Stereotyping, Discourse, and Speaker Status
From Social Psychology to Democratic Theory

Paper prepared by:
Edana Beauvais, M.A., PhD Candidate (University of British Columbia)
edana.beauvais@gmail.com

Paper prepared for:
"Democratic Citizens and Democratic Practices" Panel
Western Political Science Association Conference
02-04 April 2015
Las Vegas, Nevada, USA

Work in progress, please do not circulate without author's permission

Abstract

There is a growing consensus that democratic legitimacy depends upon the quality of discourse that informs opinion- and will-formation. As we know, moral, legal, and socioeconomic inequality can vitiate discourse, and prevent deliberation from achieving its epistemic, ethical, and democratic functions. What I show in this paper is that even in the general absence of moral and legal inequality, and where socioeconomic inequality has been neutralized, a form of inequality—what I refer to as *discursive inequality*—can still undermine discourse's functions. This paper theorizes the concept of discursive inequality, which refers to when features of cognition or language undermine speaker status, reducing the ability of certain speakers from contributing knowledge or opinions on a subject, impeding the ability for discourse to generate mutual respect, or preventing certain voices from being included at all. I use insights from research in social psychology to show when and how discursive inequalities undermine the functions of discourse. I conclude by offering suggestions for addressing the ways in which stereotypes threaten to undermine the normative functions of discourse.

I. Introduction

"I could have stayed home and baked cookies and had teas, but what I decided to do was fulfill my profession which I entered before my husband was in public life," snapped prominent Arkansas lawyer, Hillary Rodham Clinton. Clinton made the comment to the press in response to the suggestion that her dual role as a lawyer and wife to the 1992 Democratic presidential hopeful, Bill Clinton, represented a conflict of interest. Clinton was trying to express that, given the broadening career opportunities for women, the stereotype of the political wife as a helpmate was outdated. However, her comments were taken as slighting women in traditional roles, and drew sharp criticism. The first lady of the United States is more than just the spouse of the president,[1] she is the symbolic representation of the American wife and mother. As Jackie Judd, a PBS correspondent, commented, "Americans are most comfortable, for example, with first ladies who are gracious, stand-by-your-man wives… Any more than that… makes people uneasy" (Judd & Koppel, 1992). Clinton was made to eat her words: to quell the negative response to her comments she donned an apron and baked chocolate oatmeal cookies on national television.

Public discourse—communicating about matters of common concern that result in either binding decisions by the state, or decisions that are only loosely binding in a societal sense—is supposed to produce better decisions (fully informed, and where different arguments are weighted according to publicly justifiable standards), is meant to promote mutual respect, and should include a diverse range of voices and concerns (Mansbridge et al., 2012). When political and social decisions are reached through discursive processes that achieve these functions, those who are bound or affected by the decisions are more likely view them as legitimate. As we

[1] Presumed (by stereotyping assumptions) to be a heterosexual man.

know, different forms of inequality—moral, legal, and socio-economic—can undermine any one of these functions. For instance, if certain members of society lack equal moral or legal status, they will have a much harder time having their voices heard or weight given to their arguments; if discussion partners are separated by vast gulfs in socio-economic status, discourse can contribute to polarization, rather than mutual respect.

But what about the case of Rodham Clinton from 1992? Her contribution to the discourse on the roles of the wives of high-powered men was not only given little weight, it backfired—it contributed to the counterargument that women *should* stay home and bake cookies. Clinton enjoyed moral and legal equality, and she certainly did not suffer from a reduced economic status. While it is true that—as a woman—she suffered some social disadvantage relative to men, she was uttering her contribution to the discourse in what was a nominally egalitarian context. Despite this, by violating the wife-as-helpmate stereotype, she elicited deep, negative affect. Clearly, the concepts of moral, legal, and socio-economic equality are not sufficient for understanding the ways in which inequality can inhibit the normative functions of discourse. We must also understand how language and features of cognition (such as social categories)—the very currency of communication and rational critique—can threaten to impede discourse from achieving its normative functions. I refer to this form of inequality as *discursive inequality*.

This paper theorizes the concept of discursive inequality, which refers to when features of cognition or language undermine speaker status, reducing the ability of certain speakers from contributing knowledge or opinions on a subject, impeding the ability for discourse to generate mutual respect, or preventing certain voices from being included at all. I begin by outlining the analytic framework of my analysis, which bridges both democratic theory (I use the systemic approach for understanding deliberative democracy) as well as social psychology (I use the

cognitive approach for understanding social categorization and stereotyping). After introducing my theoretical framework, I move on to my analysis—using insights from research in social psychology to show when and how discursive inequalities undermine the functions of discourse. I conclude by offering suggestions for addressing the ways in which stereotypes threaten to undermine the normative functions of discourse, so that we might achieve discursive equality.

II. Theoretical Framework

There is a growing consensus that democratic legitimacy depends upon the quality of discourse that informs opinion- and will-formation. However, as I have begun to suggest, a problem occurs when cognitive or linguistic features vitiate the quality of discourse. Understanding the relationship between legitimacy, discourse, and will-formation, and the ways cognitive or linguistic structures challenge this relationship, requires an analytic framework that continues what Habermas (1984) refers to as the reconstructive tradition of linking normative theory and empirical knowledge. Deliberative democratic theory provides a research paradigm that helps us understand the extent, nature, and effects of discourse, as well as normative standards by which we can judge discourse. In this section I will introduce my theoretical framework for analyzing public discourse, the *systemic approach* to deliberative democracy, as well as offer a brief introduction to the leading paradigm for understanding social categorisation and stereotypes, the *cognitive approach* to stereotyping.

A. A Systemic Approach to Deliberative Democratic Theory

My analysis adopts what Mansbridge et al. (2012, p. 2) call a systemic approach to deliberative democracy, which recognizes that "democracies are complex entities in which a wide variety of institutions, associations, and sites of contestation accomplish political work." The systemic approach for studying deliberation outlines the boundaries of the deliberative system, describes the three primary functions within the system, and offers normative standards by which the system should be evaluated. The boundaries of the system are defined very loosely: the decision-making area must be democratic, and 'decision-making' includes not only the binding decisions of the state and the activities related to preparing those binding decisions, but also societal decisions that are emergent rather than definite, and binding only in the loosely societal sense.

The systemic approach also outlines three general functions of the deliberative system: an epistemic function, an ethical function, and a democratic function. The *epistemic function* refers to whether or not deliberation produces opinions, preferences, and attitudes that are informed by facts and information, and where relevant arguments and concerns are aired, discussed, and "appropriately weighted" (Mansbridge et al., 2012, p. 11). The weighting of arguments and concerns can vary, but should have a degree of publicity (if certain arguments or concerns are given more weight than others, the reason for this weighting should be public or at least publicly justifiable). The *ethical function* of deliberation is to generate mutual respect. I would add that the ethical function can be expanded to include generating other positive sentiments that strengthen democracy and civil society, including interpersonal trust, generalized trust, and mutual liking (positive affect). Some of the literature on small groups (see Mendelberg, 2002a for a review of this literature) and on social capital and associational activity (for instance, see

Portes, 2000; Putnam, 2000) acknowledge these other outcomes. Finally, the *democratic function* refers to how well deliberation includes a diversity of voices, concerns, interests, and arguments.

The deliberative system can be normatively evaluated based on how well different aspects of the system (informal networks, the media, legislatures, the courts, etc.) achieve different, or multiple functions. It is possible for components of the system to perform certain functions well, while failing to achieve the other functions—the overall health of the deliberative system is the sum of how well these interacting components perform these three primary functions of deliberation. So, in short, the purpose of this paper is to describe how features of human cognition can undermine speaker status—in other words—the ways in which discursive inequalities can inhibit discourse from achieving its epistemic, ethical, or democratic functions. But before I illustrate how and when social categorisation impedes the normative functions of discourse, I must briefly introduce my readers to the leading social psychological framework for understanding stereotypes: the cognitive approach.

B. Social Cognition Approaches to Stereotyping

The social cognition approach treats stereotypes as cognitive categories, and focuses on the structure of these cognitive schema as well as on their effects on interpersonal interactions and information processing (David L. Hamilton & Tina K. Trolier, 1986). However, while categorisation is one of the most important and basic processes of human cognition—categorization prevents cognitive overload by simplifying large quantities of information, and provides a way of organizing our perception of the world (Hinton, 2000)—stereotypes can and should be normatively distinguished from other categories. Stereotypes are generated under conditions we find reprehensible (conditions of social inequality), and, as I will show, help to

maintain these conditions by silencing the voices of, or undermining the persuasive force of utterances from members of subordinate groups in public discourses.

So how should category structure (in general) be understood, and how do categories organize information? Within the social cognitive approach to the study of categorization, there are two (compatible) leading theories for understanding category structure: prototype models and associative networks. I will briefly consider each in turn. The prototype model proposes two general principles for the formation of categories. The first of these is familiar—*cognitive economy*. The second is *perceived world structure*, which asserts that objects in the world are perceived to possess a high correlational structure. Rosch (1999) gives the example of a person who, upon perceiving the attributes of feathers, fur, and wings, realizes that feathers co-occur more often with wings than with fur. Of course, as Rosch points-out, our association of feathers with wings is not merely a product of the physical environment, but also the social environment; specifically, that we have already possess a cultural and linguistic category called "birds."

A prototype refers to the 'most typical' member of a category (Operario & Fiske, 2001). The prototype theory of categorization follows Wittgenstein's insight[2] that humans deal with categorization on the basis of "clearest cases," defined as "peoples' judgments of goodness of membership in the category," and *not* by reference to boundaries (Rosch, 1999, p. 36). Thus, category membership is determined by goodness-of-fit in the category (which is subjective, but—although there are variations across individuals—these perceptions tend to be inscribed in culture, and so fairly consistent across members of the same cultural group), and not by strict boundaries delineating category membership. The more prototypical a category member is rated, the "more attributes it has in common with other members of the category and fewer attributes in

[2] See Wittgenstein (1953).

common with members of the contrasting category" (Rosch, 1999, p. 37; see also Rosch & Mervis, 1975). Thus, for instance, a sparrow is a more prototypical (representative) member of the category "birds," while a penguin is a less prototypical category member.

Associative network models is another theory explaining category structure, and it helps clarify some of the basic assumptions of the prototype model. According to associative network model, information is stored in mental structures called nodes, and each node corresponds to a single concept (a name, object, personality trait, place, emotional response, etc.) (Operario & Fiske, 2001; see also Carlston, 1994). Each of these nodes is interconnected by links, which map out meaningful associations between the concepts contained within each individual node. These inter-nodal links are what structure people's mental representations—stronger links denote more significant associations between concepts, while weaker links denote less meaningful associations. According to the theory, "the nature of nodal linkages fluctuates according to perceivers' experiences: Links increase or decrease in strength depending on the perceived correlation between concepts, and new nodal links can develop according to new associations between previously unpaired concepts" (Operario & Fiske, 2001, p. 29).

As I have said, stereotypes are cognitive categories containing assumptions about the attributes of social groups. Following the prototype model, another way of phrasing this is to say that stereotypes are group prototypes, which are "mental representations consisting of a collection of associations between group labels (e.g., Italians) and the features that are assumed to be true of the group (e.g., a[n assumed] feature of Italians might be 'romantic')" (Charles Stangor & Mark Schaller, 1996). The prototype (e.g., of the romantic Italian) is the average or most typical value of group members (the assumption being that most Italians are romantic and Italians are, on average, more romantic than other groups). Associative network models help

explain how processes of categorization generate stereotypes: stereotypes occur from spread of activation, where the excitation of one node (e.g., the social category 'Italian') flows across links to stimulate other nodes (e.g., attributes such as romantic, artistic, wine-loving, etc.) (Operario & Fiske, 2001). Excitation travels faster along stronger links (denoting a more significant association between concepts, or nodes), and nature of nodal linkages fluctuates according to perceivers' experiences.

As such, one's experience in the natural and, more pertinently for my analysis, in the social world structures the linkages between nodes, which can include anything from names, objects, personality traits, places, emotions, etc. Of course, because of my focus on stereotypes, I am particularly interested in the fact that the social world structures linkages between semantic concepts, particularly between social groups and various human attributes (e.g., the link between the social group 'Italian', and the attribute of being 'romantic'). With greater exposure to the co-occurrence of semantic concepts (such as the co-occurrence of Italians and romance), either directly (one had a romantic love affair with an Italian), or indirectly (one has seen the film *Under the Tuscan Sun* too many times), the link between these concepts becomes stronger, denoting a more significant association between these concepts.

This analysis is congruent with findings in the stereotyping literature, particularly studies measuring response times to priming (e.g., rapidly identifying the word "good" after being primed with the word "White"). As Operario and Fiske point out, many of the most well-known experiments in social psychology use response times to priming to study stereotypes (for instance, see Devine, 1989; Higgins, Bargh, & Lombardi, 1985; Perdue, Dovidio, Gurtman, & Tyler, 1990), suggesting that "stereotypes represent the strength between two or more conceptual nodes stored in peoples' mental representations" (2001, p. 30). Because network models focus on

micro-level cognitive processes, the approach suggests that stereotyping occurs implicitly—that is, mostly outside of a person's awareness.

However, even though stereotyping is largely an implicit process, the cognitive and behavioural effects of stereotyping can be identified, and stereotyping can be inhibited. In the following sections I will discuss the consequences of stereotyping for cognition (e.g., information processing, forming judgments) and emotional responses. I will relate this to the normative concerns of democratic theory, by discussing how these discursive inequalities—how stereotypes and the related cognitive and affective consequences of stereotypes—undermine the epistemic, ethical, and democratic functions of discourse. In my conclusion, I will offer some suggestions on how address or neutralize discursive inequalities, so that discourse can achieve its normative functions.

III. Discursive Inequalities: When and how they undermine the Functions of Discourse

In this section I will review when and how discursive inequalities—forms of inequality that are internal to cognition and language, including category structures (i.e., stereotypes)—influence both human reasoning and affective responses, and discuss what this means for democratic theory. By 'category structure' I am referring to the way associations between concepts (such as associations between social groups and particular human attributes) structure judgments of degree of prototypicality (what kind of person constitutes the 'best' or 'most typical' example of their social category). The effects of category structure on human reasoning and inference are called *prototype effects*. In this section I will review four prototype effects: (1) reaction time, (2) asymmetries in similarity ratings, (3) asymmetries in generalization, and (4) asymmetries in probability estimations. Besides impacting human reasoning, stereotypes shape

affective responses. Stereotypes shape affect responses in two ways: directly, by eliciting automatic feelings (e.g., pity, contempt, etc.); and indirectly, by generating particular expectations of category members, which—when the expectations are violated—induce negative feelings aimed at the transgressing category members. I will begin this section by considering how different prototype effects bias human reasoning and inference, and considering how these biases threaten to undermine discourse's epistemic, ethnical, and democratic functions; I will then turn my attention to how stereotypes structure affective responses, and how unconscious affective responses can threaten to undermine discourse's normative functions. In my conclusion I will discuss ways of addressing or mitigating these discursive inequalities.

A. Stereotyping and Cognitive Biases: When Prototype Effects Undermine the Functions of Discourse

Recall that categories are structured in terms of associations (or links) between concepts (or nodes), and that the *strength* of associations between particular concepts—determined by how frequently the links between concepts are activated—structures judgements of prototypicality. That is to say, when links between a concept (e.g., "bird") and other concepts, (e.g., "flight," "chirping," "nesting," etc.) are activated frequently, a person's judgment of a prototypical bird will be one that includes all of those characteristics. This prototypical bird does not have to be a real bird, but in choosing the 'best case' of an actual bird, a given person will select a bird that has these characteristics (such as a sparrow, or a robin). Stereotypes, as I have discussed, are a special instance of categorization, where social groups are associated with particular human attributes. All categories, including stereotypes, are used in reasoning, and the nature of any given category's structure (the association between concepts, or prototypes)

impacts cognition. As Lakoff says, “prototypes act as *cognitive reference points* of various sorts and form the basis for inferences” (1990, p. 45, italics in original) (see also Rosch, 1975, 1983). The effects of category structure on reasoning are called *prototype effects*. As I have said, I will review four important prototype effects: (1) reaction time (or processing speed), (2) asymmetries in similarity ratings, (3) asymmetries in generalization, and (4) asymmetries in probability estimations.

(1) Reaction time

The first of these prototype effects, reaction time (RT), refers to the speed at which respondents are able to associate concepts with categories. For instance, a study in linguistics might ask a respondent to indicate true or false in response to a statement such as “An [example] is a [category name]” (e.g., “a penguin is a bird”) (Lakoff, 1990, p. 41). These studies show that response times are shorter for representative examples (prototypes). This technique has also been used extensively in social psychology to test implicit associations between concepts such as associative linkages between words with positive or negative valence and racial stimuli (e.g. words like “good,” “wonderful,” or alternatively, “bad,” “annoying”; and words or images of members of different racial groups) (see Banaji & Hardin, 1996; Blair & Banaji, 1996; Devine, 1989; Higgins et al., 1985; Perdue et al., 1990).

Unsurprisingly, studies show that RTs are faster when positive words are associated with members of dominant social groups (who are more often positively stereotyped), and RTs are faster when negative words are associated with members of subordinate groups (who are more often negatively stereotyped) (Fazio, 2001; Fazio & Dunton, 1997; Fazio & Olson, 2003). In many instances, the implicit association between positive words and members of higher status groups, and negative words and lower status groups, has been taken as evidence of implicit

prejudice[3] (Greenwald, McGhee, & Schwartz, 1998; Rudman, Greenwald, Mellott, & Schwartz, 1999). However, it seems unlikely that prejudice can account for these findings, since both high- and low-prejudiced individuals exhibit RT effects (including members of subordinate groups) (Arkes & Tetlock, 2004). An experiment by Lepore and Brown (1997) substantiates this intuition. The authors find that RT effects result from stereotype *knowledge*, not from prejudice (stereotype endorsement). In other words, RT effects are generated by awareness of the links between positively or negatively valence words and social groups—not by the endorsement that these associations are good or right.

The RT prototype effect is not simply a curiosity revealed in laboratory settings, but has serious behavioural consequences. For instance, research reveals that there is a racialized 'shooter bias'. In a series of studies respondents were asked to play a video game where they must "shoot" armed targets and not "shoot" targets holding innocuous items. As it turns out, people are much faster on the draw if they believe a Black target is armed with a gun, as compared to if they believe a White target is armed: both White and Black subjects are more likely (and quicker) to shoot Black targets, as compared to White targets (Correll, Park, Judd, & Wittenbrink, 2002, 2007). In another study, Devine (1989) primed subjects with a series of words that cued Black American stereotypes, and then asked subjects to read a vignette describing a social interaction. Subjects who were primed with stereotype-related words perceived greater hostility in the vignette, as compared to the control group.

If the association of certain social groups (typically members of more dominant social groups) with concepts related to good decision-making skills results in members of these social groups being associated with greater expertise or competence, and this would increase the status

[3] Prejudice is a more complex term, which generally refers to explicit endorsement of stereotypes, and contains a component of negative affect (an active *dislike* for members of out-groups).

of speakers from dominant groups and potentially threaten the *democratic function* of discourse (e.g., to produce decisions informed by a diverse range of voices). Evidence from studies of small groups and jury deliberations suggest that implicit associations do vitiate the democratic function of discourse. Studies of jury deliberations show that higher-status members tend to speak more, and to be perceived as more accurate (even though status does not correlate with accuracy) (Hastie, Penrod, & Pennington, 1983). One of the strongest predictors of a person's influence in a discursive interaction is whether other participants in the conversation view them as more expert or competent in the task at hand (Bottger, 1984; Kirchler & Davis, 1986; MacRae Jr, 1993; see Mendelberg, 2002b for a review of this literature).

Strong implicit associations can also undermine the *ethical function* of discourse (e.g., to generate mutual respect). If the content and/ or the context of deliberation cues negative racial stereotypes, then this could trigger more negative interpretations of discussion contributions from members of subordinate groups. Extrapolating from Devine's (1989) study, if negative racial stereotypes of Black Americans are primed during racially segregated deliberations where participants are deliberating desegregated schooling, then White participants may come to interpret the utterances of Black participants in a more negative light (i.e., they may perceive the utterances of Black participants are being more aggressive). This could help explain the mechanism for why, in their study of deliberations on desegregated schooling, Mendelberg and Oleske (2000, p. 186) found that racism produced "strong negative assumptions about the opposing deliberators," and that deliberation increased polarization, rather than contributing to mutual respect.

Finally, strong semantic associations can undermine the *epistemic function* of discourse (e.g., to produce fully-informed decisions, where different arguments were fairly or appropriately

weighted). Since positive attributes such as expertise, competence, and authoritativeness are associated with dominant social groups, while negative attributes such as aggressiveness and incompetence are associated with subordinate groups, these associations might lend greater weight to the arguments put forward by members of dominant groups and reduce the weight given to arguments put forward by members of subordinate groups. Evidence from the study of deliberation appears to substantiate this concern. For instance, Mendelberg et al. (2014; see also Karpowitz, Mendelberg, & Shaker, 2012) find that with majority rule and few women, deliberation produces decisions that are less attentive to women's concerns or opinions.

This is because gendered stereotypes construct women's voices as less authoritative, and when institutions fail to give numbers power (such as majoritarian systems with few women), then women experience a negative balance of interruptions when speaking, and they clam-up (Mendelberg, Karpowitz, & Oliphant, 2014). As such, in majoritarian settings with few women, women's voices are shut out of deliberations—and women's arguments, concerns, and points of views are not reflected in resultant decisions. This clearly undermines the epistemic function of discourse and, speaking to my early point, also undermines the democratic function of discourse. Evidence from studies of small group dynamics suggests that this might also be true of other marginalized groups: for instance, studies of jury deliberations show that Black jurors have less influence in deliberation, they participate less, and they experience less satisfaction with deliberation (Bowers, Steiner, & Sandys, 2001; see Mendelberg, 2002b for a review).

(2) Asymmetries in similarity ratings

The second prototype effect I wish to consider, asymmetries in similarity ratings, refers to the tendency for less representative category members to be considered more similar to

prototypical members than vice versa (Lakoff, 1990; Rosch, 1975; Tversky & Gati, 1978). In a series of experiments, Tversky and Gati show that less salient items (the 'variant') are judged to be more similar to the prototype than vice versa (1978, p. 85). For instance, the authors find that, among subjects in the United States, Mexico (perceived by subject as a less salient example of a country) is judged to be *more* similar to the US (perceived as a more prototypical country) than the US is to Mexico. This pattern holds with ratings of difference: the prototype is expected to differ from the variant more than the variant differs from the prototype. Thus, for instance, the US is judged to be more different from Mexico than vice versa (at least among respondents in the US).

These asymmetries in similarity ratings pose a problem for the 'common in-group identity model' for reducing in-group bias and intergroup discrimination. The common in-group identity model (or, re-categorization model) posits that in-group bias and intergroup discrimination can be reduced by redefining group categories at a higher level of group inclusiveness (Brewer & Gaertner, 2001; Gaertner & Dovidio, 2014; Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993). When there are asymmetries between groups, then superordinate categories might become dominated by the majority (or socially dominant) group, and these asymmetries can make re-categorization more difficult, as minority group members might be unwilling to accept a superordinate category membership that is dominated by the majority group (Brewer & Gaertner, 2001). Asymmetries in similarity ratings can undermine the ethical function of discourse, as dominant group members might expect to receive respect from subordinate group members, without feeling the need to reciprocate symmetrically.

Differences in group hierarchy and salience, and resultant asymmetries in similarity ratings, also matter for the nature of discourse itself. As I have already suggested, language—not

just *what* people say, but *how* they say it—matters a great deal for discourse and deliberation. People can use language to facilitate cooperation and generate an overarching sense of identity, such as by using first names and the first person plural ("we") to create feelings of shared identity (Sornig, 1989). Speech accommodation theory reveals that inter-group discourse can result in either "convergence" (when forms of speech, such as rate of speech, accent, dialect, etc. become more similar) or "divergence" (when differences in forms of speech are accentuated). Convergence increases the perception that speakers are cooperative, friendly, effective, and warm; by contrast, divergence decreases the perception that speakers have these attributes. The re-categorization model suggests that creating an overarching shared identity (as is facilitated by convergence) is exactly what is needed to reduce in-group biases and ease intergroup conflict (Brewer & Gaertner, 2001; Gaertner & Dovidio, 2014; Gaertner et al., 1993).

However, asymmetries in similarity ratings—the perception that less salient groups are more similar to dominant groups than vice versa—likely impedes cooperative language. Speech accommodation theory reveals that there is a linguistic expectation that members of less salient (typically, less dominant or lower status) groups should adopt the speech patterns of more salient groups: people tend to converge *toward* the higher group status (Giles, Mulac, Bradac, & Johnson, 1987). Again, this expectation that members of subordinate groups become 'more like' members of dominant groups (rather than meeting in the middle through mutual accommodation), undermines the ability for discourse to generate mutual respect.

Evidence from speech accommodation theory shows that when social identities are threatened, speech patterns tend to diverge, which generates less positive perceptions of interaction partners (Hogg, 1985; Thakerar, Giles, & Cheshire, 1982). Not only do speech patterns diverge when groups come into conflict, but the Linguistic Intergroup Bias (LIB)—the

tendency for group members to use abstract terms for describing their own positive attributes, concrete terms for describing their own negative attributes, and doing the opposite when describing attributes of out-group members—is heightened (Maass, Corvino, & Arcuri, 1994). And the LIB further undermines feelings of attraction, closeness, and empathy (Rubini & Kruglanski, 1997). Asymmetries in similarity ratings, which generate the *expectation* that less salient groups become more like dominant group members (rather than, say, the expectation that group members meet halfway), can generate speech patterns that undermine the ability to create an overarching group identity. This is particularly a concern when group differences are made salient or when groups feel threatened, such as when there are conflicts over "exclusive group identity" (Warren, 1992, p. 21). This not only undermines the ability for discourse to produce mutual respect, it also potentially undermines the ability for discourse to promote the inclusion of diverse voices.

(3) Asymmetries in generalization

The third prototype effect of interest, asymmetries in generalization, refers to a similar asymmetry. Specifically, it refers to the tendency for subjects to be more likely to generalize *from* a prototypical case *to* a less prototypical case, than vice versa. For instance, in one study Rips (1975) tells subjects about a fictional island containing robins (a more prototypical bird) and ducks (a less salient example of a bird), and finds that subjects are more likely to fear the spread of disease from the robins to the ducks, than vice versa. Some social psychologists have pointed to asymmetries in generalization to help explain stereotype stability. For instance, Rothbart and John (1985) predict that the tendency to generalize from typical cases to less typical cases should lead a bias in information processing that favours stereotype stability over stereotype change.

For example, if, while grocery shopping in the Italian quarter, a person encounters two Italian-Canadians: Daniella and Stefania. Daniella is passionately complaining about her latest romantic love-affair, while Stefania is reserved and melancholic. Stefania is not behaving like a prototypical Italian (e.g., romantic, passionate), and her stereotype disconfirming behaviour (being a melancholic, reserved Italian) is less likely to be generalized to the social group; instead, Stefania becomes "functionally isolated" (or compartmentalized) from the category (Rothbart & John, 1985). By contrast, Daniella—who is behaving like a prototypical Italian—has her attributes generalized to the category, thus contributing to stereotype stability. Longitudinal findings appear to substantiate these claims (Rothbart & John, 1993). Thus, it appears that "the same processes that individuate the category member—that is, distance the member from the category—work against generalization from the individual to the category" (Rothbart, 2001, p. 55)

Asymmetries in generalization may also contribute in part to the tendency for higher-status groups to have more influence in-group discussions and in deliberations. Because certain social groups (e.g., men, wealthier people, the university educated) are more strongly associated with attributes related to decision-making (authoritative, decisive, intelligent, etc.), then there may be an underlying expectancy that ideas and decisions will flow *from* these individuals, *to* social groups that are less strongly associated with attributes related to decision-making. This might help us understand Mendelberg et al.'s (2014) finding that when women experience a negative balance of interruptions they withdraw from the discussion, but this effect does not occur when men are negatively interrupted. That is to say, even though particular woman may deviate from the non-authoritative/passive-woman stereotype, the expectation that ideas and decisions flow from men to women—which is reinforced when she is interrupted—leads her to

withdraw her voice from deliberations. By contrast, negative interruptions would not have this effect on a male discussion partner, as there is an implicit expectation that ideas and decisions should flow from him. As I discussed earlier, this potentially undermines both the democratic and epistemic functions of discourse, and further entrenches structural advantages and disadvantages. If conversational patterns shut women's voices out of discussions, decision-making will be neither inclusive nor informed by all relevant sides of the argument. And the exclusion of women and women's voices reconstitutes a long history of excluding women from politics, and of empowering men's political decision-making capacities at the expense of women.

(4) Asymmetries in probability estimations

The final prototype effect I will consider are asymmetries in probability estimations. Prototypes—being the 'clearest cases' of any given category—are the most salient examples in any given category, and people tend to use familiar or salient examples to comprehend categories (Lakoff, 1990). When making probability estimations under conditions of uncertainty, people often rely on a *representativeness heuristic* (Tversky & Kahneman, 1974). That is, their estimates of the probability of something occurring (A producing B) are driven by the degree to which A is representative of B (the degree to which A resembles B). That people see probability and similarity in exactly the same way leads to a number of serious biases,[4] but for brevity I will focus on two: the *conjunction fallacy* (Tversky & Kahneman, 1983), and the *illusion of validity* (Tversky & Kahneman, 1974).

One of the most basic rules of probability theory is that the probability of a conjunction occurring (the probability that A & B will both occur) is always less than the probabilities of its constituent parts occurring (the probability that either A or B will occur). However, when a

[4] Other biases include: insensitivity to prior probability of outcomes, insensitivity to sample size, insensitivity to predictability, misconceptions of regression, and anchoring (see Tversky & Kahneman, 1974).

conjunction is more representative or familiar than one of its constituent parts, people falsely predict that the conjunction is more likely to occur than either of its constituent parts. For instance, if subjects are told that a police officer is prone to violence, then subjects rate the likelihood that the officer was involved in the heroin trade as being *lower* than the likelihood that the officer was involved in the heroin trade *and* that he recently assaulted a suspect (despite the fact that the second condition is a conjunction, and so probabilistically less likely to be true). In this example, the assault is not causally linked to the heroin trade allegation, but "it made the combined allegation more representative of the suspect's disposition" (Tversky & Kahneman, 1983, p. 307).

As Tversky and Kahneman state "the natural assessments of representativeness and availability do not conform to the extensional logic of probability theory" (1983, p. 295). As the authors stress, the conjunction fallacy is important for the study of trial outcomes (since judges and jurors may be making judgments on the basis of partial and fallible data). But of course, this is also important for the study of deliberative processes: the conjunction fallacy represents clear threat to the epistemic function of discourse, since it leads to worse judgements; it also potentially represents a threat to the ethical function of discourse, as I will explain. Consider an instance where neighbourhood residents are deliberating whether units of a new condominium development should be reserved as subsidized housing for low-income families. If participants are told that low-income residents are prone to drug crimes, then participants might rate the likelihood that the subsidized housing units would benefit local low-income families as *lower* than the likelihood that the subsidized housing units would benefit local low-income families *and* that drug-related crime would rise (despite the fact that the second condition is a conjunction, and so probabilistically less likely to be true).

Another biased estimation is the *illusion of validity*, which refers to the "unwarranted confidence which is predicted by any good fit between the predicted outcome and the input information" (Tversky & Kahneman, 1974, p. 1126). Thus, for instance, upon reading a vignette depicting an aggressive interaction between two individuals, a perceiver may exaggerate the likelihood that one or both of the interactions partners in the vignette is Black (as opposed to any other race), because of the fit between the predicted outcome (that social interactions should be congruent with social stereotypes, e.g. that Black people are more aggressive) and the input information (an aggressive interaction).

As I have already discussed, the associations between social groups and particular concepts, and the tendency for social stereotype to benefit members of dominant social groups (e.g., stereotypes of authoritative-men, intelligent-White people, etc.) and to reinforce to subordination of lower-status groups (e.g., stereotypes of passive/emotional-women, aggressive-Black people, etc.), can undermine the epistemic and ethical functions of discourse. Both the ethical and epistemic functions are undermined if people are more likely to judge an utterance made by White males as authoritative and intelligent, and to judge an utterance made by Black females as emotional and aggressive (regardless of the content of the utterance); or to predict that an intelligent comment must have been made by a White speaker, and an aggressive comment must have been made by a Black speaker (without knowing either speaker's race).

B. Stereotyping and Affective Responses: When Emotional Reactions Undermine the Functions of Discourse

As I have mentioned, stereotypes shape emotional responses in two ways: directly, by eliciting automatic feelings (e.g., pity, contempt, etc.); and indirectly, by generating particular

expectations of category members, which—when the expectations are violated—induce negative feelings directed toward the transgressing category members. In this section I will first consider how stereotypes prompt immediate affective responses, following the theory of the universal dimensions of social cognition (Cuddy, Fiske, & Glick, 2008; Fiske, Cuddy, & Glick, 2007). I will then examine how stereotypes generate 'normal' expectations (Lakoff, 1990), and show how—when a person violates the expectations generated by stereotypes—negative feelings aimed at the norm-transgressing group member are aroused in others (Andersen & Miller, 1997; Baldwin & Blattner, 2003; Bennett, 1982). Finally, I will relate this discussion back to normative theory, showing how automatic affective responses and emotional responses to violated expectations undermine discourse's epistemic, ethical, and democratic functions.

(1) Theory of the universal dimensions of social cognition

Research suggests that there are two basic, universal dimensions of social cognition: warmth and competence. The 'warmth' attributes include traits such as fair, generous, helpful, righteous, honest, tolerant, and understanding, and the 'competence' attributes include clever, creative, competent, efficient, foresighted, intelligent, ingenious, and knowledgeable (Fiske et al., 2007). These basic dimensions account for 82 percent of the variation in perceptions of everyday social behaviours (Wojciszke, Bazinska, & Jaworski, 1998). Fiske et al. (2007) note that certain groups represent prototypes, or social stereotypes, and that these groups map onto the two-dimensional warmth-by-competence space such that there is one cluster of in-groups (that is, one grouping of prototypes that rank *highly* on both warmth and competence), and three clusters of out-groups (groupings that rank *low* on one or both of the warmth-competence dimensions). At the present-time, in the United States, prototypes that constitute societal in-groups include: housewives, Christians, Irish, middle-class, and Americans.

The other important implication of the theory of the universal dimensions of social cognition is that most feelings toward out-groups are ambiguous. That is, although traditionally psychologists and lay people have viewed out-group prejudice as a reflection of strictly negative feelings, this model shows that most feelings toward social out-groups are a mixture of positive and negative feelings: some groups are perceived as being low in competence but high in warmth (such as the elderly, or the disabled), while some groups are perceived as being high in competence but low in warmth (such as Jewish and Asian people). Only certain social groups are unambiguously viewed negatively (low warmth and low competence), such as: poor Black people, Turkish people, welfare recipients, feminists, and—most markedly—the homeless (see Figure 1).

From Fiske et al., "Universal dimensions of social cognition: Warmth and competence" (2007, p. 80)

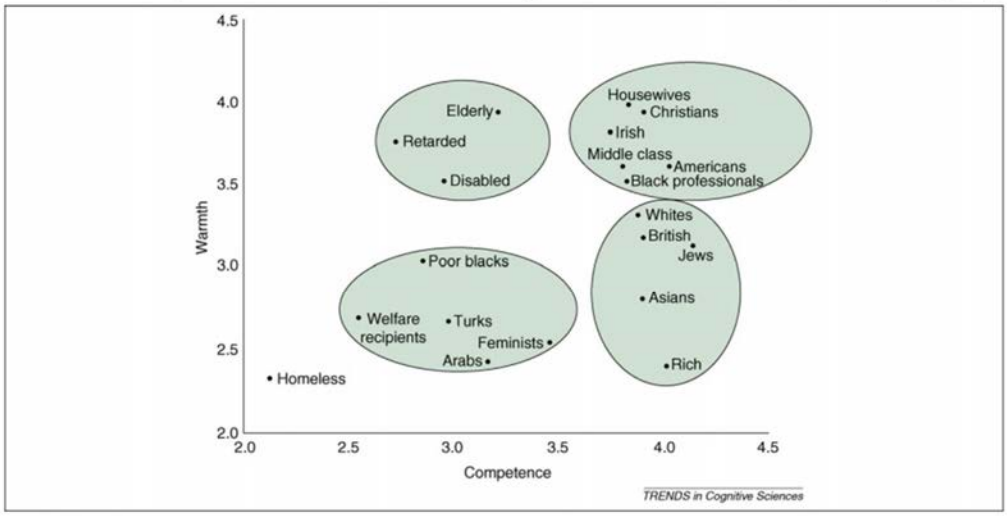

Figure 1. Scatter plot and cluster analysis of competence and warmth ratings for 20 groups. Averaging across US respondents, each group receives warmth (warm, friendly) and competence (competent, capable) scores, which are submitted to cluster analyses to determine number and membership of clusters. Groups near the center of their cluster replicate cluster membership most reliably across studies. Ratings on other variables (emotions, behaviors) cross-validate the cluster solutions. Different group names were used in different studies. Usually, an initial sample of respondents generated group names that were later rated by a second set of respondents on warmth and competence. The 20 names shown here were selected from prior sets and for various theoretical reasons. Warmth and competence were rated on five-point scales. Related to data from Refs [30,31,33,38,39,43]. Reproduced, with permission, from Ref. [31].

Each of these groupings is associated with particular automatic affective responses. For instance, groups rated as having low competence but high warmth (such as the elderly and disabled) elicit feelings of pity and sympathy (Cuddy, Fiske, & Glick, 2007; Fiske et al., 2007; Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Xu, Cuddy, & Glick, 1999), and groups rated as having high competence but low warmth (such as Jewish and Asian people) elicit feelings of envy,

jealousy, and distrust. By contrast, in-groups rated uniformly highly (such as housewives and Christians) elicit feelings of pride and admiration, and out-groups rated uniformly low across both dimensions (such as poor Black people and the homeless) elicit feelings of contempt and disgust.

From Fiske et al., "Universal dimensions of social cognition: warmth and competence" (2007, p. 81)

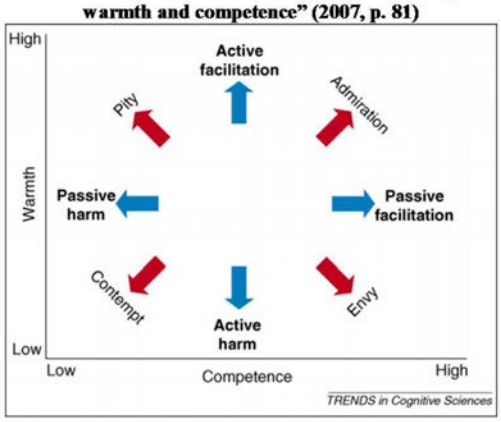

Figure 2. Schematic representation of behaviors from intergroup affect and stereotypes (BIAS) map. Competence and warmth stereotypes are represented along the *x* and *y* axes. Emotions are represented by red arrows on diagonal axes. Thus, groups in different quadrants are rated as receiving one predicted emotional prejudice and two predicted behaviors. Behavioral tendencies are represented by blue arrows on horizontal and vertical axes. Reproduced, with permission, from Ref. [31].

Not only do social cognitive groupings of prototypes elicit particular affective responses, but these affective responses generate distinct types of discrimination (the behavioural component of prejudice). The warmth dimension predicts active behaviours (either helping or harming), while the competence dimension predicts passive behaviours (passive association or neglect) (Cuddy et al., 2007; Fiske et al., 2007). Social in-groups elicit helping and associating behaviours, while low-low out-groups elicit active attacks and passive neglect. By contrast, the other two groups elicit more ambiguous behavioural responses: sympathetic groups elicit passive neglect and active helping (consider the institutionalization of the elderly in seniors' homes), while envied groups elicit passive association and active harm.

Automatic emotional reactions to social cognitive groupings of prototypes threaten to undermine the democratic function of discourse, since positive and negative feelings can play a role in determining who is included in political talk in the first place. The preference to associate with, and to help members of in-groups mean that members of these categories—which tend to be members of dominant social groups—are more likely to be included in political dialogue in

the first place. By contrast, members of out-groups, who are often neglected or harmed, are much less likely to be included in social or political discourse.

This brings me to the less obvious consequence of affective responses: the way in which feelings towards others colours how people interpret what others say. Members of social groups who are perceived as having high competence but low warmth (such as Asian or Jewish people), and who elicit passive associating behaviours might be included in discursive events, but their voices might still be discounted due to the negative feelings (e.g., envy, distrust) that they evoke. And of course, if members of low-low out-groups (e.g., poor Black or Turkish people) are included in discourse at all, it seems very likely that the negative feelings they elicit (e.g., disgust, contempt) will reduce their status as speakers, and undermine the force of their conversational contributions. This would further undermine the democratic function of discourse (since neglect is a form of exclusion), and would also undermine the epistemic function of discourse (since the weight assigned to arguments would differ depending on the ascribed status of the speaker). Automatic affective responses could also vitiate the ethical function of discourse, since social intercourse that calls up automatic feelings of disgust or contempt will have a hard time generating feelings of mutual respect or liking.

(2) Affective responses to violations of stereotypic-expectations

The way that stereotypes generate normative expectations is also important when considering how negative affect undermines the status of speakers, as norm violations elicit negative affective responses. Consider the prototype of a “housewife-mother,” which arises from a stereotypical view of the nurturance model of motherhood and femininity (Lakoff 1990, 80). The stereotype of the housewife-mother generates a prototype effect, whereby the ‘best example’

of a mother (the prototypical mother) is a biological mother, who is a housewife concerned with nurturing, married to the offspring's male progenitor, and not working at a paid position in the labour force. These stereotypes help define normal expectations (of the woman as working in the home, being nurturing, financially dependent upon the father of her children, etc.). As I have suggested, cognitive structures (stereotypes and resultant prototype effects) generate affective responses. Studies show that *violating* the expectations generated by stereotypes—for instance, when women in positions of power act authoritatively, and violate the expectation that women are nurturing—generates negative feelings aimed at the transgressor (Dion, 2008; Statham, Richardson, & Cook, 1991).

The automatic affective response to violated expectations has real consequences: for instance, women university instructors who deviate from nurturing-woman stereotypic expectations are consistently given lower Student Teaching Evaluations (Andersen & Miller, 1997; Baldwin & Blattner, 2003; Bennett, 1982). This issue is compounded by another stereotype: the stereotype of university instructors as being authoritative (a traditionally 'masculine' attribute, which contradicts the 'nurturing-female' stereotype). Female university instructors who are perceived as being *too* nurturing and not authoritative enough are more often judged as being unprofessional or unqualified (Statham et al., 1991). Clearly, these discursive inequalities create unequal capacities to influence the outcome of discourse. Female instructors have a more difficult time effectively communicating their expertise than their male colleagues, because they are faced with the difficult task of meeting the contradictory expectations of two paradoxical stereotypes (nurturing and authoritative).

As I discussed in my section on automatic affective responses, negative affect can undermine the democratic function of discourse (since related avoidance behaviours will exclude

disliked participants from the discussion), the epistemic function (since less weight may be assigned to disliked participants arguments), as well as the ethical function (since feelings like disgust are unlikely to produce positive mutual feelings such as mutual respect and trust). The findings from the universal theory of social cognition and studies of stereotype violation help explain what went wrong in 1992, when the public responded furiously to Clinton's cookies-and-tea statement. First of all, Clinton—who was not only gainfully employed, but, furthermore, was fulfilling a career that commands authority—gained the reputation as being an "ultra-feminist" (Judd & Koppel, 1992). As we know from the universal theory of social cognition, the prototypical feminist is rated as being low on both warmth and competence, and elicits unambiguously negative automatic emotional responses (such as disgust and contempt). Not only did Clinton elicit automatic negative feelings by behaving like a feminist, but furthermore—by violating the housewife-mother prototype that prospective first ladies are expected to embody—Clinton incurred additional dislike. In order to come across as a likeable, warm, and competent prospective first lady, Clinton (at least, Clinton circa 1992) had to try and approximate the housewife-mother prototype (hence the fresh-baked, oatmeal chocolate chip cookies).

IV. Conclusion: Achieving Discursive Equality

My integration of the cognitive approach to stereotyping—and the relevant empirical findings concerning what I call discursive inequalities—with a systemic approach to deliberative democracy, has a number of implications. As I have shown, discursive inequalities—social categories and the resultant prototype effects and affective responses—threaten to undermine the epistemic, ethical, and democratic functions of discourse. However, the fact that discursive inequalities present a challenge to achieving the normative goals of deliberative democratic

theory does *not* mean that the normative aims of deliberative democratic theory cannot be achieved. Rather, it means that democratic theories recognizing the central role that discourse plays in generating good and inclusive decisions, and in generating mutual respect among discussants, need to find ways to address and potentially neutralize these discursive inequalities.

I propose three potential avenues that theorists and practitioners could explore, in order to address the discursive inequalities resulting from stereotyping. These avenues are:

(1) Promoting crosscutting communication.

(2) Drawing attention to, or making stereotypes and their consequences explicit.

(3) Investigating institutional-contextual solutions to minimizing discursive inequalities in more formal deliberative settings.

The first of these—promoting crosscutting exposure—borrows heavily from the contact mechanism literature (Allport, 1954; Brown & Hewstone, 2005; Pettigrew, 1979). The basic premise of the contact mechanism is that promoting interactions between members of different groups—say, for instance, members of an in-group and members of an out-group—under certain conditions will promote more positive intergroup relations. Crosscutting communicative interactions promote more positive intergroup relations because forging interpersonal relationships (such as through discourse) and learning about disparaged out-groups through narratives and perspective-taking increases feelings of empathy, and reduces prejudicial attitudes (Galinsky & Ku, 2004; Galinsky & Moskowitz, 2000; Vescio, Sechrist, & Paolucci, 2003).

The second solution for addressing the problem of how stereotyping generates discursive inequalities is to explicitly talk *about* discursive inequalities and their effects. To borrow a phrase from Foucault (2003), this will help “reveal subjugating power’s external face.” That is to say, because most people today are (or at least can be) convinced that inequalities based on

social group membership are not justifiable, and by forcing people to realize that the language they employ, or the implicit attitudes they might hold, contribute to maintaining social inequalities, most people will try and suppress the kinds of automatic evaluations and associations that generate discursive inequalities. There appears to be empirical support for this hypothesis. For instance, Mendelberg (2001) finds that members of the electorate can be primed by implicit messages on race, but that these primes lose their effectiveness when the racialised content of the message is made explicit. This behavioural evidence is substantiated by studies using fMRI scans, which illustrate that people can engage in explicit evaluations (controlled processes) to override automatic, implicit evaluations (Cunningham, Raye, & Johnson, 2004).

Finally, the third means for minimizing discursive inequalities, and achieving *discursive equality*, borrows from the work done by Mendelberg and her colleagues (Mendelberg, Karpowitz, & Goedert, 2014; Mendelberg, Karpowitz, & Oliphant, 2014), who find that both context (i.e., the proportion of women in a deliberative setting) and institutional features (i.e., whether decisions are made by unanimous agreement or majority vote) impact conversation dynamics and deliberative authority and, ultimately, the participation of subordinate group members (e.g., women). Further studies on how institutional or contextual features shape discourse and deliberation would be useful to better understand how deliberation can be organized to produce desirable outcomes.

Despite formal democratic institutions, even Western democracies continue to be marred by inequalities embedded in cognitive and linguistic categories, that grant members of certain social group greater status as speakers, while undermining the speaker-status of members of other groups. Understanding when discursive inequalities are generated, and how they impact discourse is of fundamental importance to talk-centric democratic theories. Discourse—both

public deliberation on matters of collective concern that will result in binding decisions by the state, as well as social discourses that will only generate decisions in the purely social sense—should contribute to decisions based on thoughtful opinions informed by relevant, accurate information; to decisions that are inclusive of a diverse variety of voices, opinions, and concerns; and to the mutual respect and trust of participants in discourse. However, as my discussion shows, the structure of cognitive categories (of stereotypes) can impact cognition and affect in ways that potentially interfere with the ability for discourse to fulfil these epistemic, democratic, and ethical functions.

References

Allport, G. (1954). *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.

Andersen, K., & Miller, E. D. (1997). Gender and student evaluations of teaching. *PS: Political Science & Politics*, *30*(02), 216–219.

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or“ would Jesse Jackson’fail’the implicit association test?.” *Psychological Inquiry*, *15*(4), 257–278.

Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations: What every faculty member needs to know. *College Teaching*, *51*(1), 27–32.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, *7*(3), 136–141.

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, *74*(2), 170.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*(6), 1142.

Bottger, P. C. (1984). Expertise and air time as bases of actual and perceived influence in problem-solving groups. *Journal of Applied Psychology*, *69*(2), 214.

Bowers, W. J., Steiner, B. D., & Sandys, M. (2001). Death sentencing in black and white: An empirical analysis of the role of jurors’ race and jury racial composition. *U. Pa. J. Const. L.*, *3*, 171.

Brewer, M. B., & Gaertner, S. L. (2001). Toward reduction of prejudice: Intergroup contact and social categorization. *Blackwell Handbook of Social Psychology: Intergroup Processes*, 451–472.

Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. *Advances in Experimental Social Psychology*, *37*(37), 255–343.

Carlston, D. E. (1994). *Associated Systems Theory: A systematic approach to cognitive representations of persons*. Lawrence Erlbaum Associates, Inc.

Charles Stangor, & Mark Schaller. (1996). Stereotypes as Individual and Collective Representations. In *Stereotypes and Stereotyping* (pp. 3–40). New York: Guilford Press.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer’s dilemma: using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*(6), 1314.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2007). The influence of stereotypes on decisions to shoot. *European Journal of Social Psychology*, *37*(6), 1102–1117.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, *40*, 61–149.

Cunningham, W., Raye, C., & Johnson, M. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Cognitive Neuroscience, Journal of*, *16*(10), 1717–1729.

David L. Hamilton, & Tina K. Trolier. (1986). Stereotypes and Stereotyping: An Overview of the Cognitive Approach. In *Prejudice, discrimination, and racism: Historical trends and contemporary approaches*. (pp. 127–163). Academic Press.

Devine, P. G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5.
Dion, M. (2008). All-Knowing or All-Nurturing? Student Expectations, Gender Roles, and Practical Suggestions for Women in the Classroom*. *PS: Political Science & Politics*, *41*(04), 853–856.
Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion*, *15*(2), 115–141.
Fazio, R. H., & Dunton, B. C. (1997). Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology*, *33*(5), 451–470.
Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*(1), 297–327.
Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83.
Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878.
Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, *55*(3), 473–489.
Gaertner, S. L., & Dovidio, J. F. (2014). *Reducing intergroup bias: The common ingroup identity model*. Psychology Press.
Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology*, *4*(1), 1–26.
Galinsky, A. D., & Ku, G. (2004). The effects of perspective-taking on prejudice: The moderating role of self-evaluation. *Personality and Social Psychology Bulletin*, *30*(5), 594–604.
Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708.
Giles, H., Mulac, A., Bradac, J. J., & Johnson, P. (1987). Speech accommodation theory: The first decade and beyond. *Communication Yearbook*, *10*(1), 13–48.
Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464.
Habermas, J. (1984). *The Theory of Communicative Action: Reason and the rationalization of society*. Beacon Press.
Hastie, R., Penrod, S., & Pennington, N. (1983). *Inside the jury*. The Lawbook Exchange, Ltd.
Higgins, E. T., Bargh, J. A., & Lombardi, W. J. (1985). Nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1), 59.
Hinton, P. R. (2000). *Stereotypes, cognition and culture*. Psychology Press.
Hogg, M. A. (1985). Masculine and feminine speech in dyads and groups: A study of speech style and gender salience. *Journal of Language and Social Psychology*, *4*(2), 99–112.

Judd, J., & Koppel, T. (1992, March 26). Making Hillary Clinton an Issue [Television]. *Frontline: Nightline Transcripts*. PBS.
Karpowitz, C. F., Mendelberg, T., & Shaker, L. (2012). Gender inequality in deliberative participation. *American Political Science Review*, *106*(03), 533–547.
Kirchler, E., & Davis, J. H. (1986). The influence of member status differences and task type on group consensus and member position change. *Journal of Personality and Social Psychology*, *51*(1), 83.
Lakoff, G. (1990). *Women, fire, and dangerous things: What categories reveal about the mind*. Cambridge Univ Press.
Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, *72*(2), 275.
Maass, A., Corvino, G., & Arcuri, A. (1994). Linguistic Intergroup Bias and the Mass Media. *Revue de Psychologie Sociale*, *1*, 31–4.
MacRae Jr, D. (1993). Guidelines for policy discourse: consensual versus adversarial. *The Argumentative Turn in Policy Analysis and Planning*, 1–17.
Mansbridge, J., Bohman, J., Chambers, S., Christiano, T., Fung, A., Parkinson, J., … Warren, M. E. (2012). A systemic approach to deliberative democracy. *Deliberative Systems*, 1–26.
Mendelberg, T. (2001). *The race card: Campaign strategy, implicit messages, and the norm of equality*. Princeton University Press.
Mendelberg, T. (2002a). The deliberative citizen: Theory and evidence. *Political Decision Making, Deliberation and Participation*, *6*(1), 151–193.
Mendelberg, T. (2002b). The deliberative citizen: Theory and evidence. *Political Decision Making, Deliberation and Participation*, *6*(1), 151–193.
Mendelberg, T., Karpowitz, C. F., & Goedert, N. (2014). Does Descriptive Representation Facilitate Women's Distinctive Voice? How Gender Composition and Decision Rules Affect Deliberation. *American Journal of Political Science*, *58*(2), 291–306.
Mendelberg, T., Karpowitz, C. F., & Oliphant, J. B. (2014). Gender Inequality in Deliberation: Unpacking the Black Box of Interaction. *Perspectives on Politics*, *12*(01), 18–44.
Mendelberg, T., & Oleske, J. (2000). Race and public deliberation. *Political Communication*, *17*(2), 169–191.
Operario, D., & Fiske, S. T. (2001). Stereotypes: Content, structures, processes, and context. *Blackwell Handbook of Social Psychology: Intergroup Processes*, *1*, 22–44.
Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). Us and them: social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology*, *59*(3), 475.
Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, *5*(4), 461–476.
Portes, A. (2000). Social capital: Its origins and applications in modern sociology. *LESSER, Eric L. Knowledge and Social Capital. Boston: Butterworth-Heinemann*, 43–67.
Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon and Schuster.
Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 665–681.
Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, *7*(4), 532–547.
Rosch, E. (1983). Prototype classification and logical classification: The two systems. *New Trends in Conceptual Representation: Challenges to Piaget's Theory*, 73–86.

Rosch, E. (1999). Principles of categorization. *Concepts: Core Readings*, 189–206.
Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605.
Rothbart, M. (2001). Category dynamics and the modification of outgroup stereotypes. *Blackwell Handbook of Social Psychology: Intergroup Processes*, 45–64.
Rothbart, M., & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *Journal of Social Issues*, *41*(3), 81–104.
Rothbart, M., & John, O. P. (1993). Intergroup relations and stereotype change: A social-cognitive analysis and some longitudinal findings. *Prejudice, Politics, and the American Dilemma*, 32–59.
Rubini, M., & Kruglanski, A. W. (1997). Brief encounters ending in estrangement: Motivated language use and interpersonal rapport in the question–answer paradigm. *Journal of Personality and Social Psychology*, *72*(5), 1047.
Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition*, *17*(4), 437–465.
Sornig, K. (1989). Some remarks on linguistic strategies of persuasion. *Language, Power and Ideology. Studies in Political Discourse*, 95–115.
Statham, A., Richardson, L., & Cook, J. A. (1991). *Gender and university teaching: A negotiated difference*. SUNY Press.
Thakerar, J. N., Giles, H., & Cheshire, J. (1982). Psychological and linguistic parameters of speech accommodation theory. *Advances in the Social Psychology of Language*, 205–255.
Tversky, A., & Gati, I. (1978). Studies of similarity. *Cognition and Categorization*, *1*(1978), 79–98.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.
Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293.
Vescio, T. K., Sechrist, G. B., & Paolucci, M. P. (2003). Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *European Journal of Social Psychology*, *33*(4), 455–472.
Warren, M. (1992). Democratic Theory and Self-Transformation. *American Political Science Review*, *86*(01), 8–23.
Wittgenstein, L. (1953). Philosophical investigations. Philosophische Untersuchungen.
Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, *24*(12), 1251–1263.