent judgement set in which $a$ is ranked equal to $b$, and $b$ is ranked equal to $c$ while $a$ is ranked strictly higher than $c$. A second binary change represented by replacing $\neg(c,a)$ with $(c,a)$ in the next step, will make the judgement set consistent again. The Duddy-Piggings measure will, under certain conditions, allow for multiple binary changes in a single step to avoid these inconsistent intermediate judgement sets. To be more specific, in Duddy–Piggins we are allowed to change the relative ranking of an alternative with respect to a *set* of alternatives (as opposed to only one in Kemeny–Snell), but only if the alternatives in the set are all ranked equally.

The *Cook–Seiford* distance is calculated by first assigning numbers, call them the *CS-numbers*, to the alternatives in a ranking, starting with 1 for the top alternative, 2 for the next-best alternative, etc. In case of a tie, we assign the average number to the tied alternatives. For instance, if there is one alternative at the top and two alternatives are tied just below, each of the latter gets the average of 2 and 3, i.e., 2.5. Let $x_i^r$ be the CS-number of alternative $a_i$ in ranking $r$. The CS-distance between rankings $r_1$ and $r_2$ is the sum of the absolute differences between the CS-numbers of the alternatives:

$$d(r_1, r_2) = |x_1^{r_1} - x_1^{r_2}| + \cdots + |x_n^{r_1} - x_n^{r_2}|.$$

Starting with our two rankings $r_1 = (a_1, a_2, a_3)$ and $r_2 = (a_3, a_2, a_1)$ again, we get

$x_1^{r_1} = 1$, $x_2^{r_1} = 2$, $x_3^{r_1} = 3$ and $x_1^{r_2} = 3$, $x_2^{r_2} = 2$, $x_3^{r_2} = 1$:

$$d(r_1, r_2) = |1 - 3| + |2 - 2| + |3 - 1| = 4.$$

The Cook–Seiford measure is characterized uniquely by the same set of axioms as the Kemeny–Snell measure. Their difference hinges on the different underlying notions of "betweenness" that they use.

The notion of betweenness for Cook–Seiford is not defined with respect to the corresponding judgement sets. Instead it is given in terms of the CS-numbers assigned to the alternatives. In this case we say that the ranking $r$ is between rankings $r_1$ and $r_2$ if for each alternative, its CS-number in $r$ is between its CS-numbers in $r_1$ and $r_2$, respectively. That is, for each alternative $a_i$,

$$x_i^{r_1} \leq x_i^{r} \leq x_i^{r_2} \qquad \text{or} \qquad x_i^{r_2} \leq x_i^{r} \leq x_i^{r_1}.$$

To see how this makes a difference, consider, for example, the profile of three rankings corresponding to the Condorcet Paradox:

$$r_1 = (a, b, c), \qquad r_2 = (b, c, a), \qquad r_3 = (c, a, b).$$

Here the intuitive result of aggregating this profile is for the group to settle on ranking the three alternatives equally. This option is however excluded by the Kemeny-Snell measure. It is easy to check that the distance between the equal ranking and the given profile is higher than the distance between each of the rankings $r_1, r_2$, or $r_3$ to the profile. This is however not the case for the Cook–Seiford measure. With this measure the equal ranking would indeed be the one to minimise the distance. The reason is that for the Kemeny–Snell measure the equal ranking does not fall between any two of the rankings in the profile, while with the Cook–Seiford definition of betweenness it does.