BEYOND FACTS AND NORMS: HOW PSYCHOLOGICAL TRANSPARENCY THREATENS AND RESTORES DELIBERATION'S LEGITIMATING POTENTIAL

ARTHUR LUPIA[*]

YANNA KRUPNIKOV[†]

ADAM SETH LEVINE[‡]

The tension between normative approaches, which are constantly in danger of losing contact with social reality, and objectivistic approaches, which screen out all normative aspects, can be taken as a caveat against fixating on one disciplinary point of view. Rather, one must remain open to different methodological standpoints . . . .

—Jürgen Habermas[1]

I. INTRODUCTION

In *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, Jürgen Habermas describes a challenge to modern

1. JÜRGEN HABERMAS, BETWEEN FACTS AND NORMS: CONTRIBUTIONS TO A DISCOURSE THEORY OF LAW AND DEMOCRACY 6 (William Rehg trans., MIT Press ed. 1996) (1992).

democracies and a procedure for adapting to this challenge. The challenge is that in the absence of natural law, or any other universally accepted moral or ethical code, no common framework informs people about what kinds of laws are, and are not, legitimate. Hence, if laws are to be accepted by, and hence binding on, the populations for whom they are intended to apply, an alternate legitimating mechanism is required.

Habermas describes communicative procedures that have the potential to provide legitimacy to collective decisions. In this ideal discourse ("ID"), as we describe it, all citizens have an equal right to speak and an obligation to listen.[2] If deliberations culminate in agreement on the validity of socially relevant moral, ethical, or technical propositions, then the propositions serve as substantive foundations for subsequent legislative decisions and bureaucratic actions.[3]

This ID fuels legitimacy by providing citizens with confidence that all assent to the fundamental propositional basis of law.[4] The procedures also assure citizens that if subsequent evidence calls previously validated propositions into question, the discourse remains open to reconsidering past agreements.[5] In other words, the ID generates legitimacy for collective decisions not only by giving people the ability to assent to potentially foundational social propositions but also by providing the population with a means for keeping these propositions perpetually accountable to new evidence and to the continuing evolution of diverse worldviews.[6]

Here, we make a parallel argument. Every attribute ascribed to Habermas's argument above will be true here as well. The difference between the two efforts is that we argue that an expanded domain for transparency is required for deliberation to generate legitimacy in many cases. We argue that greater transparency about factors that *Between Facts and Norms* does not examine directly can help increase the legitimating potential of deliberation. The factors in question pertain to how actual and potential deliberation participants think about one another and how such thoughts affect what they choose to say.

One reason that understanding how people think and act matters is that a deliberative endeavor's legitimating potential depends on its ability to counter power asymmetries. If citizens believe that seemingly equalizing

---

2.   *Id.* at 122–23.
3.   *Id.* at 151.
4.   *Id.* at 38–39.
5.   *Id.* at 179–80.
6.   *Id.*

discourse procedures are, in actuality, rigged in favor of more powerful actors, then any propositions generated by such procedures may not be seen as the same propositions that would emerge if the discourse were not so rigged. In such cases, citizens may lack confidence that all truly assent to an alleged propositional basis of law. Citizens may lack confidence that the deliberative process is truly open to new evidence. In such cases, deliberation's legitimating potential is diminished.

We identify factors that cause seemingly ideal discourses to magnify, rather than reduce, the effect of preexisting social inequalities on communicative outcomes. These factors operate via basic communicative phenomena such as who interrupts whom, how the order in which people speak affects subsequent elocutionary decisions, and how different kinds of people are judged after they say something that others find disagreeable. Our findings imply that if deliberation advocates' expectations about communicative outcomes do not remain open to evidence of such communicative phenomena, then the power imbalances that advocates believe they are stopping at the front door of a deliberative chamber are barging in unnoticed through the chamber's back door. We argue that remaining perpetually open to relevant psychological evidence provides an improved basis for revising deliberative designs and recovering deliberation's legitimating potential.

It is important to note that we do not view this project as a critique of Habermas's thesis. Instead, we see it as a natural extension of *Between Facts and Norms*'s argument and as necessary to achieve its goals in light of evidence that has emerged since he initially formulated his argument—evidence that calls into question the reliability of common idealized propositions about deliberative communication.

That said, our argument does include a critique. Today, prominent advocates of deliberative projects base claims about their endeavors' consequences on idealized notions of how deliberators think and behave. Consider, for example, a claim about how a national holiday dedicated to deliberation will change citizens' thought processes:

> After spending a day in deliberation, most will begin to appreciate the shallow basis for their prior views. . . . [T]hey will hold their opinions in a different spirit—one befitting a responsible citizen who refuses to jump to a knee-jerk conclusion but has taken the time and trouble to think seriously about the public good.[7]

---

7. BRUCE ACKERMAN & JAMES S. FISHKIN, DELIBERATION DAY 183 (2004).

Inherent in such claims are multiple statements about how deliberators are thinking. As such, their thoughts can be evaluated in the light of, and demonstrated to be inconsistent with, psychological evidence.

When evidence undermines the validity of such claims, many of the promises of deliberation cannot be kept. These broken promises have tangible consequences. These promises are often made to foundations (in exchange for funding) and citizens (in return for their participation), each of whom have limited resources. Broken promises leave foundations and citizens without the benefits they had hoped for. Such outcomes are also tragic because the psychological phenomena that prevent such promises from being kept have been knowable for a number of years and could have been incorporated into deliberative designs.

To better support the normative project of providing legitimacy to collective decisions, idealized conjectures about deliberative participants' behavior should be evaluated with respect to evidence on relevant psychological and communicative phenomena. Remaining open to such evaluations can clarify the extent to which specific deliberative mechanisms can—and cannot—facilitate collective legitimacy.

We continue as follows. In Part II, we present key aspects of Habermas's argument without criticism. In Part III, we argue that achieving deliberation's full legitimating potential requires expanding the domain of procedural transparency to include openness to evidence about how people think when they communicate with one another. In Part IV, we compare our findings to related works. In Part V, we conclude.

To augment our Part III argument, we integrate findings from a number of empirical research streams. Our criteria for selecting studies on which to focus were as follows. We looked for studies that (1) provide evidence of basic psychological attributes of intergroup communication that are relevant to deliberation and (2) have been in the public domain (that is, published in reputable outlets) for at least five years. The rationale for focusing on psychology, rather than existing empirical work on deliberation or juries, is to identify basic communicative phenomena that exist in many contexts. These basic phenomena provide a basis for explaining how an expanded range of variations in the design of existing deliberative institutions can affect deliberation's legitimating potential. Our rationale for focusing on published research is, in the spirit of Habermas's proposal, that it leads us to empirical claims that have already survived considerable public scrutiny and whose validity has been evaluated with respect to subsequent evidence. Within this set of psychological findings,

we focus further on studies that show how gender—and to a lesser extent, race—affect communicative outcomes. This focus is motivated by a common normative concern with deliberation—its potentially unequal treatment of women and minorities. Focusing on these groups provides us with a concrete basis for showing how greater attention to relevant psychological and communicative phenomena can clarify and improve deliberation's legitimating potential.

In sum, many normative claims about deliberation's legitimating potential are based on idealized premises about how people think and act. There is ample evidence that many of these premises are unreliable. If the idealized premises are known to be sufficiently unreliable, then the legitimating potential of many deliberative endeavors is imperiled. To recover deliberation's legitimating potential, we should remain open to relevant evidence and open to reconsidering past decisions about the normative implications of various deliberative institutions. Expanding the domain of procedural transparency to account for such evidence, rather than ignoring it, is consistent with the normative project of using deliberative democracy to provide legitimacy to collective decisions.

II. HABERMAS'S MODERN CHALLENGE

[D]emocratic will-formation does not draw its legitimating force from the prior convergence of settled ethical convictions.

—Jürgen Habermas[8]

In a democracy, effective governance requires various kinds of collective action. Professional legislators must communicate and coordinate, as must groups of bureaucrats and supporters of particular candidates or policies. The material consequence of such collective endeavors is a series of laws and a set of mechanisms for the laws' implementation and enforcement.

For laws to provide desired outcomes, citizens must perceive them in certain ways. In many contexts, it is helpful if citizens regard laws as legitimate. If they do not, the laws may be followed less often and less effective in creating reliable expectations about key aspects of social life. What factors make laws legitimate?

For many societies, natural law supplies legitimating foundations. In these cases, an all-encompassing theology or more localized stories of

---

8.   HABERMAS, *supra* note 1, at 278.

divination provide a set of basic social facts on which judgments of right and wrong are based, and on which coherent legal theories are built. The foundations of most legal traditions of many modern nations can be traced to such origins.

In recent centuries and decades, however, there has been a growing challenge to the validity of these origins. Nationalisms that may have originated in contexts of dominant theologies now seek to have influence that transcends those theologies. At the same time, changes in the economies of scale of information transmission and travel expose people to an increasing range of cultures and worldviews. The growth of nationalism and changes in "who can communicate what to whom and when" has helped to fuel skepticism of traditional ways of thinking. In places where such skepticism is allowed to become part of the social discourse, questions are raised about whose notions of right and wrong should inform (or bind) collective decisionmaking.

This is the circumstance that Habermas examines. He seeks a method for generating legitimacy in the absence of a preexisting universally accepted moral code and in the presence of positive attitudes toward the possibility of progress through change. Habermas's thesis is that this challenge of modernity can be managed through the adoption of a procedural intervention. He argues that "democratic procedure should ground the legitimacy of law."[9] The procedural intervention is a form of deliberation. Habermas, quoting Frank Michelman, describes the intervention or deliberation as "refer[ring] to a certain attitude toward social cooperation, namely, that of openness to persuasion by reasons referring to the claims of others as well as one's own."[10]

A deliberative strategy can be implemented in many different ways. Habermas asks readers to contemplate the consequences of an ideal discourse. He offers a set of necessary conditions for an ID to exist. These are "necessary conditions under which legal subjects in their role of enfranchised citizens can reach an understanding with one another about what their problems are and how they are to be solved."[11]

One necessary condition that can offer legitimacy to collective decisionmaking is that all citizens have an opportunity to participate and that people treat all others as having equal participatory rights. According

---

9.   *Id.* at 151.

10.   *Id.* at 273 (quoting Frank I. Michelman, *Conceptions of Democracy in American Constitutional Argument: The Case of Pornography Regulation*, 56 TENN. L. REV. 291, 293 (1989)).

11.   *Id.* at 445.

to Habermas, "[T]his principle explains the performative meaning of the practice of self-determination on the part of legal consociates who recognize one another as free and equal members of an association they have joined voluntarily."[12] The actual participation of all is not the necessary condition. Habermas assumes that citizens "concede one another the right to *remain* strangers."[13] Rather, the necessary condition is that all citizens have the opportunity to participate.

Hence, citizens make choices about how involved to be. They need not be directly involved in other aspects of the legislative process, such as formal lawmaking or bureaucratic rulemaking. Habermas's treatise spells out a rationale justifying a professional legislature and a bureaucracy to whom citizens can delegate extensive authority: "[T]he communication structures of the public sphere relieve the public of the burden of decision making; the postponed decisions are reserved for the institutionalized political process."[14]

The requirement of any such professional institutions is that they be continuously connected to the public sphere through deliberative endeavors that generate valid social propositions on which laws should be based, and to which laws should remain responsible.

> Discourses conducted by representatives can meet the condition of equal participation on the part of all members only if they remain porous, sensitive, and receptive to the suggestions, issues and contributions, information and arguments that flow in from a discursively structured public sphere, that is, one that is pluralistic, close to the grass roots, and relatively undisturbed by the effects of power.[15]

Habermas does not presume that deliberation will necessarily change how participants feel about particular policies or government in general. He works hard to distance his thesis from those that would impose on citizens a particular mindset towards governance. In *Between Facts and Norms*, citizens pursue their lives as they see fit. They can agree to disagree and they can consent to having important decisions made by others.[16]

Once such participation rights are established, deliberation can begin. In the ID, a subsequent set of procedures guides the deliberations.

---

12. *Id.* at 110.
13. *Id.* at 308.
14. *Id.* at 362.
15. *Id.* at 182.
16. *Id.* at 103.

> [W]e assume that conditions of communication obtain that (1) prevent a rationally unmotivated termination of argumentation, (2) secure both freedom in the choice of topics and inclusion of the best information and reasons through universal and equal access to, as well as equal and symmetrical participation in, argumentation, and (3) exclude every kind of coercion—whether originating outside the process of reaching understanding or within it—other than that of the better argument, so that all motives except that of the cooperative search for truth are neutralized.[17]

While perhaps restrictive in the sense that everyone is obligated, at least in principle, to consider the claims of all others, it is not restrictive in the sense that a wide manner of considerations can be introduced. People are encouraged to present public issues from the perspective of the lives in which they live, even if these kinds of lives are not lived by many others. Those whose viewpoints would be seen by many as fringe instead of mainstream are invited to offer their point of view.[18]

Once basic speech rights and communicative norms are established, there is the question of information aggregation. Namely, what are participants to do with the array of claims and arguments made? Habermas's procedures allow for debate and disagreement. He describes a communicative outcome in which participants will seek to offer reasons that could be accepted by all others, including those whose world views are quite different than their own. Legitimacy comes from a communicative procedure in which all participants reach a consensus about the validity of certain propositions.[19]

The only propositions that "count" are ones that "all the participating parties *together* find acceptable."[20] In other words, in order for a proposition offered by some participant to be accepted as a legitimate basis for law, all participants must agree to treat the proposition as valid.[21] Habermas is careful to distinguish validity from truth. That a proposition is treated as valid implies that a group accepts it as a basis for subsequent action. But they need not treat the proposition as true in any absolute sense. All such propositions are regarded as potentially falsifiable by new information.

---

17. *Id.* at 230.
18. *See id.* at 182, 230.
19. *See id.* at 119.
20. *Id.*
21. Seyla Benhabib, *Toward a Deliberative Model of Democratic Legitimacy*, *in* DEMOCRACY AND DIFFERENCE: CONTESTING THE BOUNDARIES OF THE POLITICAL 67, 70 (Seyla Benhabib ed., 1996).

While the ID has other attributes, we have described a set of necessary components. These components are also sufficient for a conversation about the extent to which an ID's legitimating potential persists in settings that are less than ideal.[22]

III. HOW PSYCHOLOGICAL TRANSPARENCY THREATENS AND RESTORES DELIBERATION'S LEGITIMATING POTENTIAL

Normative theories are open to the suspicion that they take insufficient notice of the hard facts . . . .

—Jürgen Habermas[23]

Habermas addresses the challenge of modernity by proposing that an ID can yield collective decisions with desirable normative properties. This proposal can succeed. It can also fail. If human communication and reasoning have properties that Habermas's treatise does not anticipate, then the legitimating power of deliberative endeavors can be different than they would be under ideal conditions. To better understand how this power is lost, and how it can be recovered, we raise questions and offer evidence about common psychological characteristics that affect two critical foundations of deliberation's legitimating potential: the absence of communicative coercion in Part III.A and the possibility of unanimous assent in Part III.B.

In Part III, we not only shed light on how individual pieces of evidence related to coercion and assent affect deliberation's legitimating potential, we also clarify cumulative effects of these legitimacy-altering factors. We conclude in Part III.C by arguing that a decision to remain open

---

22. In *Deliberative and Democratic Legitimacy*, Joshua Cohen pursues a broader-ranging treatment of deliberation and legitimacy, meaning that *Between Facts and Norms* is not his only material starting point. Cohen offers "five main features" that characterize deliberative democracy: (1) "A deliberative democracy is an ongoing and independent association," (2) "free deliberation among equals is the basis of legitimacy," (3) "[t]he members have diverse preferences, convictions, and ideals concerning the conduct of their own lives," (4) "[members] prefer institutions in which the connections between deliberation and outcomes are evident to ones in which the connections are less clear," and (5) "[t]he members recognize one another as having deliberative capacities, i.e., the capacities required for entering into a public exchange of reasons and for acting on the result of such public reasoning." Joshua Cohen, *Deliberation and Democratic Legitimacy*, *in* DELIBERATIVE DEMOCRACY: ESSAYS ON REASON AND POLITICS 67, 72–73 (James Bowman & William Rehg eds., 1997). Our work in this paper focuses on his second, fourth, and fifth features. We argue that even in circumstances where deliberative participants believe themselves to be sympathetic to deliberation's normative ends, common psychological phenomena inhibit coercion-free conversation and the extent to which members see one another as equal. We conclude that remaining open to evidence of such phenomena can improve deliberation's legitimating potential.

23. HABERMAS, *supra* note 1, at 6.

to such evidence provides a basis for adjusting communicative expectations and revising the design of deliberative institutions in ways that advance the normative project of conferring legitimacy to collective decisions.

A. IS COERCION-FREE DELIBERATION PSYCHOLOGICALLY REALISTIC?

Habermas builds his argument based on a vision of "the public use of unhindered communicative freedom . . . as enabling rational opinion- and will-formation: the free processing of information and reasons, of relevant topics and contributions is meant to ground the presumption that results reached in accordance with correct procedure are rational."[24] He sought to distinguish "the consensus-achieving force of a communication aimed at reaching understanding" from "the capacity for instrumentalizing another's will for one's own purposes."[25] Such distinctions include propositions about what people are thinking about when they converse with others—and just how free they are to think particular thoughts or express certain ideas in deliberative contexts.

What is coercion-free processing and would we know it if we saw it? If processing is not free because the manner in which people think about a particular topic is influenced by the social context in which they communicate with others, then how far—and in what directions—can participants deviate from free processing and still have their deliberations confer legitimacy on collective outcomes?

The main product of deliberation—the coding of social propositions as valid or invalid—requires participants and nonparticipants alike to proffer judgments on such questions. Deliberation's legitimating potential depends on the extent to which people see the procedures as having certain qualities. A key quality is the lack of coercion.

In contrasting his thesis with that of Immanuel Kant and Jean-Jacques Rousseau, Habermas argues,

> Both [Kant and Rousseau's] conceptions miss the legitimating force of a discursive process of opinion- and will-formation, in which the illocutionary binding forces of a use of language oriented to mutual understanding serve to bring reason and will together—and lead to convincing positions to which all individuals can agree without coercion.[26]

---

24. *Id.* at 147.
25. *Id.* at 148.
26. *Id.* at 103.

The question then becomes "to what extent can any communication be coercion-free?" We contend that it is difficult for communication to be coercion free because most communicative relationships entail at least two important sources of coercion—dependence and persuasive power. Furthermore, the ID's unanimity requirements can introduce additional coercive force. We now address each of these factors in turn.

1. When Can Dependence Introduce Coercion Into Deliberation?

By "dependence," we mean the fact that conversations do not happen in isolation. Conversations occur among people who have relationships that can or do persist outside of the conversation. In many cases, conversations entail an expectation of future interactions, in which the relevant future under consideration can range from a few seconds from now to an entire lifetime.

To visualize how dependence affects communication, think about a speaker who makes an argument to a listener. An idealized view of communication posits that the listener's reaction is based solely on the argument's content. But what if the listener hopes to have further conversations or other relationships with the speaker or others in attendance after the conversation ends? If there are some responses to the content that will affect the tenor or likelihood of future relations, the listener may have an incentive to factor such phenomena into the response. It may mean presenting oneself as more agreeable to a statement than one would be if the conversation could be held in isolation of all other matters or people (if one seeks good future relations), or it could mean presenting oneself as less agreeable (if one has the opposite desires). The main point is that conversations tend not to occur in isolation of other aspects of a social relation. Dependence raises questions about the extent to which processing is "free" in many situations.

Consider, for example, the work of Serena Chen, David Shechter, and Shelly Chaiken, which shows that when people discuss a range of topics, particularly those that may be controversial, they are impression motivated and seek to "satisfy current social goals."[27] In order to have a pleasant interaction, impression-motivated individuals express attitudes that fit opinions of their discussion partners.[28] These people are also more likely to

---

27.   Serena Chen, David Shechter & Shelly Chaiken, *Getting at the Truth or Getting Along: Accuracy- Versus Impression-Motivated Heuristic and Systematic Processing*, 71 J. PERSONALITY & SOC. PSYCHOL. 262, 263 (1996).

28.   *Id.* at 272.

view consensus information as a normative cue about what attitudes they ought to express. Chen, Shechter, and Chaiken show that such effects occur even amongst people who appear to be fully engaged in a conversation and who want to form well-reasoned, "objectively valid judgments" characterized by a "relatively impartial, open-minded treatment of information."[29]

A more general theme of research on information processing that impression motivation influences information processing and the judgments that result. Desires to please others can deeply affect information processing, with such effects often going undetected.[30] Hence, pressures to adhere to group conversational norms and trends cannot be separated from understanding how people think about and react to what others have said.

2. When Can Persuasive Power Introduce Coercion into Deliberation?

By persuasive power, we mean attributes of the source, content, or context of a message that makes others more likely to pay attention to it and attempt to reconcile it with their previous beliefs. Individuals are not equally endowed with such persuasive powers. Habermas recognized this fact explicitly.[31]

An important way in which persuasive power allows communicative coercion and undermines attempts to mitigate preexisting power imbalances is by affecting the extent to which men and women, and racial majorities and minorities, have equal opportunity to contribute to deliberations that, at least on their face, are procedurally fair. We will discuss four manifestations of such effects: (1) how people speak, (2) how people who speak are perceived by others, (3) how people react to others who have said something with which they disagree, (4) and the frequency with which people choose to interrupt one another.

---

29.   *Id.* at 263.

30.   *See generally* Robert B. Cialdini & Noah J. Goldstein, *Social Influence: Compliance and Conformity*, 55 ANN. REV. PSYCHOL. 591 (2004) (reviewing literature in compliance and conformity, and exploring the effect of outside influences).

31.   Habermas argues:

   For as soon as the public space has expanded beyond the context of simple interactions, a differentiation sets in among organizers, speakers, and hearers; arenas and galleries; stage and viewing space. The actors' roles that increasingly professionalize and multiply with organizational complexity and range of media are, of course, furnished with unequal opportunities for exerting influence. But the political influence that the actors gain through public communication must ultimately rest on the resonance and indeed the approval of a lay public whose composition is egalitarian.

HABERMAS, *supra* note 1, at 363–64 (emphasis omitted).

a. Perceptions of Female Speakers

In conversation, individuals often expect men and women to behave differently.[32] A common expectation is that women will be less competent than men.[33] Such expectations shape how many women speak. Women tend to speak more tentatively (for example by giving more disclaimers) in the presence of men.[34] Women are less likely to promote themselves or speak about their own expertise or accomplishments than men.[35] They are also less likely to display emotions of anger or aggressive behavior,[36] and take fewer potentially beneficial conversational risks such as telling jokes or making humorous commentary.[37] In general, "[m]embers of a group who have lower external status characteristics than their groupmates have a difficult time achieving proportionate influence over group decisions."[38]

Gender-based competence expectations also affect how men hear women. In a series of group decision settings, Melissa Thomas-Hunt and Katherine Phillips first identified participants' relative skills and then examined the participants' reactions to "expert" women.[39] They observed

---

32. *See* Monica Biernat & Diane Kobrynowicz, *Gender- and Race-Based Standards of Competence: Lower Minimum Standards but Higher Ability Standards for Devalued Groups*, 72 J. PERSONALITY & SOC. PSYCHOL. 544, 550 (1997) (discussing differing expectations and treatment of male and female applicants in an employment context); Joseph A. Bonito, *A Social Relations Analysis of Participation in Small Groups*, 70 COMM. MONOGRAPHS 83, 84 (2003) (examining participation and participator judgments in small same- and mixed-sex groups); Cecilia L. Ridgeway, *Status in Groups: The Importance of Motivation*, 47 AM. SOC. REV. 76, 85–86 (1982) (examining the influence of male and female members upon a group based on their external status and perceived motivation).

33. Some women are expected to be particularly incompetent. For example, Susan Fiske and her coauthors found that people expect housewives to be largely incompetent, rating this group approximately as competent as the elderly and the disabled. Susan T. Fiske et al., *A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceived Status and Competition*, 82 J. PERSONALITY & SOC. PSYCHOL. 878, 887. Over a series of studies, Fiske found that this rating of housewives to be one of the most stable group ratings. *Id.* at 886. Furthermore, businesswomen and feminists were rated highly competent, but were judged to be less tolerant and less sincere than housewives. *Id.* at 886–87.

34. Linda L. Carli, *Gender, Language and Influence*, 59 J. PERSONALITY & SOC. PSYCHOL. 941, 941–42 (1990) (briefly reviewing past studies on the relationship between status and language).

35. Laurie A. Rudman, *Self-Promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management*, 74 J. PERSONALITY & SOC. PSYCHOL. 629, 629–30, 634 (1998).

36. Ursula Hess, Sylvie Blairy & Robert E Kleck, *The Influence of Facial Emotion Displays, Gender, and Ethnicity on Judgments of Dominance and Affiliation*, 24 J. NONVERBAL BEHAV. 265, 267, 279–80 (2000).

37. Dawn T. Robinson & Lynn Smith-Lovin, *Getting a Laugh: Gender, Status and Humor in Task Discussions*, 80 SOC. FORCES 123, 125, 139–42 (2001) (finding that in an interactive group setting, men are (1) more likely than women to tell jokes; and (2) more likely than women to use humor in an aggressive and status-improving manner).

38. Ridgeway, *supra* note 32, at 76.

39. Melissa C. Thomas-Hunt & Katherine W. Phillips, *When What You Know Is Not Enough:*

that expert women were perceived to be significantly less competent than similarly expert men. And while expert men had more influence on others' behaviors and group outcomes than did nonexpert men, expert women were not more influential than their nonexpert counterparts. Thomas-Hunt and Phillips concluded that "the combination of receiving lower performance assessments than men for similar contributions and reductions in their own confidence may diminish the influence of women experts on their groups."[40]

High-status women face other challenges. Lisa Sinclair and Ziva Kunda examined reactions of men towards women whose statements could threaten male participants' self-esteem.[41] In one example, they show that female managers who provide negative feedback are later rated as significantly less competent than male managers who made identical statements.[42] There were no corresponding differences in competence judgments when managers gave positive feedback.[43] In other words, women get no credit for complimenting men (men treat the compliment as deserved). However, women's reputations are harmed when their words threaten men (men are more apt to treat the negative feedback as reflecting on the woman's competence rather than their own). Similar results were found in another study examining white respondents' reactions to statements by blacks.[44] They found that antiblack stereotypes about competence emerged in white subjects after the "out-group" member (black participant) said something disagreeable to a member of the "in group" (white participant), but not after blacks made agreeable statements.[45]

b. Conversational Turn Taking

The relevance of these communicative expectations to deliberation's legitimating power becomes greater when they shape communicative behavior. A critical conversational behavior is turn taking—in particular, who interrupts whom and when. Lynn Smith-Lovin and Charles Brody

---

Expertise and Gender Dynamics in Task Groups, 30 PERSONALITY & SOC. PSYCHOL. BULL. 1585, 1594 (2004).

    40.   *Id.* (citation omitted).

    41.   Lisa Sinclair & Ziva Kunda, *Motivated Stereotyping of Women: She's Fine If She Praised Me But Incompetent If She Criticized Me*, 26 PERSONALITY & SOC. PSYCHOL. BULL.1329, 1331, 1333 (2000) (confirming through three experiments the hypothesis that "women will be judged less competent than men when they provide negative feedback but not when they provide positive feedback").

    42.   *Id.* at 1336.

    43.   *Id.*

    44.   Ziva Kunda et al., *The Dynamic Time Course of Stereotype Activation: Activation, Dissipation, and Resurrection*, 82 J. PERSONALITY & SOC. PSYCHOL. 283, 287 (2002).

    45.   *Id.* at 289.

found that while women interrupt both genders equally, men are much more likely to interrupt women.[46] Moreover, when men interrupted other men, they did so with positive, reinforcing interruptions, while men interrupted women more often with negative interruptions.[47]

Many studies document the magnitude of such conversational imbalances. In one well-known study, Don Zimmerman and Candace West recorded a wide array of conversations from contexts including coffee shops, drug stores, and various places in a university community.[48] They then analyzed same-sex and cross-sex conversations separately.

In cross-sex conversations, 96 percent of all interruptions entailed a man interrupting a woman.[49] In other words, the ratio of men interrupting women to women interrupting men was 23:1. Moreover, despite finding same-sex interruption patterns to be clustered, meaning that some same-sex conversations had many interruptions while others had few, the asymmetric pattern of cross-sex interruptions was more uniform.[50] Many men in the sample interrupted women at comparable rates.[51]

Zimmerman and West also examined what happened to the conversations after the interruptions. They found another gender-specific difference. After being interrupted by a member of the other sex, men continued talking.[52] Women fell silent. In the few instances in which a woman interrupted a man, "there was no ensuing silence prior to the male speaking again" after the woman finished interrupting.[53] By contrast, nearly 70 percent of women who were interrupted by a man fell silent for more than three seconds after the conclusion of his interrupting utterance—an amount of time that is typically long enough for a person to forfeit their next turn to speak in ensuing conversation.[54]

---

46. Lynn Smith-Lovin & Charles Brody, *Interruptions in Group Discussions: The Effects of Gender and Group Composition*, 54 AM. SOC. REV. 424, 430 (1989).

47. *Id.* at 431.

48. Don H. Zimmerman & Candace West, *Sex Roles, Interruptions and Silence in Conversation*, *in* LANGUAGE AND SEX: DIFFERENCE AND DOMINANCE 105, 111 (Barrie Thorne & Nancy Henley eds., 1975).

49. *Id.* at 116.

50. *Id.*

51. *Id.* at 116–17.

52. *Id.* at 118.

53. *Id.*

54. *See generally* Michael B. Walker, *Smooth Transitions in Conversational Turn-Taking: Implications for Theory*, 110 J. PSYCHOL. 31, 32 (1982) (discussing studies on the duration of pauses within conversation).

c. Further Implications

So far, we have described how implications of dependence and persuasive power affect how men and women think and act when conversing with one another. When we combine these findings with research on conversational-order effects, the reliability of idealized versions of how people communicate is further challenged. For example, even when people are highly engaged in a conversation, the order in which things are said has significant effects on the rest of the conversation and subsequent opinions.[55] Under a "primacy effect," the first piece of information introduced in a conversation tends to have larger effects on postcommunicative opinions.[56] Alternatively, a "recency effect" occurs where the second piece of information has a larger effect on postcommunicative opinions.

More generally, things said earlier in a conversation tend to influence the subsequent content of the conversation more than things said later.[57] Hence, the earlier in a conversation that someone speaks, the more likely it is that his views will be remembered and referenced as the conversation continues. As Martin Pickering and Simon Garrod's work shows, it is natural for individuals in dialogue to engage in "interactive alignment"—that is, even when people are engaged in a dialogue and motivated to achieve a mutual understanding, they are also cognitively predisposed to align their comments with those that came before.[58]

What are the consequences for deliberation's legitimating potential when we combine the content of the order-effects studies with the gender-inequality-in-communication studies described above? One implication is that people who speak earlier will have a greater effect on the subsequent discussion. If men are far more likely to interrupt women than women are to interrupt men, and if order effects make things said early in a conversation more influential than things said later, then actual opportunities for influence can be very unequal even in a deliberative

---

55. Curtis P. Haugtvedt & Duane T. Wegener, *Message Order Effects in Persuasion: An Attitude Strength Perspective*, 21 J. CONSUMER RES. 205, 211, 214 (1994) (discussing two experiments examining message order effects within an attitude strength framework and finding a significant primacy effect where the personal relevance of the message topic was high and a significant recency effect where relevance of the message topic was low).

56. *Id.* at 205.

57. *See* Martin J. Dennis & Woo-Kyoung Ahn, *Primacy in Causal Strength Judgments: The Effect of Initial Evidence for Generative Versus Inhibitory Relationships*, 29 MEMORY & COGNITION 152 (2001) (presenting several experiments showing primacy effects in causal learning).

58. Martin J. Pickering & Simon Garrod, *Toward a Mechanistic Psychology of Dialogue*, 27 BEHAV. & BRAIN SCI. 169, 169–72, 174, 177 (2004).

environment that appears to offer all participants equal opportunities to speak and be heard. When men think less of women after women say something with which the men disagree, the conversational inequality is further magnified.

If citizens understand that deliberative mechanics are so skewed, then they may lack confidence in deliberative processes and outcomes. They may come to believe that coercion has been introduced into the deliberative chamber—that certain people are not as free to think particular thoughts or express certain ideas as others. In such cases, deliberation's legitimating potential is diminished. The first step in recovering this lost potential is to replace, in communicative expectations and in the design of deliberative institutions, evidence-inconsistent communicative idealizations with premises about communication that are more evidence consistent. We will return to such endeavors in Part III.C.

3. When Can the Ideal Discourse's Unanimity Criterion Introduce Coercion into Deliberation?

Above and beyond the coercive potential of persuasive power and dependence is the coercive potential found in the ID's unanimity criterion. Habermas, referencing the work of Rousseau and Kant, explains that "the claim to legitimacy on the part of a legal order built on rights can be redeemed only through the socially integrative force of the 'concurring and united will of all' free and equal citizens."[59] With this criterion, a new question arises: How many of the participants must agree? In the ID, agreement is needed at several points. Unanimous agreement is required not only on whether any given proposition is valid, but also on the decision to enter a deliberative chamber as a means of legitimating social consensus in the first place and on the decision to stay in the chamber once the discussions start.

Under what conditions would people agree to participate in such an endeavor? In other words, what must people believe about the behaviors of others, the time commitment, and the likely consequence of the endeavor in order to be willing to participate? While many deliberation advocates cite public spiritedness as being sufficient to induce participation,[60] there is

---

59. HABERMAS, *supra* note 1, at 32.

60. *See generally* ACKERMAN & FISHKIN, *supra* note 7 (arguing that public spiritedness would be sufficient to induce millions of people to participate in deliberation if it were organized and compensated); E. ALLAN LIND & TOM R. TYLER, THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE (1988) (arguing that people like having an opportunity to express their views before public decisions are made).

reason for doubt. An innovative set of experiments that offered citizens an opportunity to deliberate and interact with members of Congress found substantial interest in deliberations among people who feel alienated from more traditional forms of representative government.[61] By contrast, other scholars have examined how much people would be willing to pay to deliberate as a means of estimating the depth of interest in such endeavors. Mathew McCubbins and Daniel Rodriguez, for example, offered opportunities to deliberate on collectively beneficial topics.[62] The key part of their design was to vary the cost of participating. They found that even tiny costs had a significant negative effect on how groups performed on the same tasks: participants bearing costs did not perform as well as those who could freely deliberate.[63] Judge Richard Posner's rebuttal to Bruce Ackerman and James Fishkin's *Deliberation Day*, which proposes giving everyone a day off from work and paying them $150 each to deliberate national issues, supports this view.[64] Posner argues that if deliberation was seen as inherently valuable "you wouldn't have to pay voters to do it."[65]

Unanimity requirements also introduce forms of dependence. In many cases, decisions to participate or remain in a conversation cannot be decoupled from broader social relationships. People may enter a deliberative chamber because they want to establish or reinforce relationships with others that are largely or wholly independent of the content of the deliberative session.[66] But what happens if someone goes into the chamber and then wants to leave because they are disappointed in the proceedings, or for some other reason? In *Between Facts and Norms*, people are free to remain strangers.[67] But to the extent that a person wishes to continue relationships with others in the deliberative chamber, they will feel some pressure to stay. This force may be sufficient to induce people to remain part of the conversation even though they would exit if the

---

61.   Michael A. Neblo et al., *Who Wants to Deliberate—And Why?*, 104 AM. POL. SCI. REV. 566, 574 (2010) (explaining that "[y]ounger people, racial minorities, and lower-income people expressed significantly *more* willingness to deliberate" than other demographic groups) (footnote omitted).

62.   Mathew D. McCubbins & Daniel B. Rodriguez, *When Does Deliberating Improve Decisionmaking?*, 15 J. CONTEMP. LEGAL ISSUES 9, 19–22, 25–29 (2007) (offering each participant of a group experiment ten dollars for every math problem answered correctly by every member of the group, thereby encouraging deliberation).

63.   *Id.* at 30–31 & n.82 (finding that charging even a dollar for participants to receive information significantly decreased participation in deliberation). *See also* JOHN RH. HIBBING & ELIZABETH THEISS-MORSE, STEALTH DEMOCRACY: AMERICANS BELIEFS ABOUT HOW GOVERNMENT SHOULD WORK 187 (2002).

64.   ACKERMAN & FISHKIN, *supra* note 7, at 3.

65.   Richard Posner, *Smooth Sailing*, LEGAL AFF. Jan.–Feb. 2004, at 40, 41.

66.   *See supra* Part III.A.1.

67.   *See supra* note 13 and accompanying text.

deliberation and all other relationships could be separated in everyone's perceptions and memories.

The presence of people who are coerced into staying and the absence of people who choose to leave a deliberative conversation because of how it progresses raise important questions about the extent to which unanimous assent can be claimed. In particular, should leaving because of unhappiness with the conversation mean forfeiting one's voice? Since participants who leave for such reasons may be reticent to say so, for reasons of dependence, the explanations that they give for their departures will be of limited help in determining whether unanimous assent was actually attained. The coercive possibilities inherent in seeking unanimous assent introduce further challenges to deliberation's legitimating potential.

B. DOES UNANIMOUS ASSENT REQUIRE SHARED MEANINGS?

[O]nly those reasons count that all the participating parties together find acceptable.

—Jürgen Habermas[68]

A key element of the legitimating procedure described in *Between Facts and Norms* is that the assent of all is required to validate a statement as a propositional basis of law. Moreover, it is assumed that when all assent to accept a proposition, the proposition's content is universally understood. This conclusion follows from an assumption about the relationship between thoughts and sentences. Habermas argues that "[t]he important point is that we can read the structure of thoughts from the structure of sentences"[69] and that "[e]very complete thought has a specific propositional content that can be expressed by an assertoric sentence."[70]

In this part, we first examine the extent to which shared meaning is possible. We then proceed to examine the extent to which such possibilities are threatened by people who do not or cannot remain completely focused the content of a conversation, and finally the extent to which common psychological biases complicate the achievement of common meaning.

1. Why Achieving Shared Meaning May Be Difficult

What is the relationship between intent and expression in communicative contexts? Habermas's argument asks us to see this

---

68.   HABERMAS, *supra* note 1, at 119 (emphasis omitted).

69.   *Id.* at 11.

70.   *Id.* at 12.

relationship in a particular way. His view builds from assumptions about psychological foundations of language. Some of these are questionable, such as the assumption that "[t]he members of a language community must proceed on the performative assumption that speakers and hearers can understand a grammatical expression in identical ways."[71]

That people can make this assumption is uncontroversial. Of interest is whether people actually make such assumptions, and whether those who do so are those who do so well served by them. In this subsection, we argue that the range of situations in which people's actions are consistent with such assumptions are more limited than is often appreciated in the deliberation literature.

Consider, for example, the sentence, "I threw the ball." If you throw balls and are never hit by them and if I am hit by balls but have never thrown one, our internal representation of what a ball is and how it relates to other objects or to be will likely be nonidentical. Moreover, when making this claim, I have a mental representation of a ball. My representation (a baseball) need not be—and likely is not—identical to yours (a beach ball or a meatball). So I may say "ball" being perfectly confident that you know what I mean. But I should not be confident that we are sharing an identical meaning. As Paul Churchland explains, "[W]hile we all participate in the richness of sensory life, we struggle to communicate to others all but its coarsest features. Our capacity for verbal description comes nowhere near our capacity for sensory discrimination."[72]

To be sure, people who expect communication to successfully convey meaning must believe that there is a reasonably high likelihood that others interpret a grammatical expression as they do.[73] But they need not believe that others always do so. In fact, evidence that people generally do not believe that others interpret expressions as they do is that people often repeat themselves or they attempt to establish the validity of a proposition using several consecutive and nonidentical utterances, as we are doing in this paragraph. If two people always interpret something the same way, there would be no need to use alternate statements to convey something.

Probabilistic beliefs about what inferences other people draw from utterances are particularly manifest when attempting to converse with people whom we do not know well. Such practices suggest that people use

---

71.   *Id.* at 11.

72.   PAUL M. CHURCHLAND, THE ENGINE OF REASON, THE SEAT OF THE SOUL: A PHILOSOPHICAL JOURNEY INTO THE BRAIN 21 (1995).

73.   PETER J. STEINBERGER, THE CONCEPT OF POLITICAL JUDGMENT 157–58 (1993).

and think about language in a very different way than the ideal described in *Between Facts and Norms*. Habermas explains that "[t]he ideal character of semantic generality shapes communicative action inasmuch as the participants could not even intend to reach an understanding . . . if they did not *presuppose*, on the basis of common (or translatable) language, that they conferred identical meanings on the expressions they employed."[74] As illustrated above, this does not account for real life communicative practices.

A more accurate description about how deliberators use language to confer legitimacy to collective decisions is that they seek convergent understanding rather than identical understanding. Hence, communicative actions depend on probabilistic calculations

of the

> in settings that—to me—resemble a circumstance similar to that in which our conversation now occurs—I perceive a particular utterance to have conveyed my meaning with a sufficient degree of success. I will try it now. If your response suggests that you are coming within some distance of my intended meaning, I will consider this communicative attempt a success and continue as if it succeeded. If your response suggests that your understanding is not sufficiently close to my intent, I will try again using another utterance, or I will give up.

2. How Low Involvement Challenges the Formation of Shared Meanings

A lingering question in the context of using deliberative endeavors to generate legitimacy involves the situation where such nonidentical understandings provide a sufficient basis for regarding a particular proposition as having a single meaning to which all can be said to assent. If the communicative process remains open to settling subsequent disputes about the meaning of a previously accepted proposition, then a nonidentical meaning poses less of a threat to deliberation's legitimating potential. Problems of nonidentical meaning increase when common psychological phenomena, of the type we have described above and of the type that we shall describe below, cause the communicative process to become closed to such inquiries.

For communicative processes to remain open, deliberators must be willing and able to continue to pay attention to, and think about, others' views. In many cases, however, people will do no such things. Many will instead rely on simple information shortcuts or cues. Research on this topic

---

74.   HABERMAS, *supra* note 1, at 19.

shows that even when people are highly involved and motivated to pay attention to one another, they remain highly susceptible to factors other than content of utterances when forming judgments about what was said. Hence, even for participants who appear to be highly engaged in a conversation, their openness to adjusting their views can be severely inhibited, with the groups' abilities to reach common meaning also limited as a result.

Richard Petty, John Cacioppo, and David Schumann offer a framework for thinking about information processing variations.[75] The framework describes two processing routes.[76] On the central route, attitude change is the product of "a person's diligent consideration of information that s/he feels is central to the true merits of a particular attitudinal position."[77] On the peripheral route, attitude change occurs "because the person makes a simple inference about the merits of the advocated position based on various simple cues in the persuasion context."[78]

These authors' findings support the position that when people are more involved, for example, when the issue under discussion is highly relevant for their own lives, their processing is more central.[79] Less involved people tend to rely on peripheral processing and are more influenced by factors other than an argument's content.[80] Given that a deliberative episode brings together individuals who pay differing levels of attention to different parts of a conversation, it is likely that even if they are exposed to identical utterances, they are processing the information in very

---

75.   Richard E. Petty, John T. Cacioppo & David Schumann, *Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement*, 10 J. CONSUMER RES. 135, 135 (1983) (performing an experiment to determine how different arguments impacted subjects with different levels of issue involvement); Richard E. Petty & John T. Cacioppo, *The Effects of Involvement on Response to Argument Quantity and Quality: Central and Peripheral Routes to Persuasion*, 46 J. PERSONALITY. & SOC. PSYCHOL. 69, 78–79 (1984) (same).

76.   Petty, Cacioppo & Schumann, *supra* note 75, at 135.

77.   *Id.*

78.   *Id.*

79.   *Id.* at 141, 143 (arguing that attitudes formed through the central route are more predictive of behavior given that attitudes "were better predictors of behavioral intentions" for high-involvement participants as opposed to those with low involvement). In one experiment, subjects who were highly involved were more impacted by the quality of arguments than low-involvement subjects, and their intentions were more impacted by argument quality than low-involvement subjects. *Id.* at 141. In another experiment, the number of arguments had a weaker impact on high-involvement subjects as opposed to weak-involvement subjects, and strong arguments produced more agreement than weak ones among high-involvement but not low-involvement subjects. So, in high-involvement situations, "attitudes were affected by quality but not by number." Petty & Cacioppo, *supra* note 75, at 75.

80.   *See* Shelly Chaiken, *The Heuristic Model of Persuasion*, *in* 5 SOCIAL INFLUENCE: THE ONTARIO SYMPOSIUM 3, 4 (Mark P. Zanna et al. eds., 1987) (discussing how, according to the heuristic model of persuasion, people use minimal effort when determining the validity of a message).

different ways. At the same time that high-involvement people may be attempting to think through complex implications of what they are hearing, low-involvement people may be making less of an effort and seeking quick cognitive shortcuts.

Common shortcuts that have been shown to have greater influence than a message's content on the opinions regarding low-involvement peripheral processors include the communicative roles of other people, such as whether the argument is attributed to the minority or the majority,[81] or whether others in an audience react positively or negatively.[82] Danny Axsom, Suzanne Yates, and Shelly Chaiken, for example, manipulated three factors simultaneously: (1) audience response (enthusiastic or unenthusiastic), (2) argument quality (strong or weak), and (3) subject involvement (high motivation or low motivation).[83] They found significant interactions between audience response and involvement, and between audience response and argument quality.[84] For low-involvement individuals who had little motivation to systematically process the information, the audience's response was more influential than the content of the message itself.[85]

Some people, such as those with low involvement in the topic, may come to believe an utterance not because they have a shared understanding of the meaning of the content with others in the deliberative chamber, but simply because they perceive that the majority of others seemed to think it was a good idea. Hence, not everyone in deliberative contexts will think about what others say in an identical—or even approximately identical—manner. Reactions to what others say can be affected by factors such as how many other people hold a particular opinion, whether a man or woman is speaking, and the order in which people speak.[86] Such findings are consonant with a concern of Ackerman and Fishkin that "[i]f too high a proportion of [Deliberation Day] participants enter the deliberations with a

---

81. Diane M. Mackie, *Systematic and Nonsystematic Processing of Majority and Minority Persuasive Communications*, 53 J. PERSONALITY & SOC. PSYCHOL. 41, 41, 50 (1987).

82. Danny Axsom, Suzanne Yates & Shelly Chaiken, *Audience Response as a Heuristic Cue in Persuasion*, 53 J. PERSONALITY & SOC. PSYCHOL. 30, 36–37 (1987).

83. *Id.* at 31.

84. *Id.* at 34–37. Based on their data, the authors found that audience response influenced only the opinions of low-involvement subjects and not high-involvement subjects. The authors also found that audience response impacted low-involvement opinions only where the quality was high, although this was somewhat less straightforward.

85. *Id.* at 35. The authors conclude that "low-involvement subjects were more responsive to the audience cue, presumably because the lesser systematic processing they engaged in failed to provide them with information that contradicted the consensus heuristic." *Id.* at 36.

86. *See supra* Part III.A.2.

closed mind, they will spoil the environment for everybody else. They warn that this "can do lasting damage, since a bad initial experience might lead many ordinary Americans to steer clear of future engagements."[87]

While they stated this concern in the context of activists who might attempt a communicative coup, several decades of research in psychology shows that various forms of what some might label closed-mindedness are actually common attributes of information processing. In other words, the idealization of a completely open mind does not square well with accumulated evidence. This is discussed in the following section.

3. How Information-Processing Biases Challenge the Formation of Shared Meanings

The extent to which minds are open is further challenged by decades of research on information processing biases. Two biases, confirmatory bias and perseverance bias, are especially relevant.

Confirmatory bias is a phenomenon where individuals misinterpret new evidence as supporting what they already believe.[88] In Charles Lord, Lee Ross, and Mark Lepper's classic experiment, views about a scientific study are first elicited before subjects were presented with critiques of that study. Holding the study's content constant, they found that people are less critical of the study if its conclusion matched their initial attitudes. In other words, identical information elicits "entirely opposite evaluations from people who hold different initial beliefs."[89] Hence, and as John Bullock argues, exposure to a common stimulus need not lead a group to a common conclusion. If confirmatory biases are present, shared exposure can produce divergent views.[90]

When confirmatory biases are present, and absent strong motivation to question the manner in which they are processing information with which they agree and that which threatens their current understandings, people will not systematically stop to consider other possible points of view.

---

    87.   ACKERMAN & FISHKIN, *supra* note 7, at 128.
    88.   Charles G. Lord, Lee Ross & Mark R. Lepper, *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098, 2100–01 (1979).
    89.   *Id.* at 2102. Where subjects were initial proponents of capital punishment, a study confirming the deterrence effect of the death penalty was significantly more convincing than an antideterrence study regardless of research design criticism. Alternatively, opponents of capital punishments found the prodeterrence study significantly less convincing than a study against the deterrent effect of the death penalty, also regardless of prominent criticisms of the studies. *Id.* at 2100–02.
    90.   John G. Bullock, *Partisan Bias and the Bayesian Ideal in the Study of Public Opinion*, 71 J. POL. 1109, 1117–19, 1123 (2009).

Matthew Rabin calls this phenomenon "hypothesis-based filtering."[91] "While it is sensible to interpret ambiguous data according to current hypotheses, people tend to use the consequent 'filtered' evidence inappropriately as further evidence for these hypotheses."[92] So, if confirmatory biases stifle consideration of diverse points of view in deliberative settings, then they may also limit the extent to which deliberation can confer legitimacy on collective decisions.

Perseverance bias describes a situation in which beliefs persist after supporting evidence is discredited. In a study by Lee Ross, Mark Lepper, and Michael Hubbard, people received a performance evaluation, and then, in a debriefing, were told that their evaluation was completely unrelated to their actual performance.[93] Later, they were asked to rate their ability. Despite the debriefing, people who received a positive evaluation rated their ability much more favorably than those who were evaluated negatively.[94] So, if perseverance biases render deliberative conversations less responsive to new information, then they may also limit the extent to which deliberation can confer legitimacy on collective decisions.

Findings of the kind described above, that previous commitments can affect how people process new information, are also consistent with a philosophy of mind that Quentin Skinner attributes to Hobbes.

> Hobbes came to believe that most people are moved less by force of reason than by their perceived sense of their own self-interest. By contrast with the optimism of *The Elements* and *De Cive*, he additionally insists in *Leviathan* that, if the requirements of reason collide with people's interests, they will not only refuse to accept what reason dictates, but will do their best to dispute or suppress even the clearest scientific proofs if these seem liable to affect their interests in an adverse way.[95]

Hence, these biases are not newly discovered. Instead, the psychological work clarifies when and how these biases affect deliberation's legitimating potential. The value of this evidence is that it can inform communicative expectations and help us design more effective deliberative institutions.

---

91. Matthew Rabin, *Psychology and Economics*, 36 J. ECON. LITERATURE 11, 28 (1998).

92. *Id.* at 28.

93. Lee Ross, Marc R. Lepper & Michael Hubbard, *Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm*, 32 J. PERSONALITY & SOC. PSYCHOL. 880, 882–83 (1975).

94. *Id.* at 883–84.

95. QUENTIN SKINNER, REASON AND RHETORIC IN THE PHILOSOPHY OF HOBBES 427 (1996).

C. ADJUSTMENTS TO CONSIDER

Collectively, differences in involvement and biases in information processing reduces deliberation's ability to produce the kinds of shared meaning that can give a populace confidence that propositions to which all presumably assent are, in fact, understood identically by all. When the challenges to shared meaning described in Part III.B are combined with Part III.A's evidence on communicative gender inequities, it becomes apparent that diverse demographic subpopulations vary greatly in their ability to speak and be taken seriously by others as well as their ability to understand the intended meaning of what others say in public forums.

For example, people who are not prone to speaking early or interrupting are far less likely to have an impact on the conversation than those who are prone to speaking early or interrupting. Moreover, even if people come to a deliberative session with the intention of being open-minded, variations in their level of involvement, as well as confirmatory and perseverance biases, can lead them to pay differential attention to the ideas that are presented in a deliberative forum. Hence, even in deliberative environments that are seemingly ideal—in which all participants have equal opportunities to speak and be heard—regular and predictable psychological phenomena can fail to mitigate preexisting power imbalances.

A cumulative consequence of such phenomena is that when deliberation advocates fail to account for such psychological factors, the power imbalances that they think they are stopping at the front door barge in through the back door. Moreover, if people who are asked to view ID-generated propositions as valid understand that power imbalances affect deliberation in this way, then they will have a reason to see the deliberative endeavor as rigged. In both situations, deliberation can lose its legitimating potential.

We close this section by using psychological research to inform potentially effective adjustments to deliberative institutions. One adjustment entails front loading a deliberative endeavor with team-building exercises in which individuals can develop a basis for trusting one another.[96] This interpersonal foundation can give people a stronger basis for interpersonal trust, which can increase their motivation to try to see

---

96.   Roger C. Mayer, James H. Davis & F. David Schoorman, *An Integrative Model of Organizational Trust*, 20 ACAD. MGMT. REV. 709, 710 (1995) (briefly reviewing work on trust and the work place).

matters from others' points of view—rather than reverting to conscious or unconscious gender stereotypes—when disagreements arise.

Gender imbalances in conversational turn taking can also be reduced by design.[97] Approaches include introducing rules against interruption, making interruption costly (for example by making people press a button, or giving people a limited number of nontradable interruption credits that they can use during a session). These rules make people aware of conversational gender imbalances. They can help people who may be motivated to work toward social progress and shared understandings, but cannot otherwise restrain themselves from responding to those who interrupt.[98]

If gender affects not only interruption patterns but also how people listen to one another, then other adjustments can be beneficial. One approach to limiting the communicative influence of gender (or other) stereotypes is to make individuals aware of stereotype use in the exact deliberative setting in which people are sitting.[99] Role assignment is another approach. Jennifer Richeson and Nalini Ambady show that when white respondents are given titles suggesting leadership or supervisorial power within a group, they are more likely to rely on negative racial stereotypes when evaluating others as compared to the case when no such titles are given.[100] Hence, women and racial minorities may have better deliberative experiences when all members of a group are explicitly assigned equal roles.

Different strategies can reduce confirmatory or perseverance biases. Stefan Schulz-Hardt, Marc Jochims, and Dieter Frey argue that group discussion settings often exacerbate an individual's tendency to rely on biases.[101] Such biases are most influential when group members know at the beginning of a conversation that they have similar opinions about the

---

97.   Emanuel A. Schegloff, *Overlapping Talk and the Organization of Turn-Taking for Conversation*, 29 LANGUAGE IN SOC'Y 1, 22–45 (2000).

98.   In 2009, Lupia served on a jury and, during lunch on the first day of deliberations, he described the interruption literature. Jurors subsequently referenced these statistics many times during the deliberations. Without instruction, they adopted a norm of making sure that everyone had a chance to be heard. When a high-status white male interrupted from that point on, it was noticed and self-corrected. No other sanctions were needed to limit interruptions. These behaviors became a source of camaraderie for the group and a source of unifying humor in the days that followed.

99.   Irene V. Blair, *The Malleability of Automatic Stereotypes and Prejudice*, 6 PERSONALITY & SOC. PSYCHOL. REV. 242, 248–50 (2002).

100.   Jennifer A. Richeson & Nalini Ambady, *Effects of Situational Power on Automatic Racial Prejudice*, 39 J. EXPERIMENTAL SOC. PSYCHOL. 177, 181 (2003).

101.   Stefan Schulz-Hardt et al., *Biased Information Search in Group Decision Making*, 78 J. PERSONALITY & SOC. PSYCHOL. 655, 655–56 (2000).

discussion topic. One method for reducing confirmatory bias in such instances is to create "contrived dissent."[102] For example, one participant can be asked to serve as a "devil's advocate," who "generates counterarguments to [a proposed] solution and tries to identify all weaknesses inherent in it."[103] A variation on devil's advocacy is "dialectical inquiry."[104] Instead of challenging the group only after a proposed solution emerges, dialectical inquiry involves deliberately debating and questioning various opinions and solutions as they are raised by "pitting diametrically opposed assumptions and recommendations . . . against each other."[105] While both approaches have been shown to reduce confirmatory bias in settings where people are predisposed to agree with one another,[106] viewing these findings in light of the gender-effect-neutralizing role assignment suggests such roles must be assigned with care. If it is not possible to neutralize communicative power imbalances through the methods stated above, then contrived dissent methods may reduce bias at the expense of increasing inequality.

In sum, many normative claims about deliberation's legitimating potential are based on idealized premises about how people communicate. If these idealizations are known to be sufficiently incorrect, then the legitimating basis of many deliberative endeavors is imperiled. To recover deliberation's legitimating potential, we should remain open to relevant evidence. We can then use such evidence to inform our communicative expectations and develop more effective deliberative strategies.

IV. RELATION TO PREVIOUS WORK

[T]he ideal model does not take account of the information and decision costs of the communication process itself. It does not consider the limited capacities for cognitive processing afforded by simple horizontal networks of communication; in particular, it abstracts from the unequal distribution of attention, competences, and knowledge within a public. It also ignores attitudes and motives at cross-purposes to the orientation to

---

102.   Stefan Schulz-Hardt, Marc Jochims & Dieter Frey, *Productive Conflict in Group Decision Making: Genuine and Contrived Dissent as Strategies to Counteract Biased Information Seeking*, 88 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 563, 567 (2002).

103.   *Id.*

104.   David M. Schweiger, William R. Sandberg & James W. Ragan, *Group Approaches for Improving Strategic Decision Making: A Comparative Analysis of Dialectical Inquiry, Devil's Advocacy, and Consensus*, 29 ACAD. MGMT. J. 51, 53 (1986).

105.   *Id.*

106.   *See id.* at 52–53 (briefly reviewing studies about dialectical inquiry and devil's advocacy and their effects on decisionmaking, but noting weaknesses in experimental design raised by a different study).

mutual understanding and is thus blind to egocentrism, weakness of will, irrationality, and self-deception on the part of participants.[107]

In this paper, we have summarized Habermas's argument for using deliberative procedures to generate legitimacy for collective decisions in modern societies and we have raised questions about deliberation's legitimating potential. In so doing, we argue that deliberation's legitimating potential is threatened by gaps between Habermas's idealized attributes of human communication as presented in *Between Facts and Norms*, and inferences about these attributes that scholars might draw given better evidence.

We also argue that remaining open to evidence about differences between idealized communication and observed communication is a means for deliberative endeavors to recover legitimating potential. This aspect of our work is complementary to the previous efforts of the authors we discuss below. These authors, in various ways and with diverse objectives, document problems caused by the lack of exchange between normative and empirical literatures. We take a moment to discuss how our work relates to these efforts. In general, the distinction between our work and previous efforts is that we seek to clarify and improve deliberation's legitimating potential by giving greater exposure to psychological evidence that is relevant to this normative aim. Past work either discusses legitimacy without reference to psychological evidence or introduces related evidence without tying it specifically to *Between Facts and Norms*'s normative focus on deliberation's legitimating potential.

Our work relates to that of Dennis Thompson, first and foremost, by following his headline advice. He calls for closer ties between specific normative claims and empirical evaluations—in contrast to empirical arguments that summarily dismiss normative endeavors and normative work that treats empirical scholarship in an analogously derisive manner.[108] Where we part from Thompson is our interest in the mechanics underlying our main normative focus—deliberation's legitimating potential. Legitimacy is one of several normative topics that Thompson covers. On legitimacy, he argues that "[o]ne of the most important benefits that theorists ascribe to deliberative democracy is that the decisions it produces are more legitimate because they respect the moral agency of the

---

107.   HABERMAS, *supra* note 1, at 325.
108.   Dennis F. Thompson, *Deliberative Democratic Theory and Empirical Political Science*, 11 ANN. REV. POL. SCI. 497, 498 (2008).

participants."[109] We agree that deliberative democracy *can* accomplish this goal.

However, Thompson explains that the benefit of legitimacy is "inherent in the [deliberative] process, not a consequence of it. It is not appropriately subjected to direct empirical investigation."[110]In contrast, we contend that legitimacy is not an inherent byproduct of deliberation. Indeed, Webster's Dictionary defines "legitimate" as "*to give* legal status or authorization to."[111] In *Between Facts and Norms*, our normative starting point, the relevant authority for judging deliberation's legitimacy is the populace itself. If citizens believe that conversational dynamics reinforce power imbalances, they may see deliberations as rigged in favor of the powerful. Such perceptions reduce deliberation's legitimating potential.

We view deliberation's legitimating potential as at least partially determined by interactions among basic psychological dynamics and choices about deliberative rules. Hence, given the choice of whether to ignore evidence of such dynamics and interactions or to remain open to the prospect of adjusting communicative expectations and deliberative institutional designs on such evidence, we see the normative project as better served by remaining open.

Like us, John Dryzek focuses on underlying mechanics of deliberation's legitimating potential.[112] He emphasizes that "legitimacy for Habermas is secured by public acceptance of *procedural* responsiveness, not by the actual responsiveness of legislation to the substance of public opinion on an issue."[113] Our efforts differ in which threats to deliberation's legitimating potential we examine. Unlike *Between Facts and Norms*, which asserts that the actual participation of all is not a necessary condition for deliberation to confer legitimacy to subsequent collective decisions,[114] Dryzek argues that deliberation's legitimating potential is threatened by the impracticality of having every member of large modern societies participate in deliberative endeavors.[115] Given the difficulty in reconciling how to "secur[e] legitimacy while respecting the basic constraint of deliberative economy," Dryzek then focuses on alternative deliberative

---

109. *Id.*
110. *Id.*
111. WEBSTER'S NINTH NEW COLLEGIATE DICTIONARY 683 (9th ed. 1988).
112. John S. Dryzek, *Legitimacy and Economy in Deliberative Democracy*, 29 POL. THEORY 651, 657 (2001).
113. *Id.* (emphasis omitted).
114. *See supra* note 13 and accompanying text.
115. Dryzek, *supra* note 112, at 651–52, 657.

institutions that do not require universal participation.[116] He compares their output to the content of what propositions "a whole" would validate if it could gather together.

Our efforts are complementary, but orthogonal for the following reason. While Dryzek's alternatives describe creative ways to expose people to one another's views, the psychological factors we identity do not simply apply to discourses amongst entire polities, they are also present in smaller conversations. Hence, while the evidence we introduce can adversely affect Dryzek's alternative designs' legitimating potential, his normative project can be aided by keeping these designs open to this evidence.

The collaborative effort headed by Jane Mansbridge also complements our work.[117] They argue that "[i]f . . . the exercise of power is inevitable in human politics, then we must . . . design democratic institutions that incorporate that power rather than ignore it."[118] Like us, they recognize it may be impossible to remove all power inequities from deliberation. Hence, their work suggests that in light of this, coercive forces should be reduced to a "practical minimum."[119]

Our approach differs from theirs in that we identify basic psychological phenomena as a source of coercive potential and we present evidence on such matters as a means of updating communicative expectations and informing the design of deliberative institutions. Where we focus on how these coercive forces affect deliberation's legitimating potential, Mansbridge regards noncoercion as a distinguishing attribute of deliberation and, hence, a normative end unto itself.[120] As a result, they do not delve into coercion's psychological foundations. Instead, they seek to defend deliberation vis-à-vis other means of decisionmaking. As is the case with Dryzek's proposed solutions, accounting for the coercive forces present in deliberative communications can improve the legitimating potential of the endeavors that the Mansbridge group endorses.

Tali Mendelberg is also concerned with the role of power and self-interest in deliberative democracy. She begins with the idea that "[d]espite thin or non-existent empirical evidence for the benefits that deliberative

---

116. *Id.* at 653–56 (discussing alternatives such as random sampling or nondemocratic ways to restrict participation).

117. Jane Mansbridge et al., *The Place of Self-Interest and the Role of Power in Deliberative Democracy*, 18 J. POL. PHIL. 64 (2010).

118. *Id.* at 93.

119. *Id.* at 82.

120. *Id.* at 69.

theorists expect, many theorists argue forcefully for more citizen deliberation."[121] Drawing examples from small-group discussion studies, Mendelberg describes ways in which conversations violate deliberative ideals.[122] She offers reason to doubt that ignoring evidence about how people communicate is consistent with the normative aims of many deliberative democrats. Our work builds on hers by identifying—as Thompson recommends—a specific normative attribute of deliberative democracy and then "trying to discover the conditions in which deliberative democracy does and does not work well, while paying more attention to the question of to what extent the unfavorable conditions could change."[123]

Lynn Sanders explores scholarly advocacy of deliberative democracy rather than discussing deliberation's legitimating potential specifically. She highlights an antidemocratic strain of argument supporting deliberation.[124] Sanders recognized that "some Americans are more likely to be persuasive than others" and "some Americans are apparently less likely than others to be listened to," meaning that in reality, discourse is "neither truly deliberative nor really democratic."[125] Like us, she draws attention to gender differences in communication.[126] Like Dryzek and Mansbridge, she concludes by voicing support for an alternative to deliberation. The alternative is testimony, as in "telling one's own story, not seeking communal dialogue."[127]

Our focus on evaluating and improving deliberation's legitimating potential leads us to provide evidence about the psychological mechanisms underlying the behaviors that Sanders describes. So, where Sanders uses evidence of inequality to argue against deliberation, we use evidence about the mechanisms underlying inequality to clarify how institutions can be differently designed to improve deliberation's legitimating potential.

---

121.   Tali Mendelberg, *The Deliberative Citizen: Theory and Evidence*, *in* POLITICAL DECISION-MAKING, DELIBERATION AND PARTICIPATION 151, 154 (Michael X. Delli Carpini et al. eds., 2002)

122.   *Id.* at 155–61.

123.   Thompson, *supra* note 108, at 500.

124.   Lynn M. Sanders, *Against Deliberation*, 25 POL. THEORY 347, 350–51 (1997).

125.   *Id.* at 349.

126.   *Id.* at 367.

127.   *Id.* at 372.

V. CONCLUSION

[T]he idealistic content of normative theories has been evaporating under the sun of social science.

—Jürgen Habermas [128]

Interest in deliberative democracy continues to grow. Its appeal is understandable. Deliberation, with its emphasis on distributed speech rights and information exchange, has the potential to confer important normative benefits. While such benefits are easy to imagine, they can be hard to achieve.

Many advocates describe deliberation as if it is a place where ideas travel from one mind to another unadulterated—as if ideas are absorbed en masse. This is wrong. In human communication, all but the simplest utterances and stimuli are processed, sometimes in normatively unappealing ways. Physical limits in cognitive capacity force people to pay attention to only a tiny fraction of the information to which they are exposed and prevent them from recalling more than a tiny fraction of the things to which they have paid attention.[129] Such observations pose important challenges to idealized claims about deliberation's benefits. Remaining open to evidence about these processes provide new opportunities to better achieve valuable normative goals.

The purpose of deliberation, as *Between Facts and Norms* describes it, is to generate legitimacy for collective decisions while respecting differences in individual worldviews. In *Between Facts and Norms*, Habermas sought to include citizens in the processes of collective decisionmaking, but he did not seek to make unrealistic demands of them. He sought to work from reasonable assumptions about what people want and how they see the world. He then sought to develop communicative mechanisms that permit these views to be converted into legitimate social propositions through transparent procedural means. Our goal is the same.

As Habermas recognized, "Due to their idealizing content, the universal presuppositions of argumentation can only be approximately fulfilled. Moreover, because there is no criterion independent of the argumentative process, one can judge only from the participant's perspective whether these demanding presuppositions have been

---

128.   HABERMAS, supra note 1, at 329.

129.   *See* ESSENTIALS OF NEURAL SCIENCE AND BEHAVIOR 667–94 (Eric R. Kandel, James H. Schwartz & Thomas M. Jessell eds., 1996).

sufficiently fulfilled in a given case."[130] We contend that such demands are more likely to be met if the domain of transparency expands to include deliberation-relevant psychological phenomena.[131]

To what we have argued, one could respond that the ID is just an ideal and should not be evaluated for its relationship to observed patterns of communicative behavior. That may be so, but the fact is that many entrepreneurs are using idealized representations of deliberation to promote their endeavors.[132] Such representations are the basis for appeals to philanthropic foundations that are asked to sponsor such endeavors and they are the basis of promises made to participants. To these foundations and people, the relationship between the "ideal" and observations of the "real" is highly germane.

Another response is that everyone knows that the ideal is unrealistic, but that there is legitimating benefit in attempting to get as close as possible to the ideal. In other words, as a social decisionmaking procedure becomes more similar in structure to the ID, beneficial normative properties that are often associated with the ID are more likely to follow.[133] The evidence we

---

130.   HABERMAS, *supra* note 1, at 178.

131.   Deliberation and Habermas's argument can be evaluated with respect to criteria other than the production of legitimacy. We briefly describe this point here. Jack Knight and James Johnson focused on stability rather than legitimacy, and argue that deliberation's ability to produce desired outcomes depends crucially on institutional arrangements. *See generally* Jack Knight & James Johnson, *Aggregation and Deliberation: On the Possibility of Democratic Legitimacy*, 22 POL. THEORY 277 (1994) (focusing on stability). Knight and Johnson raise questions about deliberation's normative properties by drawing attention to problems of preference aggregation from social choice theory. For other work raising questions about the correspondence between various way of aggregating preferences and the legitimacy of collective decisions see John Ferejohn, *Instituting Deliberative Democracy*, *in* DESIGNING DEMOCRATIC INSTITUTIONS 75, 82–87 (Ian Shapiro & Stephen Macedo eds., 2000) and John W. Patty & Elizabeth Maggie Penn, *A Social Choice Theory of Legitimacy*, 36 SOC. CHOICE & WELFARE 365, 365–67 (2011). We raise questions about deliberation's normative properties by drawing attention to problems of communicative equality and bias from social psychology. Our approaches are complementary, but our focus is logically prior to theirs with respect to normative questions about deliberation's legitimating potential. Social choice theory, which Knight and Johnson rely on, derives conclusions about collective outcomes from premises about individual attributes. The norm in social choice theory is to abstract away from all psychological and communicative phenomena. Arthur Lupia & Mathew D. McCubbins, *Lost in Translation: Social Choice Theory is Misapplied Against Legislative Intent*, 14 J. CONTEMP. LEGAL ISSUES 585, 585–94 (2005). Social choice theory's conclusions are not designed to clarify conditions under which deliberative outcomes reflect unanimous assent of free processors to accept a socially relevant proposition as valid. We seek more psychologically defensible premises about how deliberation affects attitude formation and change. Such premises can inform future iterations of such theorizing that seek to clarify when aggregate legitimacy is logically consistent with relevant psychological and communicative phenomena.

132.   For example, James Fishkin acquired a patent on "deliberative polling." U.S. Patent Application No. 13/337,957 (filed Dec. 27, 2011).

133.   Archon Fung, *Deliberation Before the Revolution: Toward an Ethics of Deliberative Democracy in an Unjust World*, 33 POL. THEORY 397, 400 (2005).

have presented makes this claim problematic. Simply stated, unless power imbalances that are common to communicative relationships are not recognized and reconciled, more is not always better when it comes to deliberation. If deliberation is implemented in conditions in which power relationships and other coercive factors adversely affect decisions to join the conversation, to speak, to remain in the conversation, and to react to what others have said, deliberative outcomes can have a substantially different normative character than some deliberative advocates have claimed.

Deliberative outcomes depend on interactions amongst individuals with nonideal psychological profiles. If legitimacy is to emerge from such interactions, then citizens and participants must have reason to believe that the interactions are not rigged in favor of the powerful and are capable of producing propositions to which all truly assent. To accomplish this goal in light of the evidence described above, the design of deliberative institutions should remain open to publicly validated propositions about how people reason and communicate.