

## RESEARCH ARTICLE

# Diverse Discussion in Public Deliberation on Cancer Drug Funding

Daniel Steel\*, Naseeb Bolduc\*, Reena Titoria\*, Michael Burgess\*, Julia Abelson† and Stuart Peacock‡

Structured deliberations among members of the public are increasingly viewed as useful inputs to health policy decisions that also rely on scientific evidence and expertise. Such deliberations typically aim for discussions that explore a diversity of ideas and perspectives. However, the concept of a diverse discussion has not been thoroughly examined and methods for measuring the extent to which a discussion actually was diverse are lacking. In this article, we develop a theoretical account of diverse discussion and propose a method for operationalizing it, which we illustrate by means of an analysis of transcripts from public deliberations on cancer drug funding in Canada.

**Keywords:** Diversity; information elaboration; public deliberation; deliberative mini-publics

## 1. Introduction

Policy makers tasked with difficult decisions have increasingly turned to deliberative mini-publics, which recruit members of the public to be informed and deliberate on a topic, as a source of advice or legitimacy (Brown 2006; Fishkin 2018; Fung 2007; O'Doherty & Burgess 2013; Rowe & Frewer 2000; Smith 2009). This is especially true for issues related to health policy and biomedical research, such as funding for cancer drugs and human tissue biobanking (Abelson et al. 2003; Degeling et al. 2017). Diversity is frequently noted as a desired quality of public deliberations (Burgess et al. 2016; Carman et al. 2015; Degeling et al. 2015; Longstaff & Burgess 2010). Usually, diversity is understood in reference to who is there: a deliberation is thought to be diverse if its participants are sufficiently varied in their social, professional, or attitudinal profiles. But diversity can also be viewed as a quality of the discussion itself. One group of participants might only discuss a few isolated issues, while another raises a wide range of issues and explores the interconnections among them. The latter discussion is naturally judged more diverse than the former. However, while the concept of diversity as it pertains to groups of people has been examined (Harrison & Klein 2007; McDonald & Dimmick 2003; Solanas et al. 2012; Steel et al. 2018), the concept of diverse discussion has received little attention.

In this article, we develop a theoretical account of diverse discussion and propose a method for operationalizing it, which we illustrate by means of an analysis of transcripts from public deliberations on cancer drug funding in Canada. A diverse discussion, as we propose to understand it, explores a variety of relevant information, perspectives, or heuristics and integrates them in addressing the task at hand. We explain how diverse discussion plays an important, although often implicit, role in theoretical discussions of ideal deliberation (Anderson 2006; Habermas 1990; Longino 1990, 2002). These proposals typically specify conditions that are, to all appearances, designed to foster diverse discussion, which in turn is thought to enhance, or be constitutive of, quality deliberation. Indeed, diverse discussions are often argued to produce superior results in science and elsewhere (Anderson 2006; Bohman 2006; Longino 1990, 2002; Muldoon 2016; Page 2017).

Our approach to operationalizing the notion of diverse discussion relies on a theoretical construct known as information elaboration, for which instruments have been developed for experimental contexts. Information elaboration refers to the process by which varied perspectives, knowledge, experiences, or other cognitive resources dispersed in a group are elicited and integrated in the process of discussion (Homan et al. 2007; Kooij-de Bode et al. 2008; Meyer et al. 2011). There is a straightforward correspondence between information elaboration and the intuitive notion of diverse discussion. Like diverse discussion, information elaboration is enhanced when a larger number of perspectives, items of information, and so on are expressed, explored, and brought in contact with one another. Our approach, therefore, is to adapt

\* University of British Columbia, CA

† Faculty of Health Sciences, McMaster University, CA

‡ Cancer Control Research, Provincial Health Services Authority, CA

Corresponding author: Daniel Steel ([daniel.steel@ubc.ca](mailto:daniel.steel@ubc.ca))

instruments developed for measuring information elaboration within an experimental context, wherein information elaboration is studied as an important causal mediator between demographic diversity and group performance (Homan et al. 2007; Kooij-de Bode et al. 2008; Meyer et al. 2011), to the analysis of transcripts from public deliberations.

We develop our proposal in connection with an analysis of transcripts from a series of deliberative mini-publics, or ‘citizen panels,’ conducted in 2016 under the title *Making Fair and Sustainable Decisions about Funding for Cancer Drugs in Canada* and sponsored by the Canadian Partnership Against Cancer (Bentley et al. 2017; Bentley et al. 2018; cf. Bentley et al. 2019; Costa et al. 2019). In these events, participants recruited from the general public were informed about challenging trade-offs confronted by policy makers in the arena of funding for cancer drugs and were invited to discuss and to ultimately make recommendations about the considerations that should guide decisions on this matter.

The organization of this paper is as follows. Section 2 provides a theoretical grounding for the concept of diverse discussion, links it to literature on public deliberation, and explains its connection to information elaboration. Section 3 describes the transcripts we analyze and the deliberations they came from, while section 4 describes how we adapted the information elaboration instrument to the transcripts on cancer drug funding in Canada alluded to above. Section 5 explains how data generated by the method we describe can be used to operationalize diverse discussion and gives examples of research questions that could be studied using our method. Section 6 summarizes the positive features of our proposal and considers its limitations.

## 2. Conceptualizing Diverse Discussion

Diversity is usually seen as a function of *who is present*, in which case individuals are the units of analysis and are diverse if there is sufficient variety among them along demographic or cognitive lines. However, diversity can also be a function of *what is said*, in which case the units of analysis are transcripts and the speech acts they contain. Just as a collection of people might or might not be diverse, a discussion might explore a more or less diverse array of ideas.

One hallmark of a diverse discussion is that a variety of relevant ideas are raised. However, diverse discussion as we understand it also involves substantive interaction among those ideas. Thus, a discussion in which each participant expresses a distinct perspective but in which no one engages with anyone else’s views would not be very diverse. Our concept of diverse discussion, then, embodies what Medina (2013: 7) calls ‘epistemic interaction,’ wherein ‘resources are pooled and experiences and imaginations are shared, compared, and contrasted.’ Medina proposes that epistemic interaction is closely linked to integration, in which people of varied social backgrounds are not merely present in society but regularly interact with one another on an equal basis in educational, professional, and other contexts. Anderson

(2013) makes the case that integration is essential for well-functioning democracies, and Medina (2013) proposes that openness to equal epistemic interactions with people who hold differing views is a crucial democratic sensibility.

We suggest that integration and epistemic interaction are also relevant to the concept of diversity. A stylized example and analogy may be helpful to motivate this idea. Imagine two universities. In University A, two ethnic groups are represented in the student body, but there are no interactions among them. At University B, the same two ethnic groups are represented in the same proportions, but there are frequent interactions among them as peers in a variety of contexts (e.g., classes, social events, etc.). In other words, University B is more integrated than University A, so that students at University B are more likely than those at University A to encounter students whose ethnicities differ from their own. That provides a straightforward reason to judge University B more ethnically diverse than University A. Bringing differing people in contact with one another is an important aspect of diversity, and this happens at University B but not A, at least with regard to ethnic differences.

Epistemic interaction is the analogue to integration when the diversity of a discussion, rather than the diversity of a collection of people, is in question. Note that integration alone does not ensure epistemic interaction. People in frequent contact on equal terms might, for instance, politely avoid objecting to views they disagree with out of a desire to maintain positive relationships. In parallel to the previous example, then, consider Discussion A, wherein two ideas are proposed but are never brought in contact with one another, and Discussion B, in which the same two ideas are broached and their relative merits thoroughly debated. Just as the student body of University B is more diverse than that of University A, Discussion B is more diverse than Discussion A. Of course, other discussions could be more diverse than Discussion B, for instance, by raising and integrating further ideas. The takeaway point here, however, is that both the number of ideas raised and the interaction among them matter for diverse discussion.

While the term diverse discussion is not frequently encountered in literature on public deliberation, the underlying idea is nevertheless quite common. Consider the following minimal criteria for public deliberations (Blacksher et al. 2012):

- (1) the provision of balanced, factual information that improves participants’ knowledge of the issue;
- (2) the inclusion of diverse perspectives to counter the well-documented tendency of better educated and wealthier citizens to participate disproportionately in deliberative opportunities and to identify points of view and conflicting interests that might otherwise go untapped; and
- (3) the opportunity to reflect on and discuss freely a wide spectrum of viewpoints and to challenge and test competing moral claims.

All three criteria mentioned in this passage are linked to diversity of discussion. The first criterion demands that participants be presented with balanced information that does focus attention on a single point of view, while the second insists that a variety of perspectives be present among the participants themselves. And the third criterion describes epistemic interaction, as when the merits of different viewpoints are compared.

Theoretical analyses of ideal deliberation also describe conditions apparently intended to promote diverse discussion. These typically emphasize an equitable setting wherein obstacles to the free expression and discussion of ideas, such as sexism or racial discrimination, are expunged (Anderson 2006; Habermas 1984; Longino 1990, 2002). For example, Anderson (2006) proposes universal inclusion and the expression of dissent as conditions needed for democracy to generate discussions that generate epistemic benefits from diversity. One of the aims of such models is apparently to foster diverse discussion in the way we have described it: they describe conditions in which a variety of potentially conflicting ideas can be raised and rationally debated. This requires not only that a number of different beliefs, values, and perspectives be aired, but also that they be brought into contact with one another so that epistemic interaction occurs. Of course, models of ideal deliberation typically have concerns besides diverse discussion. They often embody moral principles, such as equality and respect for human rights; provide a philosophical account of rationality or objectivity; and seek to promote more effective democratic decision-making. Nevertheless, we suggest that the concept of diverse discussion plays an important yet implicit mediating role in theoretical accounts of ideal deliberation, wherein equity and a diverse array of participants enhances the diversity of discussion, which in turn improves the quality of deliberation.

Diverse discussion, understood as the expression of a broad variety of contrasting or differing views that are jointly explored in relation to the task, is similar to information elaboration, which refers to ‘the degree to which information is shared, processed, and integrated in group interaction’ (Homan et al. 2007: 1193). Note that the term ‘integration’ is being used in this quoted passage to mean something similar to Medina’s epistemic interaction (i.e., the engagement of differing ideas or information) rather than integration in the sense of racial integration of public schools. We will use the term integration in the epistemic sense in what follows. The more items of relevant information that are shared and the more they are processed and integrated in relation to carrying out the task, the greater the extent of information elaboration. Information elaboration is motivated by the observation that diverse perspectives and information might be possessed by participants of a discussion but not be expressed, explored, or integrated with one another (van Knippenberg et al. 2004). Moreover, information elaboration is often studied as a mediator between group diversity and improved performance (Homan et al. 2007), paralleling the implicit role of diverse discussion in theoretical accounts of ideal deliberation.

These similarities suggest that information elaboration could serve as a basis for operationalizing the theoretical construct of diverse discussion, an idea we pursue in below in sections 4 and 5.

Finally, diversity of discussion is a distinct theoretical construct from quality of deliberation. For example, a discussion might raise and integrate a wide variety of issues, but also it might involve numerous false premises and fallacious inferences and thus be diverse but low quality. One might reasonably suppose that diverse discussions are more likely to be of higher quality, but we regard this as a hypothesis to be studied empirically rather than as an a priori principle. To illustrate these points, consider the Discourse Quality Index (DQI), which is intended to measure deliberation quality from a Habermasian perspective (Steenbergen et al. 2003; cf. Edwards et al. 2008; De Vries et al. 2011; Himmelroos 2017; Rowe et al. 2004; Steiner et al. 2004). DQI includes five coding categories: (1) *Participation*, which measures whether participants interrupt one another, (2) *Level of Justification*, which assesses the extent to which participants give complete reasons for positions they advocate, (3) *Content of Justification*, which assesses whether participant’s reasons are based on individual interest or the common good, (4) *Respect*, which includes three sub-indicators: respect for other groups, respect for demands of others, and respect for counterarguments, and finally (5) *Constructive Politics*, which assesses whether participants suggest mediating or compromise positions (Steenbergen et al. 2003). Because DQI does not include a coding category relating to the variety of ideas or issues discussed, it is not a measure of diverse discussion. Some of the coding categories, such as *Participation* and *Respect*, may measure factors that foster diverse discussion, and others like *Level* and *Content of Justification*, may be effects of it. However, the *Constructive Politics* category of DQI does involve a particular type of integration (namely, compromise) and thus is more directly related to diverse discussion. But even here, DQI would not be a good measure for diverse discussion, because this requires tracking themes raised in the discussion and their integrations with other themes. Consequently, we regard DQI and our approach to measuring diversity of discussion as largely complementary.

### 3. The CPAC Transcripts

The transcripts we analyze derive from six two-day deliberative events sponsored by the Canadian Partnership Against Cancer (CPAC) in 2016 (Bentley et al. 2017; Bentley et al. 2019). Our analysis of these transcripts for this research project was approved by the Behavioural Research Ethics Board at the University of British Columbia (H19-02928). The deliberations examined funding decisions regarding cancer drugs whose effects had already been studied and which had received regulatory approval. The deliberation focused on how to decide which of these drugs to place on formularies—so that their cost would be covered for patients through publicly funded healthcare systems—given that it is not financially possible to fund every drug. The deliberations aimed to provide policy-

makers with public input about the value judgments that should guide trade-offs inherent in these decisions.

The six CPAC deliberations occurred in Saskatchewan, Ontario, Quebec (separate events in English and in French), and Nova Scotia, along with a final pan-Canadian event. Each event consisted of approximately 20–25 citizens who were recruited to reflect a diversity of experiences and perspectives based on the demographics of each respective province. Participants were recruited through both random and purposeful selection. Email invitations were randomly distributed, and interested recipients were asked to complete a survey to collect demographic and experiential information. Participants were then selected to create balanced groups for deliberation.

The structure of these events was based on deliberative methods from Burgess & O'Doherty (2009) and the McMaster Health Forum (<https://www.mcmasterforum.org/spark-action/citizen-panels>). Each deliberative event occurred over one weekend (two full days) and consisted of both small and large group sessions. Prior to deliberation, participants were provided with background information through a plain-language booklet introducing the topic, an informational video, and presentations by a local patient representative and an oncologist. In each event, the 20–25 participants were divided into three smaller breakout groups consisting of 6–8 individuals. Each of these small groups had concurrent discussions, where participants were provided with a question and were prompted to deliberate by a facilitator. After approximately one hour of small group exchange, participants reconvened in a large group, and one member from each small group was tasked with summarizing their discussion through a report back. After the report backs, all 20–25 members of the large group engaged in facilitated conversation, with the aim of developing a set of recommendations. All the small and large group discussions at all six events were audio-recorded and transcribed verbatim. Results of the deliberation were analyzed in a published article (Bentley et al. 2019) and summarized in an online report (Bentley et al. 2017).

For the purposes of this study, we only examined the transcripts for the events in Halifax, Nova Scotia, Montreal, Quebec (English), and Hamilton, Ontario. And for each of these, we looked only at the first set of three concurrent small group sessions and the subsequent large group session. We limited attention in this way because our purpose was to assess the validity of our proposed method for operationalizing diverse discussion rather than to provide an overall evaluation of any of the deliberations. The sub-discussions we examined all focused on the same question: 'What should guide decisions about whether to fund new cancer drugs, or change the funding provided for existing cancer drugs?' (Bentley et al. 2018). As a result, we could use the same codebook for all of the transcripts we examined.

#### 4. The Information Elaboration Instrument

This section describes instruments developed for measuring information elaboration in the context of

experiments and how we repurposed them for a new role of analyzing transcripts from public deliberations.

Instruments for measuring information elaboration have been created for experimental contexts in which research participants are presented with a task and items of information needed to complete the task are distributed among them (Homan et al. 2007; Meyer et al. 2011). The instrument then records whether the information items are raised in the discussion and the extent to which they are elaborated. The earliest variant of this instrument we know of uses a coding scale of 0 to 5, where 0 indicates that an item of information is not raised at all, 1 indicates that the item was raised, 2 indicates that the item was raised and acknowledged by another person, 3 indicates that another person asked a question about the item, 4 indicates that another person drew an inference from the item, and 5 indicates that another person integrated that item of information with another item (Homan et al. 2007: 1193). Other authors have reported using a 0 to 3 scale, which resembles the coding scheme just described except that coding categories 3 through 5 are collapsed (Meyer et al. 2011: 267).

In our transcripts, the distinction between mentioning and acknowledging an item mentioned was not very useful as nearly every contribution to the discussion was at least acknowledged by the facilitator if not by other discussants. In addition, we found the difference between asking a question about an item and drawing an inference from it often difficult to distinguish clearly, because people may express ideas using rhetorical questions. The distinction between asking questions and making inferences was, moreover, not especially important for our purposes. Finally, we use the term theme rather than item because participants were typically suggesting considerations that should guide decisions about cancer drug funding rather than putting forward items of information.

Consequently, we decided upon the following 3-point IE coding scheme: 0 indicates that theme is not mentioned in the discussion, 1 that the theme is mentioned and possibly acknowledged by another person, 2 that the theme is mentioned and a question is asked about it or an inference is drawn from it by another person, and 3 that the theme is mentioned and integrated with at least one other theme by another person. The following are examples to illustrate codes 1 through 3 (note that both Cost and Number Treated are themes in our codebook, see **Table 1**).

- Example of Code 1: A: So, yeah, I think cost is important. B: Uh huh, can't overlook that.
- Example of Code 2: A: So, yeah, I think cost is important. B: Do you mean, for the government or for the patients? For both maybe?
- Example of Code 3: A: So, yeah, I think cost is important. B: I think that connects to the number of people it can help. Maybe if it works for a lot of people, you get more bang for your buck.

We found integration to be the most challenging aspect of coding and consequently discuss this in detail. Integration



**Table 1:** Codebook for Day 1 of the CPAC Transcripts.

Theme	Explanation, Subthemes, Examples	Non-Examples
Quality of Life	Well-being associated with a health or disease state 'Feeling like crap' Pain Independence 'able to function' Mobility Relations with family, friends	Throwing up, hair falling out, etc.—these should be coded as side-effects; quality of life should be about the well-being associated with a health/disease state, not just adverse physiological effects.
Oral Pills over IV	Ease of access of oral pills Taking medicine at home vs travelling to hospital Oral easier for rural people	Treatment should be accessible to everybody no matter where they live; this is equity
Cost	Cost to the patient Cost to government of drug Cost effectiveness Ancillary cost to system (e.g., paying nurses to administer drug) Total cost	Financial status of patient—this is a patient trait (e.g., rich people should pay out of pocket)
Effectiveness	Percentage who benefit (among those treated) Extent of benefit (e.g., cure or not?)	Drug helps for a common type of cancer; this is number treated
Target specific need	Treatment targeted to specific cancer Genetically tailored	Genetic predisposition to cancer; this code also requires a treatment targeted to the disorder
Length of life	Survival rate	Should prefer treatments that prevent harm
Prevention	Public health interventions to encourage healthy behaviour, or reduce cancer risk exposure, generally  Note: technically not relevant as answer to the question, but often raised	Merely mentioning individual behaviors that may cause susceptibility to cancer, such as smoking. Prevention code requires PH intervention
Side effects	Discussed as adverse physiological effects of the drug	'Feeling miserable'—this is quality of life because it is about the affective state
Equity	Fairness; everybody should get the same no matter ...	Health care is a human right
Number treated	Prevalence of the cancer(s) treated by the drug Multiple uses—drug treats many types of cancer or possibly other non-cancer diseases	Treatment produces more benefits; this is about effectiveness
Patient traits	Who would be treated by this drug? Old vs young, males vs females, rich vs poor Smokers, obese (people 'responsible' for their cancer) Early vs late stage Severity of the cancer the patient has	Patient values; e.g., decisions should take values of public into account; this is not something to code because it is uninformative; the question is <i>which</i> values?
Trustworthiness	Evidence from independent sources or industry only? Transparency Strength of evidence	Simply mentioning evidence without any concern about whether the evidence is strong enough or should be trusted
Alternative treatments	'do we need this drug if we've already got a treatment for it?' Delisting when (or only when) there is an alternative—sometimes linked to grandfathering May or may not be a drug	Merely mentioning a decision about whether to fund drug A or drug B; this is the basic scenario participants are asked to address, not something to code. This code is for treatment already available vs new drug for the same cancer.

occurs when (a) two or more themes are mentioned or referred to in a speaking turn directly or implicitly (e.g., if the speaking turn occurs in a thread where theme X is the topic of conversation and the speaker is responding to X by drawing a connection with Y) and (b) some relationship between those themes is asserted. Thus, simply mentioning two or more themes in a speaking turn without asserting a relationship between them (i.e., merely listing) is not integration. We found that these

relationships fell into one of three categories: *comparisons*, *interacting considerations*, and *one impacts the other*. In the first case, one theme is claimed to be more important or justified than another. Examples in the transcripts included the claim that quality of life matters more than length of life (integrating Quality of Life and Length or Life) and the claim that effectiveness of a drug should matter but not the age of those who would be treated with it (integrating Effectiveness and Patient Traits).

Interacting considerations arise when it is suggested that two or more themes need to be considered together when deciding whether to fund or delist a drug. Examples include claiming that drugs that extend life are good only if they give a decent quality of life (integrating Quality of Life and Length of Life) or suggesting that a drug should be delisted if a cheaper and more effective drug becomes available (integrating Alternative Treatments, Cost, and Effectiveness). The third type of integration occurs when one theme is said to have implications for or effects on another. Examples here include claims such as side effects matter for quality of life (integrating Quality of Life and Side Effects) and the code 3 example above (integrating Cost and Number Treated).

Normally, integration is symmetrical, that is, if theme X is integrated with theme Y, then Y is also integrated with X. However, an exception can occur when one or more of the themes being integrated is first raised in the discussion by the person doing the integration. For example, suppose quality of life has already been raised by Sue, and Joe then suggests that quality of life is related to side effects, which had not previously been raised. Then we have Side Effects as an integration partner of Quality of Life, but not vice versa. However, if a Quality of Life-Side Effects relationship is subsequently asserted by anyone else, then Quality of Life also becomes an integration partner of Side Effects. This asymmetry is a consequence of the information elaboration instrument, wherein any code higher than 1 requires an item or theme be posed by one discussant and responded to by another. The rationale here is that information elaboration involves people engaging with ideas raised by others, not merely individuals explaining their views.

Our information elaboration instrument can be applied only given a codebook of relevant themes. Because public deliberations frequently seek to elicit perspectives that may not have been considered by experts or policy makers, our codebook could not be predetermined and instead was constructed from the transcripts themselves. Our approach was to develop a codebook of themes from the Halifax transcripts and then to use this codebook for the transcripts from Hamilton and Montreal. Thematic analysis (cf. Braun & Clarke 2008; Nowell et al. 2017) was conducted by three members of the research team (DS, NB, and RT), hereafter referred to as the thematic analysis team. To start off, the thematic analysis team read the final report of the CPAC deliberations (Bentley et al. 2017) and then focused specifically on transcripts from the Halifax, Nova Scotia, deliberation, beginning with the transcript of the introductory large group session, in which participants were provided with background information about the deliberation topic. Next, the small group 1 transcript was separately examined by members of the thematic analysis team to inductively identify recurrent themes and sub-themes occurring in answers to the deliberation question ('What should guide decisions about whether to fund new cancer drugs, or change the funding provided for existing cancer drugs?'). The thematic analysis team met weekly to compare and discuss identified themes, using diagrams to represent connections among themes, until an initial

consensus was reached on which themes to include and on how to define and delineate the themes. This process was then sequentially repeated with the transcripts of the two remaining Halifax small groups and the subsequent large group session.

Clear distinctions between separate themes were preferred to facilitate consistent theme identification. Sometimes this led to consolidating themes. For example, initial themes Cost to Patient, Cost to Government, and Cost Effectiveness were combined into one theme, Cost, with more specific cost-related issues retained as sub-themes. Iterative refinements ultimately resulted in the production of the list of themes and sub-themes that comprised our codebook (see **Table 1**). With the exception of Prevention, we limited themes to answers to the question participants were asked to deliberate upon. We made an exception for Prevention because it was a focus of sustained discussion in small group 2 and the subject of a recommendation in the large group session.

Given the codebook, the thematic analysis team proceeded to code the transcripts from the three small groups and the subsequent large group of the Halifax transcripts using the 0 to 3 point information elaboration coding instrument described above. The transcripts were coded in the same order as in the thematic analysis (i.e., the three small groups in numerical order followed by the large group). For each theme, the code on the information elaboration instrument and its integration partners (i.e., the other themes integrated with it) were recorded. For example, in the Halifax small group 1 transcript, the theme Cost received a code of 3 on the information elaboration instrument and had a single integration partner, Pills versus IV. After each Halifax transcript was coded, the three members of the thematic analysis team met to compare their results and to come to consensus in all cases of disagreeing codes. After the coding of the Halifax transcripts was completed, NB and RT independently coded transcripts from the Hamilton and Montreal deliberations.

Coding results of the Hamilton and Montreal transcripts were used to assess inter-rater reliability. We assessed inter-rater reliability for the information elaboration instrument and for integration partners. Because our information elaboration instrument uses a 0 to 3 coding scale, we used the weighted Cohen's kappa, which takes into account degrees of disagreement, for instance, by making the difference between 0 and 3 count for more than 1 versus 2. Rounding to two significant digits, the weighted Cohen's kappa for the pooled Hamilton and Montreal transcripts was 0.69, which falls in the 'substantial agreement' category and comfortably above the conventional 0.41 threshold of adequacy.

We used the ordinary (unweighted) Cohen's kappa for the assessment of inter-rater reliability of integration partners, because this is a yes or no judgment (theme X is coded as integrated with theme Y or it is not). For the pooled Hamilton and Montreal transcripts, the rate of agreement was 91% and Cohen's kappa was 0.49, both above the conventional thresholds. However, the distribution of Cohen's kappa was not uniform across the transcripts.

For the first two Hamilton small group transcripts, the pooled Cohen's kappa was only 0.16. In response to these low inter-rater reliability scores, the research team leader (DS) developed a document providing detailed guidance on when, and when not, to code two themes as integrated. Inter-rater reliability improved markedly in the subsequent transcripts, where the pooled Cohen's kappa was 0.57. These inter-rater reliability results show that consistent coding of deliberation transcripts can be achieved with the approach we propose here, although consistency in coding integration partners requires careful guidance.

## 5. Operationalizing Diverse Discussion

In this section, we explain how data generated as described in the previous section can be used to operationalize the concept of diverse discussion. A diverse discussion raises and integrates a number of different ideas. In our transcript analyses, the relevant ideas are represented by the 13 themes in the codebook that we developed. Themes raised in a sub-discussion are

indicated in our data by a score on the information elaboration instrument of 1 or greater. Our data also indicate integrations among different themes in two ways. A score of 3 on the IE instrument indicates that the theme was integrated with at least one other theme. In addition, we recorded which other themes each theme was integrated with. These data allow for relatively straightforward assessments of the diversity of a discussion.

Consider the three concurrent small group sessions from the Halifax deliberation. In the first small group, 8 out of the 13 themes were mentioned in the discussion, 5 of these were integrated with another theme. Among themes that were raised, the average number of integration partners was one and no theme had more than three. By contrast, in the second small group, all 13 themes were raised and each was integrated with at least 1 other theme. The average number of integration partners was approximately 3.31, with the maximum being 8. In the third small group, 10 out of the 13 themes were raised, 9 were integrated with at least 1 other topic. Among those themes that were raised, the average number of integration partners was 3.4, and the maximum was 7.

Given these numbers, the discussions in the second and third small groups appear clearly more diverse than that in the first small group. In the second and third discussions, more themes were raised, more were integrated with other themes, and the average number of integration partners was higher. A plausible case can also be made for ranking the discussion in the second small group as more diverse than that of the third, as the second raised more themes. However, the third group had a slightly higher average number of integration partners among themes that were raised. This illustrates that evaluations of diverse discussion can involve a tradeoff when one discussion raises fewer themes than another but more thoroughly integrates them. Resolving such tradeoffs requires deciding upon a measure for ranking the diversity of discussion.

To reflect the underlying theoretical construct of diverse discussion as one in which many themes are raised and integrated, such a measure should increase when (a) the information elaboration score of a theme increases and (b) a theme adds an integration partner. The sum of the information elaboration scores and integration partners is a simple measure that possesses these two properties. For the three small group discussions in the Halifax deliberation, this sum is 16, 56, and 44, respectively, affirming the judgment that the first discussion was clearly less diverse than the others, while suggesting that the second small group discussion was somewhat more diverse than the third.<sup>1</sup> In addition, one would expect large group sessions to be more diverse than the separate parallel small group discussions that feed into them, as the large group would be likely to consider ideas raised in the several small groups. This expectation is borne out in the Halifax transcripts, wherein the diversity score for the large group session following the 3 small groups we analyzed was 57 in comparison to an average score of approximately 38 for the 3 small groups.

Most measures of diversity increase with the number of categories present and the extent to which they are represented in equal proportions (McDonald & Dimmick 2003; Solanas et al. 2012). The measure of diverse discussion described above has the first of these features but not the second. That is raising more themes in a discussion (*ceteris paribus*) increases the sum of information elaboration scores and integration partners, but no measure of their proportions is incorporated. It is, moreover, somewhat unclear what proportions would be relevant to the underlying theoretical construct of diverse discussion. One might consider the proportions of speaking turns, so that a discussion is more diverse when each theme is raised in an equal proportion of turns. Alternatively, the thought might be that in a diverse discussion themes should be integrated in roughly equal depth. In this case, the proportion for each theme might be its number of integration partners out of the total number of integrations. However, the connection between the underlying theoretical construct of diverse discussion and equality of proportions in either speaking turns or numbers of integration partners is questionable. Specifically, equality of either of these proportions is possible, and in fact even likely, in discussions that are intuitively not very diverse. For example, consider the first small group from the Halifax deliberations, in which the discussion was not very diverse. In this case, there was very little difference in the number of integration partners among themes raised for the simple reason that the average number of integration partners was 1 and none had more than 3. In general, non-diverse discussions are likely to involve a leveling down wherein themes are addressed and integrated at minimal, and therefore nearly equal, levels. Consequently, we do not suggest a measure of diverse discussion that considers equality of proportions.

However, we do not insist upon the sum of information elaboration scores and integration partners as the uniquely best approach to measuring diversity discussion.

One alternative would be to take the sum of information elaboration scores alone to measure the diversity of discussion, and to not include the counts of integration partners. Another option would be to place less emphasis on a quantitative measure of diverse discussion and focus instead on a thematic analysis of the various perspectives that emerge, where perspectives are understood as constellations of themes linked together in a specific way. For example, in small group 2 of the Halifax deliberation, the themes Oral Pills over IV, Quality of Life, Cost, and Equity were integrated with the rationale that pills patients can take at home can enhance quality of life by reducing the need for frequent travel to a medical center, which is costly and difficult for people living in rural areas. One could then assess whether different perspectives arose in separate small groups and how these were transmitted to subsequent large group sessions.

A number of fruitful research questions can be explored using operationalization of diverse discussion we propose. Because diverse discussions are desirable in public deliberations, it would be good to know more about causes that promote them. For example, studies could ask whether more demographically diverse groups tend to have more diverse discussions, and if factors such as equality of speaking turns and facilitator style impact the diversity of discussion. Additionally, one could ask whether diverse discussions tend to be of higher quality as measured, for instance, by DQI. It might also be of interest to examine the relationship between diverse discussion and some sub-components of DQI, such as level of justification. Such an inquiry would be of considerable theoretical interest given that philosophical accounts of ideal deliberation often implicitly suppose that diverse discussions are more rational or objective.

## 6. Concluding Discussion

In this article, we have developed the theoretical construct of diverse discussion and proposed a method for operationalizing it, which we illustrate by means of transcripts from public deliberations on cancer drug funding in Canada. While the concept of diverse discussion plays an important implicit role in literature on ideal deliberation, it has previously received little explicit attention from a theoretical or empirical perspective.

New theoretical constructs and measures should be introduced with care. In their paper on DQI, Steenbergen et al. (2003: 23) suggest the following criteria for an adequate measure: '(1) it should be theoretically grounded, (2) it should tap into observable phenomena, (3) it should be general, and (4) it should be reliable.' The method of measuring diverse discussion we propose satisfies these four criteria. The theoretical grounding is explained in section 2, and sections 3 through 5 document its link to observable phenomena in the form of transcripts from public deliberations. The measure is general in so far as it could be applied to many public deliberations besides the ones we discuss here. Indeed, the method could be applied to almost any discourse for which transcripts are available. Finally, we showed that satisfactory inter-rater reliability can be achieved with our method.

There are some limitations to the method we propose. Our measure of diverse discussion does not include an assessment of how in-depth each integration was, neither does it track perspectives involving several themes interconnected in a specific way. Thus, our approach to measuring diverse discussion might be supplemented with an assessment of the depth to which themes are integrated (cf. Kooij-de Bode et al. 2008: 312) or with a qualitative analysis of persistent constellations of themes that correspond to important perspectives in the discussion. Our approach to measuring diverse discussion also does not attempt to identify psychological mechanisms that underlie integrating ideas or resistance to doing so. Nevertheless, by explicating the concept of diverse discussion and providing a method for operationalizing it, we hope to further research that promotes understanding and ultimately implementation of public deliberations in health policy contexts and elsewhere.

## Note

- <sup>1</sup> However, our measure of diverse discussion does not identify a qualitative threshold between diverse and non-diverse discussions. We also caution that our measure is best viewed as providing an ordinal ranking rather than a cardinal measure.

## Competing Interests

The authors have no competing interests to declare.

## Author Contributions

Steel led the conceptualization of the project, supervised the coding and did the bulk of the writing. Bolduc and Titoria did the coding work, were involved in discussions about the conceptualization of the project throughout, and wrote some sections of the manuscript. Burgess was a co-leader on the conceptualization of the project and provided extensive feedback on drafts. Abelson and Peacock provided access to the transcript data, engaged in discussions about the conceptualization of the project, and provided feedback on drafts.

## References

- Abelson, J., Forest, P. G., Eyles, J., Smith, P., Martin, E., & Gauvin, F. P.** (2003). Deliberations about deliberative methods: Issues in the design and evaluation of public participation processes. *Social Science & Medicine*, *57*(2), 239–251. DOI: [https://doi.org/10.1016/S0277-9536\(02\)00343-X](https://doi.org/10.1016/S0277-9536(02)00343-X)
- Anderson, E.** (2006). The Epistemology of Democracy. *Episteme*, *3*, 8–22. DOI: <https://doi.org/10.3366/epi.2006.3.1-2.8>
- Anderson, E.** (2013). *The Imperative of Integration*. Princeton, NJ: Princeton University Press.
- Bentley, C., Abelson, J., Burgess, M., Peacock, S., Demers-Payette, O., Lavis, J., & Wilson, M.** (2017). *Making Fair and Sustainable Decisions about Funding for Cancer Drugs in Canada*. Retrieved from <https://cc-arcc.ca/wp-content/uploads/2017/10/EN-Making-Fair-and-Sustainable-Decisions-about-Funding-for-Cancer-Drugs-Final-Report-March-31-2017.pdf>



- Bentley, C., Costa, S., Burgess, M., Regier, D., McTaggart-Cowan, H., & Peacock, S.** (2018). Trade-offs, fairness, and funding for cancer drugs: key findings from a deliberative public engagement event in British Columbia, Canada. *BMC Health Services Research*, *18*, 339. DOI: <https://doi.org/10.1186/s12913-018-3117-7>
- Bentley, C., Peacock, S., Abelson, J., Burgess, M., Demers-Payette, O., Longstaff, H., Tripp, L., Lavis, J., & Wilson, M.** (2019). Addressing the affordability of cancer drugs: Using deliberative public engagement to inform health policy. *Health Research and Policy Systems*, *17*, 17. DOI: <https://doi.org/10.1186/s12961-019-0411-8>
- Blacksher, E., Diebel, A., Forest, P., Dorr Goold, S., & Abelson, J.** (2012). What is Public Deliberation? *Hastings Center Report*, *42*(2), 14–17. DOI: <https://doi.org/10.1002/hast.26>
- Braun, V., & Clarke, V.** (2008). Using Qualitative Analysis in Psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. DOI: <https://doi.org/10.1191/1478088706qp063oa>
- Brown, M.** (2006). Survey Article: Citizen Panels and the Concept of Representation. *Journal of Political Philosophy*, *14*(2), 203–225. DOI: <https://doi.org/10.1111/j.1467-9760.2006.00245.x>
- Burgess, M., Longstaff, H., & O'Doherty, K.** (2016). Assessing Deliberative Design of Public Input on British Columbia Biobanks. In S. Dodds & R. Ankeny (Eds.), *Big Picture Bioethics: Developing Democratic Policy in Contested Domains* (pp. 263–278). Springer. DOI: [https://doi.org/10.1007/978-3-319-32240-7\\_13](https://doi.org/10.1007/978-3-319-32240-7_13)
- Carman, K., Mallery, C., Maurer, M., Wang, G., Garfinkel, S., Yang, M., Gilmore, D., Windham, A., Ginsburg, M., Sofaer, S., Gold, M., Pathak-Sen, E., Davies, T., Siegel, J., Mangrum, R., Fernandez, J., Richmond, J., Fishkin, J., & Siu Chao, A.** (2015). Effectiveness of public deliberation methods for gathering input on issues in healthcare: Results from a randomized trial. *Social Science & Medicine*, *133*, 11–20. DOI: <https://doi.org/10.1016/j.socscimed.2015.03.024>
- Costa, S., Bentley, C., Regier, D., McTaggart-Cowan, H., Mitton, C., Burgess, M., & Peacock, S.** (2019). Public Perspectives on Disinvestments in Drug Funding: Results from a Canadian Deliberative Public Engagement Event on Cancer Drugs. *BMC Public Health*, *19*, 977. DOI: <https://doi.org/10.1186/s12889-019-7303-2>
- Degeling, C., Carter, S., & Rychetnik, L.** (2015). Which public and why deliberate? A scoping review of public deliberation in public health and health policy research. *Social Science & Medicine*, *131*, 114–121. DOI: <https://doi.org/10.1016/j.socscimed.2015.03.009>
- Degeling, C., Rychetnik, L., Street, J., Thomas, R., & Carter, S.** (2017). Influencing health policy through public deliberation: Lessons learned from two decades of Citizens'/community juries. *Social Science & Medicine*, *179*, 166–171. DOI: <https://doi.org/10.1016/j.socscimed.2017.03.003>
- De Vries, R., Stanczyk, A., Ryan, K., & Kim, S.** (2011). A framework for assessing the quality of democratic deliberation: Enhancing deliberation as a tool for bioethics. *Journal of Empirical Research on Human Research Ethics*, *6*(3), 3–17. DOI: <https://doi.org/10.1525/jer.2011.6.3.3>
- Edwards, P., Hindmarsh, R., Mercer, H., Bond, M., & Rowland, A.** (2008). A Three-Stage Evaluation of a Deliberative Event on Climate Change and Transforming Energy. *Journal of Public Deliberation*, *4*(1), Article 6.
- Fishkin, J.** (2018). *Democracy when the people are thinking: Revitalizing our politics through public deliberation*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oso/9780198820291.001.0001>
- Habermas, J.** (1990). *Moral Consciousness and Communicative Action*. Translated by S. Weber Nicholsen & C. Lenhardt. Cambridge: MIT Press.
- Harrison, D. A., & Klein, K. J.** (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199–1228. DOI: <https://doi.org/10.5465/amr.2007.26586096>
- Himmelroos, S.** (2017). Discourse Quality in Deliberative Citizen Forums – A Comparison of Four Deliberative Mini-publics. *Journal of Public Deliberation*, *13*(1), Article 3. DOI: <https://doi.org/10.16997/jdd.269>
- Homan, A., van Knippenberg, D., van Kleef, G., & De Dreu, C.** (2007). Bridging Faultlines by Valuing Diversity: Diversity Beliefs, Information Elaboration, and Performance in Diverse Work Groups. *Journal of Applied Psychology*, *92*(5), 1189–1199. DOI: <https://doi.org/10.1037/0021-9010.92.5.1189>
- Kooij-de Bode, H., Knippenberg, D., & van Ginkel, W.** (2008). Ethnic diversity and distributed information in group decision making: The importance of information elaboration. *Group Dynamics: Theory, Research, and Practice*, *12*(4), 307–320. DOI: <https://doi.org/10.1037/1089-2699.12.4.307>
- Longino, H.** (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton: Princeton University Press. DOI: <https://doi.org/10.1515/9780691209753>
- Longino, H.** (2002). *The Fate of Knowledge*. Princeton: Princeton University Press. DOI: <https://doi.org/10.1515/9780691187013>
- Longstaff, H., & Burgess, M.** (2010). Recruiting for representation in public deliberation on the ethics of biobanks. *Public Understanding of Science*, *19*(2), 212–224. DOI: <https://doi.org/10.1177/0963662508097626>
- McDonald, D., & Dimmick, J.** (2003). The conceptualization and measurement of diversity. *Communication Research*, *30*(1), 60–79. DOI: <https://doi.org/10.1177/0093650202239026>
- Medina, J.** (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination*. New York: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199929023.001.0001>
- Meyer, B., Shemla, M., & Schermuly, C.** (2011). Social category salience moderates the effect of diversity

- faultlines on information elaboration. *Small Group Research*, 42(3), 257–282. DOI: <https://doi.org/10.1177/1046496411398396>
- Muldoon, R.** (2016). *Social Contract Theory for a Diverse World: Beyond Tolerance*. London: Routledge. DOI: <https://doi.org/10.4324/9781315545882>
- Nowell, L., Norris, J., White, D., & Moules, N.** (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, 16, 1–13. DOI: <https://doi.org/10.1177/1609406917733847>
- O'Doherty, K., & Burgess, M.** (2009). Engaging the Public on Biobanks: Outcomes of the BC Biobank Deliberation. *Public health genomics*, 12(4), 203–215. DOI: <https://doi.org/10.1159/000167801>
- O'Doherty, K., & Burgess, M.** (2013). Public deliberation to develop ethical norms and inform policy for biobanks: Lessons learnt and challenges remaining. *Research Ethics*, 9(2), 55–77. DOI: <https://doi.org/10.1177/1747016113488858>
- Page, S.** (2017). *The diversity bonus: How great teams pay off in the knowledge economy*. Princeton: Princeton University Press. DOI: <https://doi.org/10.2307/j.ctvc77c0h>
- Rowe, G., Marsh, R., & Frewer, L.** (2004). Evaluation of a deliberative conference. *Science, Technology, & Human Values*, 29(1), 88–121. DOI: <https://doi.org/10.1177/0162243903259194>
- Smith, G.** (2009). *Democratic Innovations: Designing Institutions for Citizen Participation*. Cambridge, UK: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511609848>
- Solanas, A., Selvam, R., Navarro, J., & Leiva, D.** (2012). Some common indices of group diversity: Upper boundaries. *Psychological Reports: Human Resources & Marketing*, 111(3), 777–796. DOI: <https://doi.org/10.2466/01.09.21.PR0.111.6.777-796>
- Steel, D., Fazelpour, S., Gillette, K., Crewe, B., & Burgess, M.** (2018). Multiple Diversity Concepts and their Ethical/Epistemic Implications. *European Journal for Philosophy of Science*, 8(3), 761–780. DOI: <https://doi.org/10.1007/s13194-018-0209-5>
- Steenbergen, M., Bächtiger, A., Spörndli, M., & Steiner, J.** (2003). Measuring political deliberation: A discourse quality index. *Comparative European Politics*, 1(1), 21–48. DOI: <https://doi.org/10.1057/palgrave.ccp.6110002>
- Steiner, J., Bächtiger, A., Spörndli, M., & Steenbergen, M.** (2004). *Deliberative politics in action: Analyzing parliamentary discourse*. Cambridge, England: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511491153>

**How to cite this article:** Steel, D., Bolduc, N., Titoria, R., Burgess, M., Abelson, J., & Peacock, S. (2022). Diverse Discussion in Public Deliberation on Cancer Drug Funding. *Journal of Deliberative Democracy*, 18(1), pp.17–26. DOI: <https://doi.org/10.16997/jdd.949>

**Submitted:** 04 July 2020

**Accepted:** 01 April 2021

**Published:** 25 March 2022

**Copyright:** © 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Journal of Deliberative Democracy* is a peer-reviewed open access journal published by University of Westminster Press.