UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Seeing bias in irony: How recipients infer speakers' stereotypes from their ironic remarks about social-category members

Beukeboom, C.J.; Burgers, C.

DOI
10.1177/1368430219887439

Publication date
2020

Document Version
Final published version

Published in
Group Processes & Intergroup Relations

License
CC BY-NC

Link to publication

Citation for published version (APA):
Beukeboom, C. J., & Burgers, C. (2020). Seeing bias in irony: How recipients infer speakers' stereotypes from their ironic remarks about social-category members. *Group Processes & Intergroup Relations*, *23*(7), 1085-1102. https://doi.org/10.1177/1368430219887439

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:08 Nov 2024

Check for updates

Group Processes & Intergroup Relations GPIR

Article

Seeing bias in irony: How recipients infer speakers' stereotypes from their ironic remarks about social-category members

Group Processes & Intergroup Relations
2020, Vol. 23(7) 1085–1102
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1368430219887439
journals.sagepub.com/home/gpi
SAGE

Camiel J. Beukeboom[1] and Christian Burgers[1,2]

Abstract

We study whether recipients draw inferences about speakers' stereotypic impressions from their ironic (vs. literal) remarks. Following up on the Irony Bias, we conducted two experiments in which participants were exposed (in writing or sound-recording) to only either literal remarks or irony referring to either positive or negative behaviors of an unknown social category. Results of both experiments show that participants recognize speakers' biased impressions from their pattern of irony use. When speakers made ironic remarks about category members' positive behaviors (but not about negative behaviors), participants inferred that speakers had a more negative impression of the category, and that they perceived higher essentialism of negative (vs. positive) behaviors. The impression participants perceived in speakers' biased communication pattern did not translate to participants' own reported category impression. We discuss various factors that may determine when a recognized stereotypic view in a speaker contributes to the formation and maintenance of stereotypic impressions.

Keywords

irony bias, linguistic bias, social categorization, stereotypes, verbal irony

Paper received 28 March 2019; revised version accepted 10 October 2019.

Stereotypes play a fundamental role in social judgment and interaction, particularly because they are socially shared across groups of people (Haslam, Turner, Oakes, McGarty, & Reynolds, 1997; Hogg & Reid, 2006). Communication is key in perpetuating stereotypic beliefs (Beukeboom & Burgers, 2019; Collins & Clément, 2012), as interpersonal and mass-media communication contribute to consensualization of (negative) social stereotypes (Appel & Weber, 2017; Arendt, 2013; Ramasubramanian, 2011). Biased language use in communication about categorized individuals plays an important role in this process. Variations in language features like abstraction

[1]Vrije Universiteit Amsterdam, the Netherlands
[2]University of Amsterdam, the Netherlands

Corresponding author:
Camiel J. Beukeboom, Department of Communication Science, Vrije Universiteit Amsterdam, De Boelelaan 1081, Amsterdam, 1081 HV, the Netherlands.
Email: c.j.beukeboom@vu.nl

(Maass, Salvi, Arcuri, & Semin, 1989; Wigboldus, Semin, & Spears, 2000), negations (Beukeboom, Finkenauer, & Wigboldus, 2010), and linguistic labeling (Gelman, Ware, & Kleinberg, 2010; Rhodes, Leslie, & Tworek, 2012) subtly reflect and maintain existing stereotypic beliefs.

One specific linguistic means connected with biased language is verbal irony (Burgers & Beukeboom, 2016). Verbal irony is often used in interpersonal communication, i.e., about 8% of turns in conversations between friends (Gibbs, 2000) and 7.4% of e-mails sent to friends (Whalen, Pexman, & Gill, 2009) contain irony. Ironic remarks about categorized individuals can reflect stereotypic beliefs. Research on the Irony Bias (Burgers & Beukeboom, 2016) shows that speakers particularly use irony to comment on stereotype-inconsistent (vs. stereotype-consistent) behaviors. Through irony (e.g., *what a smart professor*, after a dim comment), speakers allude to an expectancy (*professors are expected to be smart*) and signal its failure simultaneously.

Linguistic biases can contribute to stereotype formation and maintenance, when recipients infer speakers' stereotypic impressions from their biased language use (Beukeboom & Burgers, 2019). In the context of the Linguistic Intergroup Bias (LIB), various studies showed that recipients used biased variations in language abstraction to draw inferences about speakers' group identity, attitudes, and communication goals toward described targets (Assilaméhou & Testé, 2013; Douglas & Sutton, 2006). With respect to the Irony Bias, initial evidence (Burgers & Beukeboom, 2016, Study 4; Burgers, Beukeboom, Kelder, & Peeters, 2015, Study 2) suggests that recipients draw inferences from ironic remarks in line with a speaker's bias. For instance, recipients infer lower essentialism of an individual's described behavior after ironic (vs. literal) remarks, indicating recognition of a speaker's implicit message that the described behavior is unexpected and onetime (Burgers & Beukeboom, 2016). This, however, does not provide conclusive evidence that recipients also infer a speaker's group stereotypes from ironic remarks.

We conducted two experiments to test whether participants being exposed to biased ironic and literal remarks about unknown social groups infer speakers' stereotypic expectancies (stereotype valence and content) and perceived essentialism of described behaviors. Additionally, we explored whether recipients self-adopt recognized stereotypic impressions in their own impressions of described targets.

Language Use in Stereotype Formation and Maintenance

Social-category stereotypes are generalized impressions consisting of features and characteristics associated with social categories. Depending on impression strength, these characteristics are expected to apply across individual category members and to be stable across situations (Beike & Sherman, 1994; Beukeboom & Burgers, 2019), although activated stereotypic information may be context dependent (Garcia-Marques et al., 2006).

To understand how stereotypes evolve, and become shared knowledge, it is crucial to focus on language use in communication about socially categorized people (Gorham, 2006; Maass, 1999; Wigboldus & Douglas, 2007). Language reflects which categories are singled out as targets for stereotyping, and is the chief carrier of stereotypic information about socially categorized individuals. In often subtle ways, language reflects, constructs, and maintains social-category beliefs (Collins & Clément, 2012). One way in which language can convey stereotypes is through differences in *how* information about categorized individuals is formulated (Beukeboom & Burgers, 2019). By means of several subtle linguistic variations, speakers communicate what is expected rather than unexpected for a categorized target or category as a whole.

Research on the LIB (Maass et al., 1989) and Linguistic Expectancy Bias (Wigboldus et al., 2000), for example, shows that speakers use more concrete than abstract language when describing stereotype-inconsistent behavior, but more abstract language for stereotype-consistent behavior.

Similarly, the Stereotypic Explanatory Bias proposes that descriptions of stereotype-inconsistent behavior contain more explanations aimed at clarifying the unexpected behavior than descriptions of stereotype-consistent behavior (Hastie, 1984; Sekaquaptewa, Espinoza, Thompson, Vargas, & von Hippel, 2003). Research on the Negation Bias (NB; Beukeboom et al., 2010, Beukeboom et al., 2019) shows that people are likelier to use negations (vs. affirmations) to communicate stereotype-inconsistent information. These biases have in common that they frame stereotype-inconsistent behaviors as exceptions, while re-affirming existing stereotypic associations with a target category (Beukeboom & Burgers, 2019).

The significance of linguistic biases in descriptions of categorized individuals thus lies in their implicit communication of stereotypes to recipients. Language used in stereotype-inconsistent messages (i.e., concrete situational, explanatory, negations, irony) induces recipients to infer lower essentialism for the described behavior (Beukeboom, 2014). This means that the behavior is seen as unexpected, an exception to the rule, and likelier caused by transient situational circumstances. In contrast, language used in stereotype-consistent messages (i.e., abstract, unmarked, literal affirmations) induces recipients to infer that the behavior is expected and likelier caused by the actor's enduring dispositional characteristics than by situational circumstances.

Moreover, in the context of the LIB, various studies showed that recipients use biased variations in language abstraction to draw inferences about speakers' group identity, and relationship, attitude, and communication goals toward targets (Assilaméhou & Testé, 2013; Douglas & Sutton, 2006). Thus, from subtle variations in speakers' (biased) language use, recipients (implicitly) draw inferences about intergroup relations and, although this received less attention, speakers' stereotypic impressions about described target categories. Understanding these recipient inferences is important because these are a requirement for interpersonal transmission and maintenance of stereotypic expectancies to occur.

Recipient Inferences and the Irony Bias

The Irony Bias (Burgers & Beukeboom, 2016) posits that irony production in stereotype-inconsistent events and subsequent recipient inferences from ironic remarks together result in stereotype maintenance. Ironic remarks contain "a literal evaluation that is implicitly contrary to its intended evaluation" (Burgers & Beukeboom, 2016, p. 415). This definition implies that irony is a pragmatic phenomenon, and that context knowledge is required to determine whether a specific comment is ironic (or not). For instance, a comment like "Well done!" is used literally when referring to an excellent performance, but used ironically when referring to an atrocious performance. Ironic comments may be perceived in different ways by addressees, such as communicating humor (Matthews, Hancock, & Dunham, 2006) or aggression (Filik et al., 2016). Irony used in particularly "bitter and derisive statement[s]" is known as sarcasm (Kreuz, Roberts, Johnson, & Bertus, 1996, p. 87), which is a particular (highly negative) sub-type of irony (Attardo, 2000; Gibbs, 2000; Whalen et al., 2009).

Previous work on the Irony Bias mostly focused on irony production and showed that speakers' stereotypic expectancies about a target surface in a biased use of ironic versus literal remarks about the target's behavior. In three studies, Burgers and Beukeboom (2016) showed that speakers find the use of ironic remarks particularly appropriate to comment on stereotype-violating (vs. stereotype-confirming) behaviors. This fits with ideas about the pragmatic function of irony. An ironic remark about stereotype-inconsistent behavior (e.g., "What a neat person" about someone's messy room) allows speakers to introduce opposite valence to the discourse and thereby implicitly allude to (Kihara, 2005) or "echo" (Wilson & Sperber, 2004) existing expectancies. Through irony, speakers could thus note a stereotypic expectancy (*the person is expected to be neat*) and signal its failure simultaneously. When processing ironic remarks, recipients need to infer that the speaker (ironically) refers to some

relevant information, like implicit stereotypic expectancies (Wilson & Sperber, 2004).

If the Irony Bias indeed functions to implicitly communicate expectancies to message recipients, then recipients should be able to infer speakers' stereotypic impressions from their ironic remarks. Inferences about speakers' stereotypic impressions can include impression valence (i.e., the speaker's general positive or negative attitude) and content (e.g., does the speaker think of the target as warm, or competent). Second, recipients can draw inferences about a speaker's perceptions of essentialism of described behaviors, and finally about perceived strength of category impressions.

A few studies provide initial support for the hypothesis that recipients draw inferences from variations in irony use. One study focused on irony reception in an intergroup setting with rival sports teams (Burgers et al., 2015, Study 2). This study showed that recipients draw inferences about an unknown commenter's position toward a category of individuals based on their usage of ironic remarks. Neutral participants were presented with ironic and literal comments about players of one of two well-known rival soccer teams. Comments referred to both competent and incompetent actions of various players. In the literal condition, the speaker commented literally on all behaviors. In the ironic-about-competent condition, the speaker commented ironically on competent behavior (e.g., "Wow, what a poor player"), and literally on incompetent behavior. In the ironic-about-incompetent condition, this was reversed (e.g., "Wow, what a great player" on an incompetent player action).

Participants inferred that speakers using irony to comment on players' competent (vs. incompetent) behavior were likelier to support this player's rival club, and had relatively negative communicative goals (i.e., creating a negative impression of the target's team). In contrast, speakers using irony to comment on players' incompetent (vs. competent) behavior were judged to be likelier a fan of the player's club, and wanted to create a positive impression. Thus, in this intergroup context, recipients relied on a speaker's irony usage to detect the speaker's group membership and position toward discussed targets. As this study focused on inferences relevant to this rival context, it did not measure perceptions of a speaker's stereotype impression valence and content.

Nevertheless, based on these results and other research on the Irony Bias (Burgers & Beukeboom, 2016), we expect that, even when a speaker talks about a new and unknown social category, recipients can pick up on a speaker's stereotypic expectancies on the basis of his/her pattern of ironic remarks. For instance, when speakers highlight incompetence through irony (e.g., "Gee, what bad work" about an excellent performance), recipients should infer that the speaker holds a negative impression and expected the target to behave incompetently. Thus, we hypothesize that:

> *H1*: Recipients infer that (a) a speaker who makes ironic (vs. literal) remarks about positive behaviors of category members has a more negative impression of the social category, while (b) a speaker who makes ironic (vs. literal) remarks about negative behaviors of category members has a more positive impression of the social category.

Second, recipients could draw inferences about a speaker's perceived essentialism. Essentialism (Medin & Ortony, 1989) is a fundamental variable in social-category perception. It refers to the extent in which associated behaviors and characteristics are perceived to be immutable to category members, and stable across time and situations, rather than due to transient situational circumstances (Beukeboom & Burgers, 2019). In line with the Irony Bias, we expect that ironic remarks communicate low levels of essentialism of the behavior, and that recipients draw inferences about the speaker's perceived essentialism.

Burgers and Beukeboom (2016, Study 4) provide preliminary support for this hypothesis. They tested recipient inferences about individual targets using brief written scenarios in which a friend made an ironic or literal evaluative comment about an individual's positive or negative

behavior. This showed that from ironic (vs. literal) comments about positive behavior, recipients infer that the behavior is atypical for the individual target, and not informative about the target's stable dispositional characteristics (i.e., lower essentialism). Even though this study focused on inferences with respect to an individual target, we expect a similar pattern for recipient inferences about a speaker's stereotypes about a generic group. This translates to:

> *H2*: After being exposed to (a) a speaker's ironic (vs. literal) remarks about positive behaviors of category members, recipients infer that the speaker expects lower essentialism of positive behaviors, while after being exposed to (b) a speaker's ironic (vs. literal) remarks about negative behaviors, recipients infer that the speaker expects lower essentialism of negative behaviors. Only in Study 2 did we measure perceived essentialism in speakers about positive and negative behaviors. In Study 1, we measured participants' own perceived essentialism (see H3).

Next to the valence and content of the stereotypic category impression and perceived essentialism of behaviors, we explored potential effects on perceived strength of category impressions. We expected that recipients from a biased communication pattern (in this case using ironic remarks) infer that the speaker has a strong category impression, which is reflected in the inference that speakers have a high perceived entitativity (uniformity) and deductive potential[1] (informativeness) of category membership.

While (implicitly) perceiving a stereotypic impression in speakers' comments is a prerequisite for stereotype maintenance, it does not necessarily imply that the recipient will immediately endorse it. Recipients could self-adopt (i.e., internalize) the communicated impression of the target group in their own category representation, but could also be indifferent to, or reject the speaker's biased beliefs. Results of Burgers et al. (2015, Study 2) suggest that a biased use of irony can also induce recipients to self-adopt the view of the speaker. That is, the communication pattern of ironic and literal comments by an implicitly biased stranger not only influenced recipients' view of this unknown commenter, but also translated to recipients' own reported attitude toward the soccer clubs.

To study whether and when perceived bias in the ironic remarks of others can induce stereotype formation in recipients, we explicitly distinguished recipient inferences about others' impression (H1 and H2) from recipients own impression and own perceived essentialism. We expect that: *H3*: After being exposed to (a) ironic (vs. literal) remarks about positive behaviors of category members, recipients themselves report to have a more negative impression of the category and lower essentialism of positive behaviors, while after being exposed to (b) ironic (vs. literal) remarks about negative behaviors, recipients report to have a more positive impression and lower essentialism of negative behaviors.

The Present Research

We conducted two experiments, both employing a 3 (Communication pattern: all literal vs. ironic about positive behavior vs. ironic about negative behavior) between-participants design. For our experiments, we chose a realistic labor context, but with hypothetical groups to prevent existing group perceptions from interfering. Participants imagined being a new employee in a company and—based on anecdotes told by their new colleagues—tried to form an impression about a distinct group of employees they did not belong to, labeled as "Ruysdaelers" (Experiment 1) or "Brinkers" (Experiment 2).[2] Before reporting their impressions, participants were presented with eight anecdotes about Ruysdaelers or Brinkers containing concrete behaviors, of which four were positive and four were negative in varying alternate order. In all conditions, participants were thus exposed to similar information with the same amount of positive and negative behaviors of target-group members. We related presented behaviors to warmth (experiment 1) and competence (experiment 2), which are considered to be

universal dimensions of stereotype content (Cuddy, Fiske, & Glick, 2008).

The main manipulation is operationalized in the remarks that are made by colleagues about the discussed positive and negative behaviors of target group members. Depending on communication-pattern condition, the eight discussed behaviors of the target group were followed by either a literal or an ironic remark: In the "all literal" condition, all eight remarks were literal (i.e., had equal valence as the behavior to which they referred); in the "ironic about positive behavior" condition, remarks made about the four positive behaviors were ironic (i.e., had opposite valence), while remarks about the four negative behaviors were literal (i.e., had equal valence). In the "ironic about negative behavior" condition, this was reversed. Thus, in line with the common definition of irony, whether a remark is literal or ironic depends on context (Attardo, 2000; Wallace, 2015). When the comment "Gee, he is really smart" refers to a clever behavior, it is literal and reflects the propositional meaning. When it refers to a dim behavior, it is ironic.

To control for potential effects of material order and content, we created eight materials sets for each of the two experiments. Each set consisted of (1) a description of a behavioral event which ended with (1) a concrete behavior description (one positive and one negative) in which valence was varied by means of verbs (e.g., "ignore" vs. "help"), adverbs (e.g., projects progressed "pretty badly" vs. "really well"), or sentence object (e.g., "organized nothing" vs. "organized a party") and (2) a positive and negative evaluative remark about the behavior (varying only in use of one of two antonym traits; e.g., "boring" vs. "sociable"), made by another colleague. These were counterbalanced across participants. Participants were shown one behavior + remark combination from each set (sets 1 through 8) in consecutive order. Participants assigned to Behavior sequence 1 received positive behaviors from the uneven sets (i.e., 1, 3, 5, 7) and negative behaviors from the even sets (i.e., 2, 4, 6, 8). For participants in behavior sequence 2, this was reversed.

To illustrate, when the positive behavior of set 1 (i.e., "helping") was presented to a participant (i.e., in sequence 1), it was followed by an ironic remark in the "ironic about positive behavior" condition (i.e., "Gee, that is a mean Ruysdaeler"; opposite valence to behavior), while it was followed by a literal remark (i.e., "Gee, that is a nice Ruysdaeler"; same valence as behavior) in the "ironic about negative behavior" and the "all literal" conditions. When the negative behavior (i.e., "ignoring") of this set was presented (i.e., in sequence 2), this was reversed (see online Appendices A, B, C, and D for a design overview and materials; https://osf.io/9g72z). This design allows presenting participants with ironic vs. literal remarks while keeping the content of the presented behaviors, as well as the content of the remarks, the same across conditions.

Finally, to be able to test whether effects generalized across modalities, we replicated the study using different modes of presentation. Experiment 1 was a self-paced reading experiment, in which participants read the anecdotes about the target group. In Experiment 2, participants listened to a recording of a lunch conversation with three colleagues talking about target-group members.

After participants received all eight anecdotes, we measured the dependent variables: perceived impression of the target group in their colleagues, perceived essentialism about positive and negative behaviors (only in Experiment 2 as perceived in the speakers), and participants' own impression of the target group on warmth and competence (Cuddy et al., 2008).

Experiment 1

Method

Participants. To determine required power, we looked for a comparable study; Burgers et al. (2015, Study 2) asked participants in an intergroup setting to infer speaker goals after exposure to different communication patterns with irony. They reported statistically large effects on speaker inferences.[3] A power analysis in G*Power

3.1.9.4 (Faul, Erdfelder, Buchner, & Lang, 2009) with power set at 0.90 and alpha at 0.05 revealed that we needed at least 84 participants to be able to find a similarly large effect. We aimed to recruit about 150 voluntary participants using snowball sampling via different (social) media (e.g., email, *Facebook, LinkedIn*). This resulted in a convenience sample of 177 native Dutch participants who completed the online experimental questionnaire; 55 male (31%), 122 female (69%), M_{age} 35.9 years (SD = 13.8). Data analyses were conducted only after completion of data collection.

Materials. We created eight material sets, each consisting of a brief scenario about how one colleague talked about a particular Ruysdaeler in a behavioral event which ended with (1) a concrete behavior description (either positive or negative valence) on the warmth dimension, and (2) a positive or negative evaluative remark about the behavior (using one of two antonym traits) made by another colleague. For instance, set 1 described how a colleague talked about the Ruysdaelers having a busy time in their department. A deadline was approaching, so all Ruysdaelers had to work hard. One Ruysdaeler was not very fast. When another Ruysdaeler noted this, he either helped (positive behavior) or ignored him (negative behavior), depending on behavior-sequence condition. Depending on Communication pattern condition, another colleague reacts to this story by saying either: "Gee, that is a nice Ruysdaeler" (positive remark) or "Gee, that is a mean Ruysdaeler" (negative remark). The same positive remarks were literal when referring to positive behavior (i.e., same valence as behavior) and ironic when referring to negative behavior (opposite valence to behavior).

Procedure. Participants read that the study dealt with communication within companies. Participants had to imagine being a new employee in a company housed in several buildings. They were told that another team worked in the Ruysdael Building, and that its team members were known as Ruysdaelers. Participants were told that they would read transcripts of lunch conversations in which colleagues from their own team discussed and commented upon Ruysdaelers.

Participants were then instructed to take their time to carefully read all materials, and to try their best to form an impression about the characteristics of Ruysdaelers. They were subsequently presented with eight behavior descriptions on the warmth dimension, followed by either a literal or an ironic remark, depending on communication pattern condition, as described above. Next, dependent variables were measured.

Dependent variables

Participants' own impression. First, participants rated their own impression of Ruysdaelers, operationalized into three variables. Four items were combined (Cronbach's α = .87) to measure *own impression of warmth*, using three Likert scales ranging from 1 = cold, unfriendly, unhelpful, to 7 = warm, friendly, helpful; and the item "To what extent do you think the Ruysdaelers are cold or warm people?" (1 = cold people, 7 = warm people). Four similar items measured *own impression of competence*, 1 = unintelligent, unprofessional, incapable, to 7 = intelligent, professional, capable; and 1 = incompetent people, 7 = competent people (Cronbach's α = .87). One item measured general *own attitude about Ruysdaelers*: To what extent is your impression of the Ruysdaelers negative or positive? (1 = negative, 7 = positive). Two items related to the perceived category strength. These did not combine into a reliable scale ($\alpha < .47$) and were analyzed as separate variables: *Own category uniformity* was measured with the item "To what extent do you think Ruysdaelers are all alike?" (1 = are not at all alike, 7 = are very much alike) and *Own category informativeness* with the item "Based on this information, can you say little or a lot about these people?" (1 = very little, 7 = a lot).[4]

Perceived impression in speakers. Next, participants were asked for their thoughts about the impression held by their colleagues discussing Ruysdaelers. This was operationalized into the variables:

Speakers' attitude about Ruysdaelers with the item "To what extent is the impression of your colleagues about Ruysdaelers negative or positive?" (1 = negative to 7 = positive); *Speakers' impression of competence* using the item "To what extent do your colleagues think that Ruysdaelers are competent people?" (1 = incompetent people, 7 = competent people); *Speakers' impression of warmth* using the item "To what extent do your colleagues think that Ruysdaelers are cold or warm people?" (1 = cold people, 7 = warm people). Finally, category strength as conveyed by colleagues was measured using *Speakers' uniformity*: "To what extent do your colleagues think that Ruysdaelers are all alike?" (1 = are not at all alike, 7 = are very much alike), and *Speakers' informativeness*: "Suppose someone only knows about a person that he/she is a Ruysdaeler. Do your colleagues think that, based on this information, they can say little or a lot about this person?" (1 = very little, 7 = a lot).

Participants' own perceived essentialism about positive and negative behaviors. Next, participants were again shown the same eight specific behaviors they were presented with before, but this time without the colleague's ironic/literal remarks. Each behavior was followed by five items measuring participants' own perceived essentialism of the particular behavior (Burgers & Beukeboom, 2016; Wigboldus et al., 2000). For instance, for set 1, instructions read: "The Ruysdaelers are working very hard to make a deadline. One Ruysdaeler was not very fast. *When another Ruysdaeler noted this, he helped him.* Please answer the following questions about the behavior of this Ruysdaeler (note: the questions refer to the italicised behavior)."

Essentialism was operationalized in the following five variables, each measured with one item: *Expectedness* (To what extent is the described behavior of this Ruysdaeler expected for a Ruysdaeler? 1 = not at all, 7 = very much); *Repetition likelihood* (How large do you estimate the chance that this Ruysdaeler will repeat the demonstrated behavior in the future? Please note a number between 0 and 100%); *Attribution* (To what extent is the described behavior of this Ruysdaeler due to the situation or the person? 1 = completely due to the situation, 7 = completely due to the person); *Typicality* (To what extent is the described behavior of this Ruysdaeler typical for Ruysdaelers? 1 = not at all typical, 7 = very typical); *Generalizeability* (How likely is it that this Ruysdaeler will repeat the same behavior in other situations. 1 = very unlikely, 7 = very likely). These did not combine into a reliable scale (Cronbach's α for participants' own perceived essentialism of positive behavior = .41, and of negative behaviors = .68) and were therefore analyzed as separate variables. No further variables were measured.

Results

We first conducted the analyses reported below including the Behavior sequences factor that controlled for the presentation order of materials. As we observed no significant main or interaction effects of presentation order (most Fs < 1), we further excluded this factor in the analyses.

Perceived impression in speakers

H1 stated that recipients of ironic remarks about positive behaviors of category members would infer that the speaker has a more negative category impression, while recipients of ironic remarks about negative behaviors would infer that the speaker has a more positive impression. A MANOVA with Communication pattern (all literal, ironic about positive behavior, ironic about negative behavior) as our independent variable, and Speakers' attitude about Ruysdaelers, Speakers' impression of competence, Speakers' impression of warmth, Speakers' uniformity, and Speakers' informativeness as dependent variables showed a significant multivariate effect, Wilks' $\Lambda = 0.68$, $F(10,340) = 7.36$, $p < .001$, $\eta_p^2 = 0.18$. This suggests that recipients draw inferences about a speakers' category impressions from their pattern of irony use in remarks about the target category of Ruysdaelers.

Follow-up univariate ANOVAs revealed, in line with H1a (but not H1b), a main effect of

Table 1. Means (and standard deviations) of dependent variables (separated for perceived impression in speakers vs. participants' own) in the three communication pattern conditions (Experiment 1).

Dependent variable	Communication pattern condition		
	All literal (n = 56)	Ironic about positive behavior (n = 54)	Ironic about negative behavior (n = 67)
Perceived impression in speakers			
Speakers' attitude about Ruysdaelers	3.59_a (1.16)	1.81_b (1.03)	3.57_a (1.65)
Speakers' impression of warmth	3.75_a (1.05)	2.02_b (1.21)	3.54_a (1.61)
Speakers' impression of competence	3.98_a (1.14)	2.35_b (1.32)	3.84_a (1.51)
Speakers' uniformity	4.04_a (1.79)	5.20_b (1.74)	4.60_{ab} (1.80)
Speakers' informativeness	4.25_a (1.68)	5.20_b (1.50)	4.51_a (1.53)
Participants' own impression			
Own attitude about Ruysdaelers	4.09_a (1.01)	4.63_b (1.12)	3.81_a (1.20)
Own impression of warmth	4.15_a (0.62)	4.60_b (1.06)	3.94_a (1.02)
Own impression of competence	4.37_a (0.67)	4.65_a (0.81)	4.42_a (0.89)

Note. N = 177. Means in rows with a different subscript (a, b) differ significantly according to Tukey posthoc tests ($p < .05$).

Communication pattern condition on *speakers' attitude about Ruysdaelers,* $F(2,174) = 32.70$, $p < .001$, $\eta_p^2 = 0.27$. This showed that participants in the "ironic about positive behavior" condition perceived their colleagues to convey a more negative attitude about Ruysdaelers, compared to participants in both the "all literal" and "ironic about negative behavior" condition (see Table 1). The same pattern was observed with respect to the more specific impression measures *speakers' impression of warmth,* $F(2,174) = 28.05$, $p < .001$, $\eta_p^2 = 0.24$, and *speakers' impression of competence,* $F(2,174) = 25.14$, $p < .001$, $\eta_p^2 = 0.22$. Since colleagues' remarks only referred to warmth, this latter effect on colleagues' impression of competence was unexpected. It appears that the communication patterns resulted in a general halo effect in evaluating the target on both warmth and competence. Results demonstrate that participants (implicitly) recognized the biased pattern in speaker comments, particularly in the "ironic about positive behavior" condition. The "ironic about negative behavior" condition did not differ from the "all literal" control condition.

Interestingly, our measures for perceived strength of category impression (i.e., uniformity and informativeness) show that when colleagues used a biased pattern of ironic remarks (again, particularly the "ironic about positive behavior" condition, see Table 1), participants inferred that speakers think that Ruysdaelers are all alike (i.e., higher uniformity), $F(2,174) = 5.93$, $p = .003$, $\eta_p^2 = 0.06$, and that being a category member is more informative, $F(2,174) = 5.46$, $p = .005$, $\eta_p^2 = 0.06$, compared to the "all literal" condition.

Participant's Own Impression and Perceived Essentialism of Positive and Negative Behaviors

H3 stated that recipients themselves would form a biased impression after being exposed to speakers' biased ironic (vs. literal) remarks about category members. However, the variables measuring participants' *own* impression show an unexpected reversed effect as predicted in H3. A MANOVA with Communication pattern as independent variable, and participants' own attitude about Ruysdaelers, own impression of competence, own impression of warmth, own uniformity, and own informativeness as dependent variables showed a significant multivariate effect, Wilks' $\Lambda = 0.88$, $F(10,340) = 2.31$, $p = .01$, $\eta_p^2 = 0.06$.

Contrary to what participants perceive in their colleagues, participants themselves reported a more *positive* impression after they heard ironic remarks about positive behaviors, compared to both the "all literal" and "ironic about negative behavior" conditions. This pattern was observed for both participants' *own attitude about Ruysdaelers,* $F(2,174) = 8.22$, $p < .001$, $\eta_p^2 = 0.09$, and their *own impression of warmth* $F(2,174) = 7.78$, $p = .001$, $\eta_p^2 = 0.08$. No effects were observed for participants' *own impression of competence* $F(2,174) = 1.90$, $p = .15$. This appears to be reactance to ironic remarks that are perceived to convey a prejudiced impression. We observed no effects on *Own uniformity* ($F(2,174) = 1.60$, $p = .21$) and *Own informativeness* ($F(2,174) = 2.16$, $p = .12$).

In Experiment 1, we measured participants' own perceived essentialism of positive and negative behaviors. H3 stated that after being exposed to ironic remarks about positive behaviors recipients themselves would have a lower perceived essentialism of positive behaviors, while after ironic remarks about negative behaviors they would have a lower perceived essentialism of negative behaviors. Because the five essentialism items did not form a reliable scale, we conducted a 3 (Communication Pattern condition) X 2 (Behavioral valence: positive, negative) MANOVA with the five separate essentialism items as dependent variables, and with repeated measures on behavioral valence. We did not observe multivariate effects of communication pattern condition on participants' own perceived essentialism (communication pattern main effect Wilks' $\Lambda = 0.94$, $F(10,340) = 1.07$, $p = .38$, $\eta_p^2 = 0.03$; interaction Wilks' $\Lambda = 0.91$, $F(10,340) = 1.73$, $p = .07$, $\eta_p^2 = .05$. Univariate repeated measures ANOVAs suggest a comparable reversed effect as above for participants' own impression (i.e., participants disagreeing with the pattern conveyed by speakers), but only on two (about negative behaviors) out of 10 measures (see Appendix E).

Discussion

In line with H1a, we found that speakers making ironic remarks about category members' positive behaviors were perceived to hold a more negative impression about the category (on both the general attitude and specific impression measures about warmth and competence) compared to the "all literal" control condition. Ironic remarks about positive behavior (e.g., "Well, that is a rude Ruysdaeler!," to comment on friendly behavior) introduce a negative evaluation and thereby imply a negative expectancy. Contrary to H1b, speakers making ironic remarks about negative behaviors were not perceived to convey a more positive impression compared to the "all literal" control condition. Note that introducing a positive evaluation ("Well, that is a nice Ruysdaeler!," to comment on unfriendly behavior) is more commonly used for politeness reasons and to convey a general positivity norm (Brown & Levinson, 1987; Dews & Winner, 1995). Such remarks are therefore less informative about target expectancies.

In addition to impression valence, we explored potential effects on category perceptions. Participants inferred that speakers making biased ironic remarks (again in the "ironic about positive behavior" condition) thought that Ruysdaelers are all alike (i.e., higher uniformity) and that being a category member is more informative.

These effects were not observed in participants' own impression of the category and perceived essentialism of positive and negative behaviors. Apparently, receiving a number of biased ironic remarks about an unknown category of individuals was not enough for participants to take over and internalize the biased impression they inferred from the speakers' remarks. In fact, we even observed some reversed effects, which may suggest participants rejected the prejudiced impression they perceived in the ironic remarks of their colleagues.

Experiment 2

In Experiment 2, we aimed to replicate the findings of Experiment 1 and test whether they generalize to different materials and modality of presentation. Participants listened to spoken (rather than reading written) text, and were

presented with anecdotes about a different target group (Brinkers rather than Ruysdaelers) showing competence-related (rather than warmth-related) behaviors. Moreover, in Experiment 2, described behaviors and (ironic/literal) remarks were formulated in a more generic manner. In Experiment 1, both the behaviors and remarks referred to individual target group members (e.g., This Ruysdaeler is . . .). In Experiment 2, both the behaviors and the (ironic/literal) remarks referred to the target group in generic terms (e.g., Brinkers are . . .). Based on research on generics (Cimpian & Markman, 2008; Gelman et al., 2010; Rhodes et al., 2012), we expected that remarks using generic labels would have a stronger effect in conveying generalized category impressions. In order to test participants' inferences about essentialism of behaviors, the essentialism measure was now aimed at what participants perceive in speakers rather than their own perceptions.

Method

Participants. For this online experiment, we again aimed to recruit a sample of about 150 native Dutch voluntary participants using snowball sampling via a link posted on various social media. However, a number of the 189 respondents who started the study did not hear the audio fragment due to technical issues. Respondents who did not complete the questionnaire, and/or replied to our control questions that they could not hear the audio fragment well, and/or did not listen uninterrupted were excluded. This resulted in a convenience sample of 98 participants, 35 male (35%), 63 female (64%), M_{age} = 44.7 years (SD = 13.0), which is lower than we aimed for. However, the sample size still meets the minimal sample size needed for our experimental design to find a statistically large effect with a power of 0.90 (see Exp. 1). Data analyses were conducted only after completion of data collection.

Materials. Like in Experiment 1, we used eight material sets. These were used to create a script of a lunch conversation between three colleagues discussing eight competence-related behaviors of Brinkers followed by either literal or ironic remarks. To increase ecological validity, discussions were complemented with some small talk. The script was played out by three actors (two women and one man) in a recording studio. Using Adobe Audition software, sound recordings from these actors were mixed into a fluent three-person conversation. This was done by entering the sentences spoken by each actor, as well as sentence parts that varied between conditions (i.e., the positive and negative behavior descriptions and positive and negative remarks), on separate tracks. The final six recordings used in our experimental conditions (i.e., 3 communication pattern X 2 behavior sequence) were created by merging the relevant tracks (i.e., setting the relevant behavior descriptions and remarks to mute or play). These recordings thus only varied in valence of behavior description and valence of remark while everything else (including timing) was identical. As in Experiment 1, irony was thus only determined by the context of the remark.

To make the recordings sound like a realistic, interactive, and continuous lunch conversation, we included cafeteria background sounds, and interjections (e.g., hmm, right, sure) from the conversation partners. By mixing these sounds on separate tracks, we ensured they were identical in all versions. The duration of the final edited recordings was 3.58 minutes.

Procedure. The procedure and instruction were identical to Experiment 1, except in differences mentioned above (i.e., group names, listening to sound recordings, etc.).

Dependent variables. Compared to Experiment 1, we more extensively, and first, measured the perceived impression in speakers. Essentialism was also measured as perceived in the speakers. Participant's own category impression was subsequently measured.

Perceived impression in speakers. To measure *Speakers' impression of competence*, we asked whether participants thought their colleagues believed

Brinkers to have a specific set of traits. Participants rated five competence-related traits—cleverness, logical thinking, can be taken seriously, have a valuable opinion, have something valuable to add—on 7-point scales ranging from 1 = not at all agree, to 7 = completely agree. One additional general item stated "The Brinkers are . . .," 1 = incompetent people, 7 = competent people." These six items were combined into a reliable scale (Cronbach's α = .96). This question matrix also included two warmth-related traits (friendly, helpful), which we combined to measure *Speakers' impression of warmth* (Cronbach's α = .92).

Two 7-point items related to speakers' category-perception strength: *Speakers' uniformity* asked participants the degree to which they perceived their colleagues to think that Brinkers are all alike (1 = not at all alike, 7 = very much alike); *Speakers' informativeness* asked: "Suppose your colleagues only know about a person that he or she is a Brinker. Do your colleagues think that, based on this information, they can say (1) very little or (7) very much about the characteristics of this person?"

Perceived essentialism in speakers about positive and negative behaviors. Next, participants were consecutively presented with written descriptions of the eight specific behaviors they heard in the recording, without the colleague's (literal/ ironic) remark. After each behavior description, participants answered five items measuring what they thought was their colleagues' perceived essentialism of the particular behavior. These items measured *Expectedness* (To what extent do your colleagues think the described behavior of these Brinkers is expected for a Brinker? 1 = not at all expected, 7 = very much expected); *Repetition likelihood* (How large do your colleagues estimate the chance that the Brinkers will repeat the described behavior in the future? Please note a number between 0 and 100%); *Attribution* (To what extent do your colleagues think the described behavior of these Brinkers is due to the situation or the people? 1 = completely due to the situation, 7 = completely due to the person); *Typicality* (To what extent do your colleagues think the described behavior is typical for the Brinkers? 1 = not at all typical, 7 = very typical); *Generalizeability* (How likely is it, according to your colleagues, that the Brinkers will repeat the same behavior in other situations? 1 = very unlikely, 7 = very likely).

Before creating an overall essentialism scale, we checked whether individual essentialism items formed an internally consistent scale across the eight behavior judgments. These were sufficient (Cronbach's α for items across judgments of the four positive and four negative behaviors were respectively: .86, .85 for Expectedness, .83, .85 for Repetition likelihood, .73, .84 for Attribution, .74 and .88 for Typicality, .69, .87 for Generalizeability). Subsequently, these items were combined to form the speaker's essentialism for positive (Cronbach's α = .78) and negative (Cronbach's α = .95) behaviors. Thus, where the essentialism items in Experiment 1 did not form a reliable scale, we can analyze speakers' essentialism of positive and negative behavior, each in one measure. Note that, in this experiment, items referred to colleagues' (rather than participants' own) impression and to the group as a whole (i.e., "the Brinkers"), whereas in Experiment 1, the items (as well as the behavior descriptions) referred to individual category members (i.e., "this Ruysdaeler").

Participants' own impression. Next, participants rated their own impression of Brinkers, based on what they had heard about them in the fragment. *Own impression of competence* (Cronbach's α = .94) and *Own impression of warmth* (Cronbach's α = .90) were measured using respectively the same six items as speakers' impression competence, and the same two items for warmth, yet formulated such that they referred to participants' own, rather than their colleagues' impression. No further variables were measured.

Results

As in Experiment 1, we observed no significant effects of the Behavior sequences factor in the main analyses reported below (most Fs < 1) and further exclude this factor in the analyses.

Table 2. Means (and standard deviations) of dependent variables (separated for perceived impression in speakers vs. participants' own) in the three communication pattern conditions (Experiment 2).

Dependent variable	**Communication pattern condition**		
	All literal (n = 39)	Ironic about positive behavior (n = 28)	Ironic about negative behavior (n = 31)
Perceived impression in speakers			
Speakers' impression of competence	4.29_a (1.38)	2.78_b (1.62)	4.55_a (1.60)
Speakers' impression of warmth	4.83_a (1.27)	3.54_b (1.76)	4.48_a (1.55)
Speakers' uniformity	5.18_a (1.41)	5.61_a (1.71)	5.32_a (1.40)
Speakers' informativeness	4.72_a (1.40)	5.29_a (1.78)	4.58_a (1.61)
Speakers' essentialism of positive behaviors	4.70_a (1.11)	4.29_a (1.29)	4.94_a (1.09)
Speakers' essentialism of negative behaviors	4.90_{ab} (1.09)	5.50_a (1.35)	4.36_b (1.08)
Participants' own impression			
Own impression of competence	4.62_a (1.28)	4.70_a (1.59)	4.63_a (1.31)
Own impression of warmth	5.04_a (1.33)	4.95_a (1.69)	4.94_a (1.81)
Own uniformity	4.23_a (1.42)	3.96_a (1.23)	4.61_a (1.61)
Own informativeness	3.26_a (1.57)	3.39_a (1.73)	3.74_a (1.59)

Note. $N = 98$. Means in rows with a different subscript (a, b) differ significantly according to Tukey posthoc tests ($p < .05$).

Perceived impression in speaker

Results for perceived impression held by the speakers were comparable to Experiment 1 and in line with H1a (but not H1b).

A MANOVA with Communication pattern as independent impression of competence, Speakers' impression of competence, Speakers' impression of warmth, Speakers' uniformity, and Speakers' informativeness as dependent variables showed a significant multivariate effect, Wilks' $\Lambda = 0.74$, $F(8,184) = 3.79$, $p < .001$, $\eta_p^2 = 0.14$. Follow-up univariate ANOVAs showed a main effect of Communication pattern on Speakers' impression of competence of the target group, $F(2,95) = 11.63$, $p < .001$, $\eta_p^2 = 0.20$. A similar pattern, although the effect is smaller, was observed on Speakers' impression of warmth, $F(2,95) = 6.17$, $p = .003$, $\eta_p^2 = 0.12$. As the discussed behaviors and remarks were merely about competence, it appears that the communication patterns result in a general halo effect in evaluating the target, like in Experiment 1.

As can be seen in Table 2, colleagues are perceived to convey a more negative impression when they make ironic remarks about positive behaviors, compared to both "all literal" and the "ironic about negative behavior" conditions. Ironic remarks about positive behavior (e.g., "Gee, those Brinkers are terrible presenters," after they gave a successful presentation) introduce a negative evaluation and thereby convey a negative expectancy. The perceived impression in the "Ironic about negative behavior" condition, was, like in Experiment 1, not different from the "all literal" control condition. The effects we observed in Experiment 1 on the perceived strength of category perception as measured in speakers' uniformity and informativeness did not replicate (Fs < 1.7, *ns*).

With respect to Hypothesis 3, in line with Experiment 1, we observed no effects of communication pattern on the four variables measuring participants' own impression about the category; Wilks' $\Lambda = 0.96$, $F < 1$, univariate Fs < 1.5, *ns*. We also did not observe any of the reverse effects shown in Experiment 1.

Perceived essentialism in speakers about positive and negative behaviors

Essentialism (i.e., in this Experiment measured as perceived in speakers) was measured for each specific positive and negative behavior to which the ironic/literal remarks of colleagues had previously referred. We conducted a 3 (Communication Pattern condition) X 2 (Behavioral valence: positive, negative) ANOVA with repeated measures on the second factor, on the speakers' essentialism scale. The predicted interaction was significant $F(2,95) = 7.82$, $p = .001$, $\eta_p^2 = 0.14$. This revealed a pattern in line with the Irony Bias and H2.

Table 2 shows that participants in the condition with "ironic remarks about positive behavior" infer that speakers have a relatively low perception of essentialism of positive behaviors, and a high perception of essentialism of negative behaviors. Ironic remarks about positive behavior introduce negative evaluations (e.g., "Oh, that's really clumsy of the Brinkers") after a good performance of the target group, and convey the impression that the good performance was unexpected and atypical. In contrast, participants in the condition with "ironic remarks about negative behavior" ("Oh, that's really good of the Brinkers," after a bad performance) inferred that speakers have a relatively low perception of essentialism of negative behaviors, and a high perception of essentialism of positive behaviors. That is, these remarks convey the impression that the bad performance was unexpected and atypical. Looking at cell comparisons, differences between the "ironic about positive behavior" and "ironic about negative behavior" conditions are significant only with respect to perceived essentialism of negative behaviors. The scores in the "all literal" condition fall in between the scores in these two conditions, but are not significantly different. We observed no main effects of Communication pattern ($F < 1$, *ns*), nor of Behavioral valence, $F(1,95) = 2.45$, $p = .12$, $\eta_p^2 = 0.03$. Additional analyses on the individual essentialism items, in line with those in Experiment 1, can be found in Appendix F.

The fact that participants only heard the ironic and/or literal remarks in the recording and that these were not shown when participants reported on the impression they perceived in their colleagues, suggests the remarks leave a lasting impression.

General Discussion

The present studies show that recipients of seemingly harmless, but biased ironic remarks about a previously unknown target group draw inferences about the speaker's stereotypic impression of this target category. In both studies, using different materials and modalities of presentation, participants recognized the biased view of their colleagues in their communication pattern. That is, when speakers made ironic remarks about category members' positive behaviors, recipients inferred that these speakers had a more negative impression of the category, and (in Experiment 2) that they perceived higher essentialism of negative rather than positive behaviors. The observed effect sizes ($\eta_p^2 \geqslant 0.14$) can be qualified as "large" (Richardson, 2011). The findings complement previous research on the Irony Bias (Burgers & Beukeboom, 2016; Burgers et al., 2015, Study 2), and on recipient inferences from biased variations in language abstraction (Assilaméhou & Testé, 2013; Douglas & Sutton, 2006), by showing that recipients draw inferences about speakers' stereotype valence, content, and level of essentialism from their use of ironic remarks.

An important note is that, against H1b, speakers making ironic remarks about negative behaviors were not perceived to convey a more positive impression compared to the "all literal" control condition. This fits with Burgers and Beukeboom (2016) who showed that both production and reception effects of the Irony Bias were less pronounced in conditions in which ironic remarks evaluate negative (vs. positive) behaviors. One reason is that irony is more often used in negative situations, as a politeness strategy to mitigate negative evaluations (Brown & Levinson, 1987; Dews & Winner, 1995). Given that these other reasons to use irony exist when referring to negative behaviors, rather than alluding to

stereotypic expectancies, it makes sense that recipients infer less information about the speaker's stereotypic expectancies from irony about negative behaviors.

Understanding recipient inferences is important because they constitute a crucial element in the interpersonal transmission and maintenance of stereotypic expectancies (Maass, 1999; Wigboldus & Douglas, 2007). Recipients are able to extract information about prevailing attitudes and impressions from subtle variations in speakers' remarks about targets (Douglas & Sutton, 2006). But, for actual stereotype transmission to occur, recipients should also internalize the stereotypic view they perceive in a speaker's biased remarks into their own category representation. In the present research, the biased communication pattern perceived in speakers did not translate to participants' own reported category impression. Future research can focus on the factors that determine whether stereotype formation occurs. We expect that the following factors are important.

First, whether induction from received information about category members to a generalized category impression occurs depends on metajudgmental cues about the value and diagnosticity of the information (Paolini, Crisp, & McIntyre, 2009). One important factor relates to the sender of the information. It is likely that recipients mainly end up endorsing a stereotype when the source of the information is an in-group speaker. Research shows that stereotypes are bolstered when learning these are consensually shared within an ingroup (Haslam et al., 1997; Sechrist & Stangor, 2001). In the present studies, the source of the information was presented as in-group members (i.e., colleagues of one's own team), yet given that the scenario was imaginary, in-group identification was probably low.

Second, people may be reluctant to form a generalized stereotypic view after a few biased remarks about a target group. Initially, recipients may resist and even correct their own impression (see Experiment 1). Such reluctance may be particularly strong when the category is unknown, as in the present studies, and impressions have not yet been formed. In such cases, the formation of new category stereotypes may require more exposure compared to maintenance of existing stereotypes. After repeated exposure to biased remarks about a particular group, recipients may, however, follow suit, and the stereotypic impression may be internalized (Arendt, 2013; Arendt & Northup, 2015).

Third, various factors relate to how social-category information is presented. In the present experiments, aside from the (biased) remarks, recipients also learned about category members' concrete behaviors. These were balanced in valence. Knowledge about individual target-group members' behaviors can mitigate the effect of overly biased communication on category formation and maintenance (Park & Hastie, 1987). When having a balanced view of category members' behaviors, biased remarks may be perceived as unnecessarily prejudiced. In such cases, biased remarks can have no effect on recipients' own impression or induce reverse effects through overcorrection (Experiment 1).

Stereotype formation also depends on whether recipients obtain knowledge about individual category members' behaviors or generic information about the category, generalizing across individual category members. In Experiment 1, both behaviors and remarks referred to individual target group members (This Ruysdaeler is . . .). In Experiment 2, described behaviors and ironic/literal remarks were formulated more generically (i.e., Brinkers are . . .). As the experiments differed in other ways as well, we cannot conclude whether this contributed to different effects, but theoretically generic labels (compared to individual labels) should more directly induce associations between labeled categories and essential characteristics (Gelman et al., 2010; Rhodes et al., 2012). Moreover, when a group of individuals is labeled using a (generic) noun label, the group is likelier perceived as meaningful and coherent (i.e., *high entitativity*), which in turn facilitates stereotype formation (Beukeboom & Burgers, 2019).

Finally, in real life, stereotypes are likely formed spontaneously during interaction, and without an explicit instruction to form an

impression about a target group. Particularly in the scenario of lunch conversations with new colleagues, when desiring to get along, recipients would react and reinforce speakers, and thereby create a shared stereotypic impression of a target group (Bratanova & Kashima, 2014; Ruscher & Hammer, 2006). Being actively involved in such in-group communication likely amplifies one's own target group beliefs (Ruscher, Cralley, & O'Farrell, 2005).

In sum, we demonstrated that recipients infer speakers' stereotypic impressions from their pattern of ironic remarks. Whether recipients self-adopt the conveyed stereotypic impression about the target category depends on various factors that may be explored in future research. Research on the mechanisms and consequences of linguistic bias allows us to reveal their subtle effects on stereotype formation and maintenance, and this in turn could help prevent potentially negative effects.

Acknowledgements

The authors like to thank Marit Mol and Lisanne de Hoop for their help in the designing and data collection of Experiments 1 and 2, respectively.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Camiel J. Beukeboom https://orcid.org/0000-0002-0364-2784
Christian Burgers https://orcid.org/0000-0002-5652-9021

Supplemental material

Supplemental material for this article is available online.

Notes

1. In the literature, the term inductive potential is often used, but as induction refers to stereotype formation and deduction refers to the process of applying stereotypes to a given target situation (Beike & Sherman, 1994), deductive potential is the proper term in this context.
2. We made sure the used labels were non-existing as category labels and neutral in meaning. Jacob van Ruysdael is a 17th-century Dutch painter; Brink is Dutch for village square.
3. Burgers et al. (2015, experiment 2) used two MANOVAs to estimate the effect of the independent variable of communication pattern on two sets of dependent variables measuring inferences about the speaker's perceived communicative goals and group membership. They reported report effect sizes of partial eta squared of 0.24 (perceived communicative goals) and 0.30 (perceived group membership) in their MANOVA, and values of partial eta squared of 0.37, 0.38, and 0.13 for the individual communicative-goal variables in the follow-up ANOVAs and values of partial eta squared of 0.39 and 0.16 for the individual group-membership variables (Burgers et al., 2015, pp. 447–449). These effect sizes can be qualified as "large" (Richardson, 2011).
4. Only in this study did we also measure *Own clarity of impression*: How clear is the impression you formed about the characteristics of the Ruysdaelers? (1 = unclear, 7 = clear). We observed no differences between conditions and results are not further reported.

References

Appel, M., & Weber, S. (2017). Do mass mediated stereotypes harm members of negatively stereotyped groups? A meta-analytical review on media-generated stereotype threat and stereotype lift. *Communication Research*. Epub ahead of print 12 July 2017. doi: 10.1177/0093650217715543.

Arendt, F. (2013). Dose-dependent media priming effects of stereotypic newspaper articles on implicit and explicit stereotypes. *Journal of Communication*, *63*, 830–851. doi: 10.1111/jcom.12056

Arendt, F., & Northup, T. (2015). Effects of long-term exposure to news stereotypes on implicit and explicit attitudes. *International Journal of Communication*, *9*, 2370–2390.

Assilaméhou, Y., & Testé, B. (2013). The effects of linguistic abstraction on evaluations of the speaker in an intergroup context: Using the Linguistic Intergroup Bias makes you a good group member. *Journal of Experimental Social Psychology*, *49*, 113–119. doi: 10.1016/j.jesp.2012.08.001

Attardo, S. (2000). Irony as relevant inappropriateness. *Journal of Pragmatics*, *32*, 793–826. doi: 10.1016/S0378-2166(99)00070-3

Beike, D. R., & Sherman, S. J. (1994). Social inference: Inductions, deductions, and analogies. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., *Vol. 1*, pp. 209–285). Hillsdale, NJ: Erlbaum.

Beukeboom, C. J. (2014). Mechanisms of linguistic bias: How words reflect and maintain stereotypic expectancies. In J. Laszlo, J. P. Forgas, & O. Vincze (Eds.), *Social cognition and communication* (pp. 313–330). New York, NY: Psychology Press.

Beukeboom, C. J., & Burgers, C. (2019). How stereotypes are shared through language: A review and introduction of the Social Categories and Stereotypes Communication (SCSC) framework. *Review of Communication Research*, *7*, 1–37. doi: 10.12840/issn.2255-4165.017

Beukeboom, C. J., Burgers, C., Szabó, Z. P., Cvejic, S., Lönnqvist, J.-E. M., & Welbers, K. (2019). The negation bias in stereotype maintenance: A replication in five languages. *Journal of Language and Social Psychology*. Epub ahead of print 31 August 2019. doi: 10.1177/0261927X19869759

Beukeboom, C.J., Finkenauer, C., & Wigboldus, D. H. J. (2010). The negation bias: When negations signal stereotypic expectancies. *Journal of Personality and Social Psychology*, *99*, 978–992. doi: 10.1037/a0020861

Bratanova, B., & Kashima, Y. (2014). The "saying is repeating" effect: Dyadic communication can generate cultural stereotypes. *The Journal of Social Psychology*, *154*, 155–174. doi: 10.1080/00224545.2013.874326

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge, UK: Cambridge University Press.

Burgers, C., & Beukeboom, C. J. (2016). Stereotype transmission and maintenance through interpersonal communication: The irony bias. *Communication Research*, *43*, 414–441. doi: 10.1177/0093650214534975

Burgers, C., Beukeboom, C. J., Kelder, M., & Peeters, M. M. E. (2015). How sports fans forge intergroup competition through language: The case of verbal irony. *Human Communication Research*, *41*, 435–457. doi: 10.1111/hcre.12052

Cimpian, A., & Markman, E. M. (2008). Preschool children's use of cues to generic meaning. *Cognition*, *107*, 19–53. doi: 10.1016/j.cognition.2007.07.008

Collins, K. A., & Clément, R. (2012). Language and prejudice: Direct and moderated effects. *Journal of Language and Social Psychology*, *31*, 376–396. doi: 10.1177/0261927X12446611

Cuddy, A. J. C., 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. doi: 10.1016/S0065-2601(07)00002-0

Dews, S., & Winner, E. (1995). Muting the meaning: A social function of irony. *Metaphor and Symbol*, *10*, 3–19. doi: 10.1207/s15327868ms1001_2

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. doi: 10.1016/j.jesp.2005.06.001

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. doi: 10.3758/BRM.41.4.1149

Filik, R., Ţurcan, A., Thompson, D., Harvey, N., Davies, H., & Turner, A. (2016). Sarcasm and emoticons: Comprehension and emotional impact. *The Quarterly Journal of Experimental Psychology*, *69*, 2130–2146. doi: 10.1080/17470218.2015.1106566

Garcia-Marques, L., Santos, A. S. C., & Mackie, D. M. (2006). Stereotypes: Static Abstractions or Dynamic Knowledge Structures? *Journal of Personality and Social Psychology*, *91*, 814–831. https://doi.org/10.1037/0022-3514.91.5.814

Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology*, *61*, 273–301. doi: 10.1016/j.cogpsych.2010.06.001

Gibbs, R. W., Jr. (2000). Irony in talk among friends. *Metaphor and Symbol*, *15*, 5–27. doi: 10.1080/10926488.2000.9678862

Gorham, B. W. (2006). News media's relationship with stereotyping: The linguistic intergroup bias in response to crime news. *Journal of Communication*, *56*, 289–308. doi: 10.1111/j.1460-2466.2006.00020.x

Haslam, S. A., Turner, J. C., Oakes, P. J., McGarty, C., & Reynolds, K. J. (1997). The group as a basis for emergent stereotype consensus. *European Review of Social Psychology*, *8*, 203–239. doi: 10.1080/14792779643000128

Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology*, *46*, 44–56. doi: 10.1037/0022-3514.46.1.44

Hogg, M. A., & Reid, S. A. (2006). Social identity, self-categorization, and the communication of group norms. *Communication Theory*, *16*, 7–30. doi: 10.1111/j.1468-2885.2006.00003.x

Kihara, Y. (2005). The mental space structure of verbal irony. *Cognitive Linguistics*, *16*, 513–530. doi: 10.1515/cogl.2005.16.3.513

Kreuz, R., Roberts, R., Johnson, B., & Bertus, E. (1996). Figurative language occurrence and co-occurrence in contemporary literature. In R. Kreuz & M. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 83–97). Norwood, NJ: Ablex.

Maass, A. (1999). Linguistic intergroup bias: Stereotype perpetuation through language. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (*Vol. 31*, pp. 79–121). San Diego, CA: Academic Press.

Maass, A., Salvi, D., Arcuri, L., & Semin, G. R. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*, 981–993. doi: 10.1037/0022-3514.57.6.981

Matthews, J. K., Hancock, J. T., & Dunham, P. J. (2006). The roles of politeness and humor in the asymmetry of affect in verbal irony. *Discourse Processes*, *41*, 3–24. doi: 10.1207/s15326950dp4101_2

Medin, D, & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge, UK: Cambridge University Press.

Paolini, S., Crisp, R. J., & McIntyre, K. (2009). Accountability moderates member-to-group generalization: Testing a dual process model of stereotype change. *Journal of Experimental Social Psychology*, *45*, 676–685. https://doi.org/10.1016/j.jesp.2009.03.005

Park, B., & Hastie, R. (1987). Perception of variability in category development: Instance- versus abstraction-based stereotypes. *Journal of Personality and Social Psychology*, *53*, 621–635. doi: 10.1037/0022-3514.53.4.621

Ramasubramanian, S. (2011). The impact of stereotypical versus counterstereotypical media exemplars on racial attitudes, causal attributions, and support for affirmative action. *Communication Research*, *38*, 497–516. doi: 10.1177/0093650210384854

Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*, 13526–13531. doi: 10.1073/pnas.1208951109

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*, 135–147. doi: 10.1016/j.edurev.2010.12.001

Ruscher, J. B., Cralley, E. L., & O'Farrell, K. J. (2005). How newly acquainted dyads develop shared stereotypic impressions through conversation. *Group Processes & Intergroup Relations*, *8*, 259–270. doi: 10.1177/1368430205053942

Ruscher, J. B., & Hammer, E. D. (2006). The development of shared stereotypic impressions in conversation: An emerging model, methods, and extensions to cross-group settings. *Journal of Language and Social Psychology*, *25*, 221–243. doi: 10.1177/0261927X06289424

Sechrist, G. B., & Stangor, C. (2001). Perceived consensus influences intergroup behavior and stereotype accessibility. *Journal of Personality and Social Psychology*, *80*, 645–654. doi: 10.1037/0022-3514.80.4.645

Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology*, *39*, 75–82. doi: 10.1016/S0022-1031(02)00512-7

Wallace, B. C. (2015). Computational irony: A survey and new perspectives. *Artificial Intelligence Review*, *43*, 467–483. doi: 10.1007/s10462-012-9392-5

Whalen, J. M., Pexman, P. M., & Gill, A. J. (2009). "Should be fun—not!": Incidence and marking of nonliteral language in e-mail. *Journal of Language & Social Psychology*, *28*, 263–280. doi: 10.1177/0261927X09335253

Wigboldus, D., & Douglas, K. (2007). Language, stereotypes, and intergroup relations. In K. Fiedler (Ed.), *Social communication* (pp. 79–106). New York, NY: Psychology Press.

Wigboldus, D. H., 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. doi: 10.1037/0022-3514.78.1.5

Wilson, D., & Sperber, D. (2004). Relevance theory. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 607–632). Oxford, UK: Blackwell.