RUNNING HEAD: Inhibiting linguistic bias

Could you mind your language? An investigation of communicators' ability to inhibit linguistic bias

Karen M. Douglas, Robbie M. Sutton

University of Kent

Katie Wilkin

University of Manchester

AUTHOR NOTE:

This research was supported by a grant from the Economic and Social Research Council (Reference: RES-000-22-0559). Pre-testing of materials was supported by a research grant from the British Academy (Reference: SG-37234).

Correspondence concerning this article should be addressed to Karen Douglas, Department of Psychology, University of Kent, Canterbury, CT2 7NP, United Kingdom.

E-mail: k.douglas@kent.ac.uk.

Uncorrected manuscript

Abstract

Three experiments are reported that examine communicators' ability to inhibit linguistic bias. Research has shown that communicators use more abstract language (e.g., "Jamie is affectionate" versus "Jamie kisses Rose") to describe more expected behavior. Recent research shows that this bias may be overwhelmed by goals to put a "spin" on actions or to manipulate audiences' impressions of actors. Similarly, the present experiments show that people who wish to communicate without bias may often be able to do so. Inhibition occurred when participants selected descriptions from a list of alternatives, and when they freely described *both* expected and unexpected behaviors. However, inhibition failed when participants were asked to freely describe *either* expected *or* unexpected behaviors alone.

BIOGRAPHICAL INFORMATION:

Karen Douglas completed her PhD in social psychology at the Australian National University in 2000. She is currently a Senior Lecturer in psychology at the University of Kent. She is interested in communication processes, social influence and social cognition.

Robbie Sutton completed his PhD in social psychology at the Victoria University of Wellington in 2000. He is currently a Senior Lecturer in psychology at the University of Kent. He is interested in communication processes, social cognition and social justice.

Katie Wilkin completed her BSc (Honours) at the University of Newcastle-upon-Tyne in 2002. She is currently a Research Assistant in the School of Psychological Sciences at the University of Manchester. At the time the data for this article were collected, Katie was a Research Assistant at Keele University.

Could you mind your language? An investigation of communicators' ability to inhibit linguistic bias

Language is the primary means by which we share our beliefs about people (e.g., Hamilton, Gibbons, Stroessner & Sherman, 1992; Maass, Salvi, Arcuri & Semin, 1989). In using language to share our beliefs, we can intentionally (e.g., Douglas & Sutton, 2003; Petty & Cacioppo, 1986; Wenneker, Wigboldus & Spears, 2005) or unintentionally (e.g., Franco & Maass, 1999; Ruscher, 2001) influence others' beliefs. Our linguistic choices may also implicitly convey much about our own attitudes towards the people we describe (Douglas & Sutton, 2006). Often however, we may not wish to communicate our attitudes, either because they conflict with our values (e.g., Moskowitz, Gollwitzer, Wasel & Schaal, 1999) or because there is a social price to pay for their expression (Cole, Sutton & Douglas, 2007; Mae & Carlston, 2005; Sutton, Elder & Douglas, 2006). In this article, we examine whether individuals are able to suppress the linguistic expression of their biased attitudes towards people. If language is a window to the mind (Edwards & Potter, 1993), we investigate when its users are able to draw the curtains.

Effects of goals and expectancies on language abstraction

The biases inherent in people's descriptions of others have been extensively studied under the framework of the *linguistic category model* (LCM: Semin & Fiedler, 1988, 1991; see Wigboldus & Douglas, 2007 for a review). Both the linguistic intergroup bias (LIB) and the linguistic expectancy bias (LEB) are subtle but important means by which beliefs are transmitted and perpetuated (e.g., Maass, 1999; Rubini & Semin, 1994; Wigboldus, Semin & Spears, 2000). Further, recent research suggests that as well as being sensitive to describers' beliefs, language abstraction is also affected by social goals, even if it is unlikely that communicators are aware they are using language abstraction strategically (Douglas & McGarty, 2001, 2002; Gil de Montes, Semin & Valencia, 2003; Maass, Ceccarelli & Rudin, 1996; Semin, Gil de Montes & Valencia, 2003; Rubini & Sigall, 2002; Schmid & Fiedler, 1998).

These strategic effects appear to be independent of those of expectancies, or pre-existing beliefs. Douglas and Sutton (2003) separately manipulated participants' expectancies and their communication goals, for example by instructing them to portray enemies' behaviors in the best possible light. These goals influenced language abstraction more powerfully than did interpersonal expectancies (see also Wenneker et al., 2005; Semin, Higgins, Gil de Montes, Estourget & Valencia, 2005), suggesting that communication goals may moderate the tendency for the LIB and LEB to perpetuate beliefs and stereotypes.

The difficulty of inhibiting linguistic bias

Communicators' ability to use linguistic bias to achieve communication goals coheres with contemporary social psychological approaches to communication, which stress that it is flexible and purposive (e.g., Edwards & Potter, 1993; Higgins, 1981; Giles & Coupland, 1991; Semin, 2000). From this perspective it would be surprising were language use deterministically bound to describers' privately held beliefs. Nonetheless a number of published findings suggest that communicators may be unable to inhibit linguistic bias.

In one demonstration, von Hippel, Sekaquaptewa and Vargas (1997) showed that the LIB was correlated with implicit but not explicit measures of prejudice. Similarly, Franco and Maass (1999) found the LIB to be unrelated to explicit measures of prejudice towards Jews, who in the Italian Catholic milieu of the research were normatively protected from prejudice. However, the LIB was related to explicit prejudice toward Islamic Fundamentalists who were not so protected. Given that people are able to control explicit indices of prejudice more easily than implicit indices, this suggests that the LIB comes in 'under the radar' for communicators who are motivated not to express prejudice. The apparent paradox, as noted by Douglas and Sutton (2003, p. 683), is that "participants in studies of interpersonal processes appear to use language abstraction flexibly" but participants in other studies appear "unable to control their language abstraction". This would seem to extinguish hopes that the linguistic transmission of stereotypes may be stemmed. However, the present studies attempt to restore hope and uncover a possible solution to this paradox. Specifically, we aim to identify the conditions that are likely to allow communicators to control the expression of their biased beliefs.

The present research

Published studies on the ability of communicators to recruit linguistic bias differ in several respects from those concerned with the ability to inhibit bias, making difficult the comparative interpretation of results. Research on recruitment of bias has relied on explicit instructions to portray actions in a biased light (Douglas & Sutton, 2003; Semin et al., 2005; Wenneker et al., 2005). Alternatively, researchers have elicited goals specific to the communicative context, such as the goal to be liked by the audience (e.g., Rubini & Sigall, 2002). These goals are highly salient, acute, and typically have a rather Machiavellian flavor. In contrast, research concerned with inhibition of bias has not explicitly instructed people to inhibit bias. Instead it relies on the operation of chronic cultural norms such as those prohibiting prejudice toward certain groups (e.g., Franco & Maass, 1999), and correlations between linguistic bias and other indices of prejudice (e.g., Franco & Maass, 1999; von Hippel et al., 1997). It is therefore possible that different findings have emerged not because of any inherent difficulty in inhibition as opposed to recruitment of bias, but because of differences in methodology.

In this article, we therefore investigate people's ability to inhibit linguistic bias under similar conditions to those in which they have been shown to be highly proficient at recruiting bias (Douglas & Sutton, 2003). The innovation is that we instruct participants to put aside their expectancies about individuals, and provide *un*biased descriptions. We aim to be able to establish whether people truly are less able to inhibit than to creatively recruit linguistic bias.

We start with the assumption that recruiting and inhibiting linguistic bias most likely depend on different psychological mechanisms. Douglas and Sutton (2003) pointed out that inhibition and initiation of behavior are conceptually different and may be governed by different regulatory systems (cf., Higgins, 1997; Semin et al., 2005). When addicts say that they "can't help themselves", they do not mean that they can't help themselves to another cigarette for example; they mean that in an important sense they can't *stop* themselves doing it. Something similar may be true of linguistic bias, even if the specific obstacles to inhibition are different. Some evidence for the idea that recruitment and inhibition might depend on different processes comes from recent findings that the LEB persists in experiments where communication goals are also manipulated, even if it is overwhelmed by the stronger effects of the latter (Douglas and Sutton, 2003; Wenneker et al., 2005).

One specific way in which inhibition may differ from the recruitment of bias is that by its nature it depends on a monitoring process (Douglas & Sutton, 2003). In order to be suppressed, unwanted thoughts must be detected (Macrae, Bodenhausen, Milne, & Jetten, 1994; Wegner, 1992). Likewise, it is likely that in order to eradicate bias from their language, communicators need to be able to monitor it for their bias. Although there is some evidence that recipients may be able to detect biased use of language abstraction in other people's descriptions (Douglas & Sutton, 2006), detecting bias in their own language may be quite a different matter. For example, believing Jamie to be "affectionate", a communicator may not be able to appreciate that to describe Jamie's behavior in those abstract terms may be biased.

In three experiments we tested participants' ability to inhibit linguistic bias. In Experiment 1, we asked participants to choose descriptions for depicted behaviors from lists of LCM alternatives, having manipulated expectancies with consistency information. Theoretically, bias should be relatively easy to monitor under these conditions, where participants are able to scrutinize and compare different descriptions. In Experiment 2 we asked participants to freely describe both expected and unexpected behaviors. Monitoring may be more difficult here because the close relation between bias in communicators' language and thought may make it difficult for them to detect linguistic bias. On the other hand, participants were prompted to think of *both* expected and unexpected behaviors, perhaps with the knowledge that their descriptions of these behaviors ought to differ in some way. In contrast, participants in Experiment 3 were asked to freely describe *either* expected or unexpected behaviors. Here, participants did not have the luxury of thinking about two different types of behaviors (and therefore the potential for describing behaviors differently), so inhibition may be most difficult under these conditions.

Experiment 1

We pre-tested a set of materials on 160 British undergraduate students. From the pre-test we selected six behavioral scenes and descriptions that satisfied Maass et al.'s (1989) criteria. We chose only pictures that were judged positively (> 3.5) or negatively (< 2.5). Pictures were excluded or modified whenever the descriptions on free-response indicated that the scene had been interpreted differently than was intended. We also excluded pictures that were only judged positively when performed "often" and similarly for other combinations of behavior valence and expectancy. We also excluded pictures where one forced-choice response was predominantly chosen. Three of the pictures depicted positive behaviors (recycling bottles, helping someone up who had fallen, studying in the library) and three depicted negative behaviors (spray-painting a wall, talking in the cinema, dropping litter on the ground). Examples of two behaviors (one positive and one negative) with the four LCM response alternatives, are given in the Appendix.

In Experiment 1, we first asked participants to observe six cartoon scenes and choose a description for each based on the four LCM categories (e.g., Douglas & Sutton, 2003; Maass et al., 1989, 1996). Participants were informed that the actor "often" or "rarely" behaves in the manner depicted (Douglas & Sutton, 2003, Study 2). Participants were asked either simply to choose a description for each behavior (control), or to disregard the expectancy information and choose the least biased description ("inhibit" condition).

We predicted that overall, behaviors performed "often" would be encoded more abstractly than those performed "rarely". However, if people are able to inhibit the LEB, an interaction between expectancy and instruction should emerge. Specifically, the LEB should be present in the control condition, but should disappear in the "inhibit" condition.

Method

Participants and Design

One hundred and sixty five British undergraduate psychology students took part in Experiment 1 in return for course credit (123 female and 42 male, with a median age of 19.00). The experiment was a 2 (behavior valence: positive/negative) x 2 (expectancy: often/rarely) x 2 (instruction: control/inhibit) design with repeated measures on behavior valence.

Materials and Procedure

Participants were presented with six pages, each depicting one of the positive or negative pictures. Before each picture, participants were given expectancy information. For positive/often performed behaviors, the wording was: "Please note that XX is a pleasant person; he[she] very often does this sort of thing", where XX denotes the name of the actor. The wording for positive/rarely performed behaviors was: "Please note that XX is an unpleasant person; he[she] very rarely does this sort of thing". For negative behaviors, the preambles were reversed such that the actors in the "often" condition were described as unpleasant while those in the "rarely" condition were described as being pleasant.

Participants in the control condition were then asked to select from four LCM alternatives "Which of the following descriptions of XX's behavior would you pick if you had to pick one?" In the "inhibit" condition, a sentence preceded this which said "Please now disregard the information you were given above the picture and think about describing the behavior in an unbiased way." Picture order was randomized. On completion, participants were debriefed and thanked.

Results

Participants' choices were coded according to Semin and Fiedler's (1989) guidelines, by assigning DAVs the value '1', IAVs '2', SVs '3' and ADJs '4'. Data were entered into a 2 (behavior valence: positive/negative) x 2 (expectancy: often/rarely) x 2 (instruction: control/inhibit) ANOVA with repeated measures on behavior valence. Means and standard deviations are reported in Table 1.

There was a main effect for expectancy, with "often" behaviors encoded more abstractly (M = 1.99) than "rarely" behaviors (M = 1.74), F (1, 161) = 12.99, p = .000, η^2 = .08. There was no main effect for instruction, F (1, 161) = 2.53, p = .114, η^2 = .02. As predicted, there was an interaction between expectancy and instruction, F (1, 161) = 6.30, p = .013, η^2 = .04. Planned comparisons revealed that the LEB was present in the control condition (M_{often} = 2.14 vs M_{rarely} = 1.71), F (1, 161) = 18.38, p = .000, η^2 = .10, but not in the inhibit condition, (M_{often} = 1.85 vs M_{rarely} = 1.77), F (1, 161) = .61, p = .437, η^2 = .004.

Although not relevant to our hypotheses, there was a main effect for behavior valence, with negative behaviors (M = 1.97) encoded more abstractly than positive behaviors (M = 1.76), F (1, 161) = 20.26, p = .000, η^2 = .11. There was also an interaction between behavior valence and instruction, F (1, 161) = 7.28, p = .004, η^2 = .04. Post-hoc analyses revealed a difference in abstraction between control and inhibit conditions for negative behaviors ($M_{control}$ = 2.10 vs $M_{inhibit}$ = 1.85), F (1, 161) = 6.47, p = .012, η^2 = .04, but not for positive behaviors, ($M_{control}$ = 1.75 vs $M_{inhibit}$ =1.77), F (1, 161) = .04, p = .843, η^2 = .000. [1] No other effects were significant.

Discussion

Results revealed that when expectancies were minimally induced and participants were able to scrutinize and select from a list of descriptions, they were able to inhibit the LEB. The conditions of this experiment are similar to those of Douglas and Sutton (2003, Studies 1 and 2) and together with them show that participants are able both to recruit and inhibit bias when so instructed. [2]

As we have noted however, inhibition may not always be this easy. When presented with a list of descriptions, participants are able to compare and contrast them, and may notice that abstract descriptions tend to be more evaluative, interpretive, and imply greater temporal stability (Semin & Fiedler, 1988), which may make the task of inhibition easier. However, in free-response mode, communicators do not have the luxury of evaluating their comparisons in the light of others that have been suggested to them. Further, it may be intrinsically difficult to spot biases in the descriptions they generate, which by their nature are likely to be closely related to their own cognitive representations of events (Semin & Fiedler, 1988, 1991).

Therefore, the aim of Experiments 2 and 3 was to determine if participants are still able to inhibit linguistic bias when they generate their own descriptions. If the availability of alternatives determines the ability to inhibit linguistic bias, participants should be able to inhibit linguistic bias in Experiment 2, where expectancy is manipulated within-participants. However, inhibition should be more difficult in Experiment 3 when expectancy is manipulated between-participants. On the other hand, if the inhibition of bias in free-response mode is intrinsically difficult, we would expect participants to fail to inhibit bias in both Experiments 2 and 3.

Experiment 2

Method

Participants and Design

Fifty one undergraduate students from a British university participated in exchange for sweets (34 female and 17 male, with a median age of 20.00). The experiment was a 2 (behavior valence: positive/negative) x (expectancy: often/rarely) x 2 (instruction: control/inhibit) design with repeated measures on behavior valence and expectancy.

Materials and Procedure

The materials and procedure were identical to Experiment 1, except that instead of choosing from four alternatives, participants were asked to “describe XX’s behavior in your own words”. They were prompted to begin their descriptions with “XX…” Participants were given six lines to write their descriptions. Picture order was randomized. On completion, participants were debriefed and thanked.

Results

Due to the free-response nature of the study, it was necessary to code the language abstraction participants used to describe the actors’ behaviors. Predicates were coded according to the criteria set out by Semin and Fiedler (1991). Only verbs and adjectives specifically referring to the cartoon character and his/her behavior were coded. A comprehensive set of guidelines for LCM coding is provided at the website Cratylus.org (2006), and our coding followed these guidelines with two minor qualifications. First, we did not code the use of nouns because we followed the earlier version of the LCM criteria which did not include nouns. Second, state action verbs (SAV: e.g., "amaze, surprise") were always treated as IAVs. Typically, participants' responses were one sentence long and contained a small number of words for coding. An example was: "Matthew is throwing bottles into the recycling bin." where "throwing" was the only predicate to be coded, and was coded as a DAV. A participant's value of language abstraction was obtained using the formula:

Abstraction = (DAVx1 + IAVx2 + SVx3 + ADJx4)/(DAV + IAV + SV + ADJ)

(where DAV, IAV, SV and ADJ represent the number of occurrences of each category). In addition to the primary coder, a second independent rater coded all communications and interrater reliability was high (r = .98). Disagreements were resolved after discussion.

Values of language abstraction were then entered into a 2 (behavior valence: positive/negative) x 2 (expectancy: often/rarely) x 2 (instruction: control/inhibit) mixed model ANOVA with repeated measures on behavior valence and expectancy. Means and standard deviations are presented in Table 2.

There was no main effect for expectancy, with "often" behaviors encoded no more abstractly (M = 2.30) than "rarely" behaviors (M = 2.17), F (1, 49) = 1.29, p = .262, η^2 = .03. There was no main effect for instruction ($M_{control}$ = 2.41, $M_{inhibit}$ = 2.06), F (1, 49) = 3.23, p = .077, η^2 = .06. However, there was an interaction between expectancy and instruction, F (1, 49) = 6.00, p = .018, η^2 = .11. As in Experiment 1, planned comparisons revealed that the LEB was present in the control condition (M_{often} = 2.61 vs M_{rarely} = 2.20), F (1, 49) = 6.92, p = .011, η^2 =

.12, but not in the inhibit condition, (M_{often} = 1.99 vs M_{rarely} = 2.13), F (1, 49) = 0.39, p = .537, η^2 = .008.

Although not salient to our hypotheses, there was a main effect of behavior valence, with positive behaviors (M = 2.33) described more abstractly than negative behaviors (M = 2.14), F (1, 49) = 4.66, p = .036, η^2 = .09. There were no other effects.

Discussion

Even when communicators were using their own words, they were still able to inhibit linguistic bias. However, it remains to be tested whether this ability is still present when participants are not given the opportunity to compare and contrast two different types of behaviors. Inhibiting bias in one's own language may only be possible so long as one is in a position to observe discrepancies in one's descriptions of expected and unexpected events. Next, we therefore manipulated behavioral expectancy *between-participants*.

We also included a "mislead" condition (Douglas & Sutton, 2003), to examine the abilities to inhibit and recruit linguistic bias within the same experiment. Participants were asked to write their descriptions in a way that is contrary to the original expectancy (often/rarely) information given. Specifically, participants were asked to "create the opposite impression" of the target. We expected that in this condition participants would describe unexpected behaviors as or more abstractly than expected behaviors (see Douglas & Sutton, 2003).

Experiment 3

Method

Participants and Design

Two hundred and forty three undergraduate students from a British university participated in exchange for sweets (170 female and 73 male, with a median age of 19.00). The experiment was a 2 (behavior valence: positive/negative) x 2 (expectancy: often/rarely) x 3 (instruction: control/inhibit/mislead) design with repeated measures on behavior valence.

Materials and Procedure

The methods in the control and inhibit conditions were identical to Experiment 2, except that expectancy was manipulated between-participants. In the "mislead" condition, an example of the wording for a target performing a positive/"rarely" behavior was as follows: "Remember that XX is an unpleasant person who rarely does this sort of thing. However, when you are describing the behavior, I would like you to create the opposite impression of XX." The words "pleasant"/ "unpleasant" and "often"/"rarely" were substituted for positive and negative behaviors as in previous experiments. Picture order was randomized. On completion, participants were debriefed and thanked.

Results

Coding of participants' responses was carried out as described in Experiment 2. Responses were typically one sentence long and contained a small number of words for coding. In addition to the primary coder, a second independent rater coded all communications and interrater reliability was high (r = .96). Disagreements were resolved after discussion. Values were then entered into a 2 (behavior valence: positive/negative) x 2 (expectancy: often/rarely) x 3 (instruction: control/inhibit/mislead) mixed model ANOVA with repeated measures on behavior valence. Means and standard deviations are presented in Table 3.

There was a main effect for expectancy, with "often" behaviors encoded more abstractly (M = 2.10) than "rarely" behaviors (M = 1.84), F (1, 225) = 7.10, p = .008, η^2 = .03. There was also a main effect for instruction ($M_{control}$ = 2.11, $M_{inhibit}$ = 2.02, $M_{mislead}$ = 1.78), F (2, 225) = 4.08, p = .018, η^2 = .04. There was an interaction between expectancy and instruction, F (1, 225) = 6.12, p = .003, η^2 = .05. However, planned comparisons revealed that the LEB survived in the control condition (M_{often} = 2.43 vs M_{rarely} = 1.80), F (1, 225) = 14.98, p = .000, η^2 = .06 and *also* in the inhibit condition, (M_{often} = 2.19 vs M_{rarely} = 1.85), F (1, 225) = 4.48, p = .035, η^2 = .02. In the mislead condition, the LEB was obviated, replicating Douglas and Sutton's (2003) findings, (M_{often} = 1.68 vs M_{rarely} =1.88), F (1, 225) = 1.31, p = .253, η^2 = .01. We conducted the analysis again having removed the mislead condition. The interaction between expectancy and instruction was again not significant, F (1, 225) = 1.28, p = .260, η^2 = .01, demonstrating no evidence of inhibition.

Although not relevant to our hypotheses, there was a main effect for behavior valence. Positive behaviors (M = 2.01) were described more abstractly than negative behaviors (M = 1.93), F (1, 225) = 8.71, p = .004, η^2 = .04. There was also an interaction between behavior valence and instruction, F (2, 225) = 10.34, p = .000, η^2 = .08. Post-hoc analyses revealed a difference in abstraction between control, inhibit and mislead conditions for positive behaviors ($M_{control}$ = 2.20, $M_{inhibit}$ = 2.10, $M_{mislead}$ = 1.74), F (2, 225) = 8.23, p = .000, η^2 = .07, but not for negative behaviors, ($M_{control}$ = 2.03, $M_{inhibit}$ = 1.94, $M_{mislead}$ = 1.83), F (2, 225) = 1.41, p = .247, η^2 = .01. No other effects were significant. [3]

Discussion

Experiment 3 provided no evidence of inhibition. The LEB survived in the control condition *and* when participants were instructed to communicate in an unbiased way. In contrast, participants' language abstraction was sensitive to instructions to create a misleading impression of the described person. Here, the typical LEB was no longer present. Aside from the inclusion of the "mislead" condition, the only substantial change in method from Experiment 2 was that participants were asked to generate their own descriptions to *either* expected or unexpected behaviors.

General Discussion

Our results demonstrate that communicators can sometimes inhibit the linguistic expression of their expectancies. However, the results also provide the first evidence of one of the potential boundary conditions of communicators' ability to suppress bias. Experiment 3 suggests that a necessary condition for inhibition might be that communicators are able to evaluate descriptions comparatively, as is the case in forced choice procedures (Experiment 1) and when participants provide multiple descriptions, including both descriptions of expected and unexpected behaviors (Experiment 2). Given that our procedure in Experiment 3 could be considered the closest to 'real' communication, this suggests that inhibition of linguistic bias might therefore be difficult in everyday conversation. It is also important to note that under similar conditions to those in which participants failed to inhibit bias, they were able to recruit bias readily in the service of communication goals.

These results suggest not only that inhibition is potentially more difficult than the recruitment of bias, but that the two may depend on different psychological mechanisms (cf. Macrae et al., 1994; Semin et al., 2005; Wegner, 1992). Participants could inhibit bias when the method arguably facilitated monitoring by allowing participants to scrutinize alternate descriptions and behaviors. However they could not inhibit bias when this was not possible. Further suggestion that the ability to monitor bias may be uniquely important for inhibition comes from the fact that communicators are typically able to recruit linguistic bias even when they use their own words and provide just one description (Douglas & Sutton, 2003). In sum, when describers intend to be *biased*, language abstraction may be their friend, but when their intention is to be *unbiased* it may not always be their friend and may indeed sometimes be their enemy.

However, there are reasons to be optimistic. Experiments 1 and 2 clearly show that under the right conditions, people are able to inhibit the LEB, and placing people in a situation where they are aware of both expected and unexpected behaviors of a target facilitates communicators' ability to inhibit. Generally therefore, the present results suggest that interventions intended to stem linguistic transmission of biases might be more likely to succeed if they enhance people's ability to monitor their language. It may be possible to train people to appreciate how their language may be informed by their existing preconceptions.

However, there is one important factor that limits how our current findings may be generalized to the inhibition of biased language in general. In our experiments, we dealt with participants' ability to inhibit *interpersonal* biases. For example, our participants were asked to 'put aside' their pre-existing knowledge that a person often/rarely behaves in the manner depicted (see also Douglas and Sutton, 2003 for recruitment of bias). However, previous studies investigating participants' ability to inhibit linguistic bias (e.g., Franco & Maass, 1999; von Hippel et al., 1997) have done so in *intergroup* contexts. It is important to mention this, because while it may be possible for participants to inhibit the biased expression of interpersonal biases, the expression of more deep-rooted and long-standing biases, such as stereotypes, may be more difficult to control.

There are also a number of moderators and even alternative accounts of some of our findings that remain to be investigated. For example, we have not investigated whether changes occur in individuals' privately held expectancies in response to the inhibition instruction. It is not yet clear whether inhibition works by resulting in temporary suppression of bias in communicators' mental representations, or only on suppression of bias in their public language (leaving any cognitive bias intact). We think that either way, processes such as monitoring may be important but to discriminate between these two types of suppression would be of theoretical and applied interest.

Further, it would be interesting to examine the extent to which monitoring processes implicated by the present experiments are available to conscious awareness. In the literature on mental regulation and stereotype suppression, one of the main purposes of monitoring processes is to prevent unwanted beliefs from entering awareness (Macrae et al., 1994; Wegner, 1992). For any monitoring process itself to be available to introspection would perhaps defeat the purpose. In the present studies however, participants' aim was to keep biases outside their descriptions, not outside their awareness. Being consciously aware of the bias in potential descriptions scarcely defeats their communicative purpose. However, such awareness may not be necessary and it is plausible that when suppressing bias in public communication, participants recruit the same implicit monitoring mechanisms they employ when suppressing bias in their thoughts. If it turns out that biases in descriptions are consciously perceived, it will be important to determine which consciously available aspects of abstract language provide the cues to bias (e.g., informativeness, disputability, see Semin & Fiedler, 1988).

A related line of investigation would be to consider whether the internal motivation to suppress linguistic bias might be more effective than the kind of external motivations provided in the present experiments. Research shows that participants who are internally motivated to suppress bias may be more successful than those externally motivated (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Gordijn, Hindriks, Koomen, Dijksterhuis & van Knippenberg, 2004). Therefore, participants who are already intrinsically motivated to reduce bias may be better at doing so in our experimental task. Also, it suggests that interventions that elicit internal motivations to inhibit bias, for example by priming egalitarian values, may be more successful than explicit instructions.

Finally, recent findings implicate the role of regulatory focus in language abstraction (Semin et al., 2005). Semin et al. found that relative to an avoidance-oriented *prevention* focus, a pursuit-related *promotion* focus elicited abstract descriptions. Although Semin et al. (2005) did not evaluate the effect of regulatory focus on the LEB, it is an interesting possibility that a strong prevention focus may result in the attenuation of this bias. Also, instructions intended to eliminate bias are likely to work more effectively when in keeping with communicators' prior regulatory focus. For example, prevention-focused communicators may respond better to instructions to avoid bias, whereas promotion-focused communicators may better follow instructions to strive to be fair-minded.

In sum, the present studies confirm that at times, when people are perhaps able to monitor their language, they can indeed rid it of bias and thus draw the curtains on the linguistic window to their minds. However when monitoring is not possible, people may not be able to inhibit linguistic bias even under the same conditions in which they are able to use language abstraction creatively in order to put a certain spin on their descriptions.

References

Cole, J.M., Sutton, R.M., & Douglas, K.M. (2007). From speech acts to dispositions: How impressions of persons are shaped by their descriptions of others. Manuscript in preparation.

Cratylus.org (2006). The linguistic category model (LCM). Retrieved from http://www.cratylus.org/Text/1111548454250-3815/pC/1111473983125-6408/uploadedFiles/1151434261594-8567.pdf and http://www.cratylus.org/Text/1111548454250-3815/pC/1111473983125-6408/uploadedFiles/1151434300359-0007.pdf on 25 June 2007.

Devine, P.G., Plant, E.A., Amodio, D.M., Harmon-Jones, E., & Vance, S.L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality & Social Psychology*, *82*, 835-848.

Douglas, K.M., & McGarty, C. (2001). Identifiability and self-presentation: Computer-mediated communication and intergroup interaction. *British Journal of Social Psychology*, *40*, 399-416.

Douglas, K.M., & McGarty, C. (2002). On computers and elsewhere: A model of the effects of Internet identifiability on communicative behaviour. *Group Dynamics: Theory, Research and Practice, 6,* 17-26.

Douglas, K.M., & Sutton, R. M. (2003). Effects of communication goals and expectancies on language abstraction. *Journal of Personality and Social Psychology*, *84*, 692-696.

Douglas, K.M., & Sutton, R.M. (2006). When what you say about others says something about you: Language abstraction and inferences about describers' attitudes and goals. *Journal of Experimental Social Psychology, 42,* 500-508.

Edwards, D., & Potter, J. (1993). Language and causation: A discursive action model of description and attribution. *Psychological Review*, *100*, 23-41.

Franco, F.M., & Maass, A. (1999). Intentional control over prejudice: When the choice of the measure matters. *European Journal of Social Psychology*, 29, 469-477.

Gil de Montes, L.G., Semin, G.R., & Valencia, J.F. (2003). Communication patterns in interdependent relationships. *Journal of Language and Social Psychology, 22,* 259-281.

Giles, H., & Coupland, N. (1991). *Language: Contexts and consequences.* Milton Keynes UK, Open University Press.

Gordijn, E. H.; Hindriks, I., Koomen, W., Dijksterhuis, A., & Van Knippenberg, A. (2004). Consequences of stereotype suppression and internal suppression motivation: A self-regulation approach. *Personality and Social Psychology Bulletin*, *30*, 212-224

Hamilton, D.L., Gibbons, P.A., Stroessner, S.J., & Sherman, J.W. (1992). Language, intergroup relations and stereotypes. In G.R. Semin and K. Fiedler (Eds.), *Language, interaction and social cognition* (pp.102-128). London: Sage Publications.

Higgins, E.T. (1981). The 'communication game': Implications for social cognition and persuasion. In E.T. Higgins, M.P. Zanna & C.P. Herman (Eds.), *Social cognition: The Ontario Symposium, Vol. 1* (pp.343-392). Hillsdale NJ: Erlbaum.

Higgins, E.T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280-1300.

Maass, A. (1999). Linguistic intergroup bias: Stereotype perpetuation through language. *Advances in Experimental Social Psychology*, *31*, 79-121.

Maass, A., Ceccarelli, R., & Rudin, S. (1996). Linguistic intergroup bias: Evidence for an in-group-protective motivation. *Journal of Personality and Social Psychology*, *71*, 512-526.

Maass, A., Salvi, D., Arcuri, A., & Semin, G. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*, 981-993.

Macrae, C.N., Bodenhausen, G.V., Milne, A.B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*, 808-817.

Mae, L. & Carlston, D.E. (2005). Hoist on your own petard: When prejudiced remarks are recognized and backfire on speakers. *Journal of Experimental Social Psychology, 41*, 240-255.

Moskowitz, G.B., Gollwitzer, P.M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*, 167-184.

Petty, R.E., & Cacioppo, J.T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Rubini, M., & Semin, G.R. (1994). Language use in the context of congruent and incongruent in-group behaviors. *British Journal of Social Psychology*, *33*, 355-362.

Rubini, M., & Sigall, H. (2002). Taking the edge off of disagreement: Linguistic abstractness and self-presentation to a heterogeneous audience. *European Journal of Social Psychology, 32,* 343-351.

Ruscher, J.B. (2001). *Prejudiced communication: A social psychological perspective*. New York: Guilford Press.

Schmid, J., & Fiedler, K. (1998). The backbone of closing speeches: The impact of prosecution versus defense language on judicial attributions. *Journal of Applied Social Psychology*, *28*, 1140-1172.

Semin, G.R. (2000). Agenda 2000 – Communication: Language as an implementational device for cognition. *European Journal of Social Psychology*, *30*, 595-612.

Semin, G.R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*, 558-568.

Semin, G.R., & Fiedler, K. (1989). Relocating attributional phenomena within a language-cognition interface: The case of actors' and observers' perspectives. *European Journal of Social Psychology*, *19*, 491-508.

Semin, G.R., & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. *European Review of Social Psychology*, *2*, 1-30.

Semin, G. R., Higgins, E. T., Gil de Montes, L., Estourget, Y., & Valencia, J. (2005). Linguistic signatures of regulatory focus: How abstraction fits promotion more than prevention. *Journal of Personality and Social Psychology*, *89*, 36-45.

Semin, G.R., Gil de Montes, L., & Valencia, J.F. (2003). Communication constraints on the linguistic intergroup bias. *Journal of Experimental Social Psychology, 39,* 142-148.

Sutton, R. M., Elder, T. J., & Douglas, K. M. (2006). Reactions to internal and external criticism of outgroups: Social convention in the intergroup sensitivity effect. *Personality and Social Psychology Bulletin, 32*, 563-575.

von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1997). The linguistic intergroup bias as an implicit indicator of prejudice. *Journal of Experimental Social Psychology*, *33*, 490-509.

Wegner, D.M. (1992). You can't always think what you want: Problems in the suppression of unwanted thoughts. In M.P Zanna (Ed.). *Advances in Experimental Social Psychology, 25,* (pp. 193-225). San Diego, CA: Academic Press.

Wenneker, C. P. J., Wigboldus, D. H. J., & Spears, R. (2005). Biased language use in stereotype maintenance: The role of encoding and goals. *Journal of Personality and Social Psychology*, *89*, 504-516.

Wigboldus, D.H.J., & Douglas, K.M. (2007). Language, expectancies and intergroup relations. In K. Fiedler (Ed.), *Social communication* (pp. 79-106). New York: Psychology Press.

Wigboldus, D.H.J., Semin, G.R., & Spears, R. (2000). How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology*, *78*, 5-18.

Footnotes

1 The main effect of valence and the valence x instruction interaction do not involve expectancy and therefore do not influence our primary hypotheses. These unexpected findings are nevertheless interesting, although are difficult to interpret because the patterns of results are not the same across the three experiments. Specifically, in Experiment 1, negative behaviors were described more abstractly than positive behaviors, but in Experiments 2 and 3, positive behaviors were described more abstractly than negative behaviors. Also, in Experiment 1, the valence x instruction interaction yielded a difference between the control/inhibit conditions for negative behaviors but not positive behaviors. This interaction was not significant in Experiment 2 and occurred in the opposite direction in Experiment 3. We may speculate that methodological differences (e.g., forced choice/free response, within/between-participants manipulations of variables) could be responsible for these differences although further research would be required to verify this. The important point to remember is that these unexpected findings do not affect our central hypotheses.

2 We also conducted an experiment using the pictorial materials from Douglas and Sutton (2003). Here, 84 undergraduate students (22 male and 62 female with a median age of 19.00) completed an almost identical questionnaire to Experiment 1. However, we used four cartoons instead of six and expectancy was even more minimally induced by simply informing participants that the protagonist behaved in the manner depicted either “often” or “rarely”. Results were the same as in Experiment 1.

3 Note that the lower degrees of freedom are due to some participants not completing all questions.

Table 1

Means (*and standard deviations)* for language abstraction as a function of behavior typicality, behavior valence and instruction – Experiment 1.

	Instruction	
Expectancy condition	Control	Inhibit
Often		
Positive behavior	1.97 *(0.55)*	1.83 *(0.45)*
Negative behavior	2.30 *(0.79)*	1.86 *(0.56)*
Total	2.27	1.85
Rarely		
Positive behavior	1.53 *(0.45)*	1.70 *(0.45)*
Negative behavior	1.88 *(0.55)*	1.84 *(0.46)*
Total	1.71	1.77

Table 2

Means (*and standard deviations)* for language abstraction as a function of behavior typicality, behavior valence and instruction – Experiment 2.

	Instruction	
Expectancy condition	Control	Inhibit
Often		
Positive behavior	2.64 *(0.94)*	2.10 *(0.96)*
Negative behavior	2.58 *(1.05)*	1.87 *(0.95)*
Total	2.61	1.99
Rarely		
Positive behavior	2.27 *(0.86)*	2.30 *(0.79)*
Negative behavior	2.13 *(0.94)*	1.96 *(1.06)*
Total	2.20	2.13

Table 3

Means (*and standard deviations)* for language abstraction as a function of behavior typicality, behavior valence and instruction – Experiment 3.

	Instruction		
Expectancy condition	Control	Inhibit	Mislead
Often			
Positive behavior	2.49 *(0.96)*	2.27 *(0.90)*	1.62 *(0.55)*
Negative behavior	2.36 *(1.02)*	2.10 *(1.08)*	1.74 *(0.53)*
Total	2.43	2.19	1.68
Rarely			
Positive behavior	1.90 *(0.48)*	1.92 *(0.66)*	1.85 *(0.63)*
Negative behavior	1.70 *(0.64)*	1.77 *(0.62)*	1.91 *(0.52)*
Total	1.80	1.85	1.88

Appendix

Example positive behavior

(a) Matthew is putting bottles in the bin

(b) Matthew is recycling

(c) Matthew respects the environment

(d) Matthew is responsible

Example negative behavior

(a) Daniel is spraying paint on the wall

(b) Daniel is vandalising the wall

(c) Daniel doesn't care about other people's property

(d) Daniel is destructive