

Examining automatic stereotyping from a propositional perspective: Is automatic stereotyping sensitive to relational and validity information?

Tal Moran, Jamie Cummins, & Jan De Houwer

Ghent University, Belgium

In press: Personality and Social Psychology Bulletin

Research on automatic stereotyping is dominated by the idea that automatic stereotyping reflects the activation of (group-trait) associations. In two preregistered experiments (total N=391) we tested predictions derived from an alternative perspective that suggests that automatic stereotyping is the result of the activation of propositional representations that, unlike associations, can encode relational information and have truth values. Experiment 1 found that automatic stereotyping is sensitive to the validity of information about pairs of traits and groups. Experiment 2 showed that automatic stereotyping is sensitive to the specific relations (e.g., whether a particular group is more or less friendly than a reference person) between pairs of traits and groups. Interestingly, both experiments found a weaker influence of validity/relational information on automatic stereotyping than on non-automatic stereotyping. We discuss the implications of these findings for research on automatic stereotyping.

Keywords: Automatic stereotyping; Implicit stereotypes; Validity; Relational information; Propositional representations

Within every society, there are widely shared beliefs about the characteristics of social groups and their members (Ashmore & Del Boca, 1981). These societal stereotypes influence, amongst other things, which demographic groups are more likely to be arrested (Spencer et al., 2016), the kinds of medical treatment members of those groups receive (Moskowitz et al., 2012) and the kinds of jobs members of those groups are offered (Koch et al., 2015).

From a cognitive perspective, societal stereotypes are typically construed as *stereotype representations* at the mental level. Activation of these representations results in *stereotyping*, that is, behaving in line with these stereotypes (see De Houwer, Gawronski et al., 2013 for a similar conceptualization in the domain of attitudes). Stereotypes are usually distinguished from evaluations; whereas evaluations vary along a positive-negative continuum, stereotypes involve nonvalenced semantic content (Amodio & Ratner, 2011; Greenwald et al., 2002). In the last two decades, researchers have started to study *automatic stereotyping*.¹ Automatic stereotyping can be defined broadly as a behavior (a stereotypic response) that is emitted under suboptimal conditions (De Houwer, 2019; De Houwer, Barnes-Holmes et al., 2013; De Houwer et al., 2020). Conditions that are typically considered to be suboptimal include those in which there is limited time to respond, where people are engaged simultaneously in other demanding tasks, or where people do not have the explicit goal to engage in stereotyping (see Moors, 2016, and De Houwer et al., 2009, for more details). Whereas non-automatic stereotyping is typically measured by directly asking people under

Correspondence concerning this article should be addressed to Tal Moran, Department of Experimental-Clinical and Health Psychology, Ghent University, Dunantlaan 2, 9000 Ghent, Belgium. E-mail: Tal.MoranYorovich@Ugent.be. This manuscript is supported by Ghent University under grant BOF16/MET_V/002. The authors declare that they have no competing interests. The materials, data and analysis scripts of the whole project are available at the Open Science Framework (osf.io/c8kyb/)

¹ Previous research has often used the term “implicit stereotypes” to describe this construct (e.g., Blair et al., 2001; Greenwald et al., 2002).

optimal conditions about what they think and do, automatic stereotyping is usually measured with indirect (implicit) measures, which attempt to capture this information under sub-optimal conditions. For example, one of the main measures of automatic stereotyping is the Implicit Association Test (IAT; Greenwald et al., 1998) which requires participants to respond quickly (i.e., participants are instructed to respond as quickly as possible) and fosters conditions where stereotyping responses will be emitted unintentionally (i.e., participants are not asked to make a stereotypic response or to report their stereotypes).

The pervading view on automatic stereotypes for decades has been that stereotype representations are mental associations (e.g., women-emotional), and that automatic stereotyping reflects the activation of these associations (e.g., Blair et al., 2001; Divine, 1989; Greenwald et al., 2002; Kawakami et al., 2017). For example, Divine, (1989) argued that stereotypes are a “well-learned set of associations... that [are] automatically activated in the presence of a member (or symbolic equivalent) of the target group” (p. 6). Greenwald et al. (2002) defined stereotypes as “the association of a social group concept with one or more (non-valenced) attribute concepts” (p. 5) and argued that stereotyping measured with indirect measures, such as the IAT, reflect the activation of these associations.

The dominance of associative theorizing on automatic stereotyping, however, does not align well with the fact that societal stereotypes typically involve relational beliefs about the characteristics of members of certain groups (e.g., Hilton & Von Hippel, 1996). The relational nature of societal stereotypes can be illustrated by the distinction between descriptive stereotypes (e.g., “women are emotional”) and prescriptive stereotypes (e.g., “woman should be emotional”; e.g., Burgess & Borgida, 1999), a distinction that can only be drawn in terms of relational terms (e.g., “are” versus “should be”). Importantly, initial evidence suggests that evidence for this distinction can be also found under sub-optimal conditions, that is, at the level of automatic stereotyping (Cummins & De Houwer, 2019).

In the current research, we suggest an alternative perspective on automatic stereotyping: that stereotype representations are propositions, and that automatic stereotyping results from the activation of propositions, not associations. Propositional representations differ from associative representations in two key ways: (a) they encode relational information (i.e., they specify how stimuli are related) and (b) they have a

truth value (i.e., they can be evaluated as being true or false). Indeed, propositional perspectives on social psychological phenomena have already shown promise elsewhere, such as in the context of automatic evaluation (De Houwer, 2014; De Houwer et al., 2020). This perspective on automatic evaluation has recently challenged the dominant associative perspective which purports that automatic evaluation reflects the activation of associations between a target and a valence concept (e.g., a Bob–positive association mediates the evaluation of Bob; Gawronski & Bodenhausen, 2006). According to a propositional perspective, automatic evaluation reflects the automatic activation of propositional representations (e.g., “Bob is a good person”; De Houwer, 2014). Critically, this novel propositional perspective has inspired investigation into a host of new findings which do not follow readily from the associative perspective (see De Houwer et al., 2020, for a review).

Similar to the context of automatic evaluation, the propositional perspective also offers new predictions in the context of automatic stereotyping. First, because propositions (e.g., “women are emotional”) have a truth value, a propositional perspective predicts a sensitivity of automatic stereotyping to validity information. Second, because propositions specify the relation between the two concepts (group and trait), a propositional perspective predicts a sensitivity of automatic stereotyping to the relation between groups and traits.

The question of whether automatic stereotyping is sensitive to relational and validity information is highly important. As part of their everyday lives, people cannot help being exposed to information about social groups. However, not all of the information people are exposed to is true or accurate. We are often exposed to false or inaccurate information about social groups in the media and through speculation and gossip (e.g., Martin et al., 2014). From the propositional perspective, this leads to the question: will automatic stereotyping be influenced by information about social groups *even when a person knows that this information is invalid?* Furthermore, information given to us about members of social groups is almost always relational in nature. For example, we can learn that members of a social group are related to hostility because they are usually suffering as a result of it, or because they usually cause hostility (Andreychik & Grill, 2012). This leads to a further question: will automatic stereotyping be influenced *by the specific relation between social groups and traits?* Except for a

few limited studies (described below), these questions have not been explored in the past in the context of automatic stereotyping.

Before we discuss the initial base of evidence for a propositional perspective of automatic stereotyping, we would like to emphasize two things. First, our goal with the present research is not to prove which of the perspectives (associative versus propositional) to automatic stereotypes is correct. From a similar debate in the domain of automatic evaluation we know that even for findings that confirm the predictions of propositional models, it will often be possible to devise alternative explanations in terms of associations, and vice versa (De Houwer et al., 2020). Our goal here is to use the generative power of the idea that automatic stereotyping results from activation of propositional representations in order to formulate new and interesting research questions (e.g., is automatic stereotyping sensitive to validity information?) that were hardly investigated before. Importantly, the answers to these can be later used to constrain different theoretical models of automatic stereotyping.

Second, we can draw from the fact that the propositional perspective has already inspired a substantial amount of research on automatic evaluation (for a review see De Houwer et al., 2020). We hope that the present research is the start of a similarly informative line of research on automatic stereotyping. Supporting the predictions of the propositional perspective, research has accumulated evidence that that *automatic evaluation* can be sensitive to relational and validity information (e.g., Moran et al., 2015, 2017; Peters & Gawronski, 2011). However, the fact that *automatic evaluation* has shown sensitivity to relational and validity information does not necessarily mean that the same pattern will be observed for *automatic stereotyping*. Debate around the nature of the link between automatic evaluation and automatic stereotyping is still on-going, with some suggesting a weak relation between them (Amodio & Devine, 2006; Amodio & Ratner, 2011; Cunningham et al., 2004; Rudman & Goodwin, 2004; Sabin et al., 2008) and others suggesting a strong relation (Kurdi et al., 2019; Phills et al., 2020). It is therefore important to elaborate on a propositional account of automatic stereotyping, in order to better understand the nature of its relationship to automatic evaluation.

We are aware of one published study that found sensitivity of a new developed automatic stereotyping measure (Truth Misattribution Procedure; TMP) to relational information (Cummins & De Houwer, 2019).

There is some evidence suggesting that automatic stereotyping can be sensitive to validity information. Jabold (2020) investigated the effect of crime reports on criminal stereotyping of novel groups, and whether stereotyping is updated when new crime statistics negate the initial learned information (i.e., suggesting that the initial crime reports are invalid). Results indicated that automatic stereotyping was sensitive to the validity of the initial stereotypic information. Additionally, Rubinstein et al. (2018) tested whether race stereotypes (e.g., Black people are less intelligent than White people) or individuating information (e.g. SAT scores) influence the automatic stereotyping of individuals as intelligent or not. When valid individuating information contradicted the race stereotypes, automatic stereotyping reflected only the individuating information. However, both these studies confounded *stereotyping* with *evaluations*. Jabold (2020) examined stereotyping against criminals versus lawful people, which are groups that clearly differ in valence. Rubinstein et al. (2018) focused on stereotypes about intelligent vs. stupid people, a distinction that is also confounded with valence. Hence, it is unclear if these previous studies provide evidence supporting sensitivity to validity information of *automatic stereotyping* or merely *automatic evaluation*.

The Current Research

In the present research, we tested if automatic stereotyping is sensitive to validity information and relational information. As in previous research on stereotyping (e.g., Crawford et al., 2002; Dotsch et al., 2013; Phills et al., 2020; Sherman, 1996; Smith & Zarate, 1990), we simulated automatic stereotyping processes in the lab by presenting information about the traits of novel groups. Experiment 1 tested whether automatic stereotyping of novel groups was influenced by the validity of the information presented. We asked participants to form impressions of two novel groups, each paired with verbal descriptions that imply that one group is warm, and the other group is competent. We then informed half of the participants that the information had accidentally been mixed up, and that the traits attributed to one group were actually about the other group, and vice versa. We tested if automatic stereotyping (measured with the IAT) was sensitive to this validity information, or whether it reflected only the initial pairing. Experiment 2 tested whether automatic stereotyping was sensitive to the relational content of the information provided. We described two groups in relational terms (e.g., Group A is less friendly than Lisa, Group B is

more friendly than Lisa) and then informed half of the participants that, accidentally, the words “more than” and “less than” had been mixed-up. We then tested whether automatic stereotyping reflected this relational information. Unlike in previous research, we took two steps to ensure that any results we find would be the result of *stereotyping* rather than *evaluations*. First, we used only positively valenced traits (warm and competent) in the manipulations and stereotyping measures. Second, we also measured participants’ (automatic) evaluations of the novel groups and controlled for their influence in our analyses.²

All experiments were preregistered (see links below), and no file-drawered studies exist. Deviations from the preregistered plan are specified within the manuscript. We report all data exclusions, manipulations, measures, and how we determined our sample sizes (Simmons et al., 2012). Participants in all experiments were recruited online via the Prolific Academic website (<https://prolific.ac>). Each experiment took 15 minutes to complete and all participants were paid £1.60. In all experiments, we planned to collect data from 270 participants to have 95% power to find a between-participants effect of medium size ($\eta_p^2 = 0.06$) assuming 30% exclusions based on the preregistered exclusions criteria (see below for details). The materials, data, and analysis scripts of the whole project are available at the Open Science Framework (osf.io/c8kyb/).

Experiment 1

In Experiment 1, to test if automatic stereotyping was sensitive to validity information, we adopted a “mix-up” manipulation that was used in previous research to test the effect of validity information on automatic evaluation (Gregg et al., 2006; Moran et al., 2017). Informing participants of such a “mix-up” in trait-group pairing implies that the initial pairing between each group and trait was invalid and needed to be updated. If we found evidence of such updating in the performances of participants, this would imply that their automatic stereotyping responses were reflective of the validity of the previous pairing between the groups and the traits, which would therefore provide evidence that automatic stereotyping is sensitive to validity information.

Method

² In all the experiments, the results of the analyses remain the same even without controlling for participants’ (automatic) evaluations

Participants. The preregistration of Experiment 1 is available at osf.io/5pbry/. 270 participants completed the study. The preregistered data exclusion involved removing participants who (a) did not fully complete all questions and tasks (1%), (b) had more than 10% trials in the IAT faster than 300ms (9%), or (c) made at least one error on the comprehension questions (17.7%, Table 1 summarize the exclusion according to this criterion as a function of experimental condition). The final sample included 200 participants (55% women, $M_{age} = 33.53$, $SD = 11.70$).

Table 1

Exclusions (%) based on errors on the comprehension questions as a function of the experimental condition in Experiments 1-2.

		Validity/relational information		
Experiment 1		mixed-up	Control	Total (across validity/relational information)
Group-trait assignment	Niffites are warm	29.8%	9.7%	19.3%
	Luupites are warm	21.8%	10%	16.1%
Total (across group-trait assignment)		25.6%	9.8%	17.7%
Experiment 2		Opposition	Control	
Group-trait assignment	Niffites are warm	23.4%	21.4%	22.5%
	Luupites are warm	26.2%	24.2%	25.2%
Total (across group-trait assignment)		24.8%	22.8%	23.8%

Materials and procedure. The full procedure, instructions, and materials of Experiments 1-2 are explained in detail in the online supplement (osf.io/6stnx/). Participants were asked to imagine that they were planning to travel to an exotic island where two tribes (called ‘Niffites’ and ‘Luupites’) live. They were told that Lisa, who has already visited this island and met these two tribes, would provide them with information about those groups. Participants were asked to learn this information and form an impression of each of the tribes. Then, we presented descriptions of the two tribes in three blocks. Each block consisted of four descriptions of each tribe, one at the time, for 4s,

of the novel groups. Details on these analyses can be found in the online supplement (osf.io/6stnx/).

in random order. The descriptions implied that one group was warm, and the other competent (see Table 2). After exposure to the descriptions, participants in the mixed-up condition were informed that, after Lisa provided the information about the tribes, she realized that she had mixed up the two tribes: everything she had told them about the Niffites was really about the Luupites, and everything she had told them about the Luupites was actually about the Niffites. Participants were asked to correct their impression of the tribes based on this new information. Participants in the control condition did not read this information.

Table 2

Descriptions presented in Experiment 1

Warm group	Competent Group
Niffites[Luupites] are friendly	Luupites[Niffites] are smart
Niffites[Luupites] are helpful	Luupites[Niffites] are scientific
Niffites[Luupites] are kind	Luupites[Niffites] are confident
Niffites[Luupites] are good-natured	Luupites[Niffites] are determined

Note. All the traits used to represent the “warm” and “competent” categories were adopted from previous studies that used these traits to represent these two categories (Carlsson & Björklund, 2010; Fiske, Cuddy, & Glick, 2002; Fiske et al., 2007).

Next, all participants completed measures of automatic and self-reported stereotyping. All participants completed the automatic measure first. To measure automatic stereotyping, we used a seven-block IAT (Nosek et al., 2005). ‘Niffites’ and ‘Luupites’ were the target categories and ‘Warm’ and ‘Competent’ were the attribute categories (Table 3 details the items for each category). In two of the IAT’s critical blocks, ‘Niffites’ and ‘Warm’ shared the same response key, and in the other two critical blocks, ‘Luupites’ and ‘Warm’ shared the same response key. The D2 algorithm was used to create stereotyping IAT scores (Greenwald et al., 2003). We computed these scores such that positive scores indicated stereotypes in line with the initial information (median split-half reliability $r = .70$, 95% CI [.63, .76]³). To measure self-reported stereotyping, we asked participants, in two questions, how much they associated the traits ‘warm’ and ‘competent’ with Niffites versus Luupites (with a

7-point Likert scale for responses: 1 = Much more strongly with Niffites, 4 = Equally strongly with Niffites and Luupites, 7 = Much more strongly with Luupites). Self-reported stereotyping scores were calculated by reversing the scores for the ‘competent’ item and then averaging the scores of the two questions, such that positive scores indicate stereotypes in line with the initial information (Cronbach’s Alpha = .96).

Table 3

Items used for each category in the IAT in Experiments 1-2

All IATs		Categories			
		Stereotyping IAT		Evaluation IAT	
Niffites	Luupites	Warm	Competent	Good	Bad
Cellanif	Maasolup	Warm	Competent	Good	Bad
Eskannif	Neenolup	Nice	*Intelligent/Smart	Positive	Negative
Lebbunif	Omeelup	Sociable	Efficient	Pleasant	Unpleasant
Zallunif	Wenaalup	Caring	Skillful	Great	Awful

Note. *Experiment 1 used “Intelligent”, and Experiment 2 used “Smart”. All the traits used to represent the “warm” and “competent” categories in the IAT, were adopted from previous studies that used these traits to represent these two categories (Carlsson & Björklund, 2010; Fiske et al., 2002, 2007).

Next, we assessed participants’ comprehension of the information provided about the two tribes. For each tribe, participants were asked to indicate the best-fitting description of the Niffites and Luupites from a series of options, including options which categorized them as warm/competent. Participants who responded incorrectly to one of these questions (based on both the initial information and the mixed-up information), were excluded from the analyses.⁴

Finally, participants completed automatic and self-reported evaluation measures. All participants completed the automatic measure first. To measure automatic evaluation, we used the IAT, which was similar to the stereotyping IAT, except that we replaced the ‘Warm’ and ‘Competent’ categories with ‘Good’ and ‘Bad’ (see Table 2). To measure self-reported evaluation, participants rated on four different questions how positive or negative, and how good or bad, Niffites and Luupites were (with a 9-point Likert scale for response: -4 = very bad/ very negative; +4 = very good /very positive). For both evaluation measures, scores

³ The reliability of all the IATs that was calculated through the Spearman-Brown corrected correlation between split-half IAT D scores split based on 500 different, random bootstraps of participant data.

⁴ In all the experiments, the results of all the analyses remain the same when including all participants (see more information in Footnotes 6, 9, 12 and 13).

were computed to reflect a preference for Niffites over Luupites (IAT: $r = .61$, 95% CI [.51, .68], ratings: $\alpha = .94$).

Design. The main independent variable was validity information (mixed-up versus control). We counterbalanced, between participants, the group-trait assignment (Niffites are warm versus Luupites are warm), the stereotyping IAT's block order (Niffites-Warm/Luupites-Competent first versus Luupites-Warm/Niffites-Competent first) and the evaluation IAT's block order (Niffites-Positive/Luupites-Negative first versus Luupites-Positive/Niffites-Negative first).

Results

For both the automatic and self-reported stereotyping measures, we tested if the stereotyping scores were predicted by the validity condition (mixed-up vs. control). For the analysis of automatic stereotyping, we controlled for group-trait assignment, stereotyping IAT block order, and their interactions (with each other and with validity condition), and for the IAT evaluation score⁵. The full results of the analysis are detailed in Table 4 (see the end of the manuscript). Importantly, the IAT stereotyping scores were predicted by validity condition, $F(1, 191) = 34.35$, $p < .001$, $\eta_p^2 = .15$, 90% CI [.08, .22], $BF_{10} > 1000$. As predicted, the effect reflected a larger IAT stereotyping score (reflecting stereotypes in line with the initial information) in the control condition ($M = 0.41$, $SD = 0.46$, difference from zero: $t(109) = 9.47$, $p < .001$, *Cohen's d* = 0.90, $BF_{10} > 1000$) than in the mixed-up condition ($M = 0.04$, $SD = 0.49$, difference from zero: $t(89) = 0.70$, $p = .48$, *Cohen's d* = 0.07, $BF_{10} = 0.15$, see Figure 1, panel A).

For the self-reported stereotyping analysis, we controlled for group-trait assignment, its interaction with

validity condition, and the self-reported evaluation of the groups. Self-reported stereotyping scores were predicted by validity condition, $F(1, 195) = 409.33$, $p < .001$, $\eta_p^2 = .68$, 90% CI [.61, .72], $BF_{10} > 1000$. The effect reflected a larger stereotyping score (reflecting stereotypes in line with the initial information) in the control condition ($M = 2.51$, $SD = 0.98$, difference from zero: $t(109) = 26.83$, $p < .001$, *Cohen's d* = 2.56, $BF_{10} > 1000$) than in the mixed-up condition ($M = -1.64$, $SD = 1.88$, difference from zero: $t(89) = -8.26$, $p < .001$, *Cohen's d* = -0.87, $BF_{10} > 1000$, see Figure 1 panel B).⁶

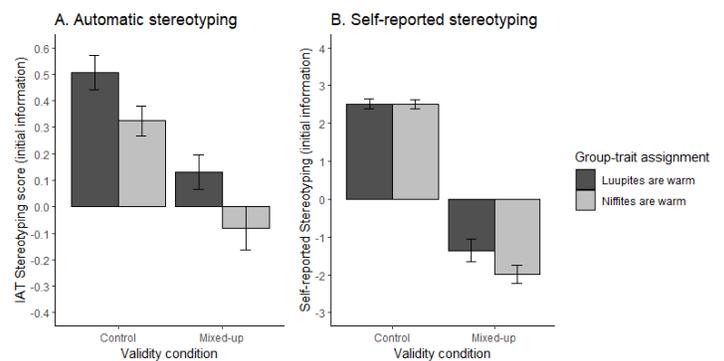


Figure 1. Experiment 1: Stereotyping scores (reflecting stereotypes in line with the initial information) as a function of validity condition (control versus mixed-up) and group-trait assignment (Niffites are warm versus Luupites are warm). Graph A presents the result for automatic stereotyping (measured with the IAT). Graph B presents the result for self-reported stereotyping. Error bars represent standard errors.

Direct comparison between self-reported and IAT stereotyping effects.⁷ To directly compare the effects of validity on the automatic versus self-reported stereotyping measures, we first computed stereotyping scores that reflected a “Niffites are

⁵ In the preregistration of Experiment 1, for both automatic and self-reported stereotyping analyses, we failed to mention that we would control for group-trait assignment. In addition, in the preregistration, we planned to control only for the main effects of all the factors. However, following a reviewer’s recommendation, we now report the results of an ANOVA which includes the validity condition, the group-trait assignment, stereotyping IAT block order (only for the IAT analysis), and their interactions, and the IAT/self-report evaluation score as a covariate.

⁶ The results without exclusions based on the comprehension questions revealed a similar pattern of results. Specifically, the IAT stereotyping scores were predicted by validity condition, $F(1, 234) = 21.12$, $p < .001$, $\eta_p^2 = .08$. The effect reflected a larger IAT stereotyping score in the control condition ($M = 0.35$, $SD = 0.50$, difference from zero: $t(121) = 7.64$, $p < .001$, *Cohen's d* = 0.69) than in

the mixed-up condition ($M = 0.06$, $SD = 0.50$, difference from zero: $t(120) = 1.38$, $p = .17$, *Cohen's d* = 0.13). The self-reported stereotyping scores were predicted by validity condition, $F(1, 238) = 154.04$, $p < .001$, $\eta_p^2 = .39$. The effect reflected a larger self-reported stereotyping score in the control condition ($M = 2.16$, $SD = 1.60$, difference from zero: $t(121) = 14.95$, $p < .001$, *Cohen's d* = 1.35) than in the mixed-up condition ($M = -0.95$, $SD = 2.23$, difference from zero: $t(120) = -4.69$, $p < .001$, *Cohen's d* = -0.43).

⁷ This analysis was not preregistered for Experiment 1. In addition, because the IAT in Experiment 1 always preceded the self-report measure, the interpretation of the results is limited.

Warm/Luupites are competent” stereotype, and standardized these scores. Next, we recoded these scores to reflect stereotyping in line with the initial information. We submitted the standardized scores to a 2 (*Measure type*) x 2 (*validity condition*) x 2 (*group-trait assignment*) x 2 (*stereotyping IAT block order*) mixed ANOVA with measure type as the only within-participants factor. Table 5 (see the end of the manuscript) summarizes the full results of the ANOVA. Importantly, the interaction between validity and measure type was significant, $F(1, 192) = 46.02, p < .001, \eta_p^2 = .19, 90\% \text{ CI } [.12, .27], BF_{10} > 1000^8$, indicating that the effect of validity on self-report stereotyping, $F(1, 192) = 407.09, p < .001, \eta_p^2 = .68$, was stronger than the effect on automatic stereotyping, $F(1, 192) = 34.42, p < .001, \eta_p^2 = .15$.⁹

Discussion

Experiment 1 showed that automatic stereotyping was sensitive to validity information. Automatic stereotyping scores, reflecting stereotypes in line with the initial information provided, were larger in the control condition (valid) than in the mixed-up (invalid) condition. Although we did not observe a full reversal in the mixed-up condition, the validity of information still exerted a clear influence over responding, leading to effects not significantly different from zero when the pairings of the groups and the traits were described as invalid. Interestingly, a direct comparison revealed a larger effect of validity on self-reported than on automatic stereotyping, although the fixed order of the two measures limits the interpretation of these results. In Experiment 2, we moved to test another prediction of the propositional perspective: that automatic stereotyping is sensitive to relational information.

Experiment 2

In order to test if automatic stereotyping was sensitive to relational information, in Experiment 2 we used a manipulation that required participants to use information about the specific relation between the two groups (Niffites and Luupites) and traits (warm and competent) to form accurate stereotypes about them. Specifically, instead of describing the traits of

the groups in absolute terms (e.g., Niffites are helpful), the traits of Niffites and Luupites were described in relational terms (e.g., Niffites are more helpful and less confident than Lisa, who is a warm and competent person). We then informed half of the participants that, accidentally, the words “more than” and “less than” were mixed-up. This manipulation resembles the validity manipulation of Experiment 1. However, in order to correctly update the stereotypes of the group based on the new information, participants would need to account for the initial relational information. Therefore, an indication of an update would imply that the representation of the groups included relational information, and would therefore provide evidence that automatic stereotyping is sensitive to relational information.

Method

Participants. The preregistration of Experiment 2 is available at osf.io/rczwn/. 270 participants completed the study. The preregistered data exclusions involved removing participants who (a) did not fully complete all questions and tasks (1.5%), (b) had more than 10% trials in the IAT faster than 300ms (5.5%), or (c) made at least one error on the comprehension questions (23.8%, see Table 1 for exclusion rate as a function of experimental condition). The final sample included 191 participants (53% women, $M_{age} = 29.51, SD = 9.47$).

Materials, procedure and design. The materials, procedure and design were similar to those of Experiment 1, except for the following changes. First, unlike Experiment 1 in which the traits of the groups were described in absolute terms, the traits of Niffites and Luupites were described in relational terms. Specifically, in one condition, Niffites were described as warmer and less competent than Lisa and Luupites were described as less warm and more competent than Lisa. In the second condition, Niffites were described as less warm and more competent than Lisa and Luupites are warmer and less competent than Lisa (see Table 6). All participants were informed, before the exposure to the descriptions, that Lisa is generally considered to be a warm and competent person. In two blocks, we presented eight descriptions of each group,

⁸ For analyses comparing effects on self-report and automatic stereotyping measures, in all the experiments, to avoid complexity, the Bayes-Factors were computed using a model that excluded the counterbalanced procedural factors.

⁹ The measure type x validity interaction was significant also in the sample without exclusions based on the comprehension questions,

$F(1, 235) = 30.20, p < .001, \eta_p^2 = .11$. This interaction indicated that the effect of validity on self-report stereotyping, $F(1, 235) = 154.46, p < .001, \eta_p^2 = .40$, was stronger than the effect on automatic stereotyping, $F(1, 235) = 21.24, p < .001, \eta_p^2 = .08$

one at a time for 4s, in random order (16 descriptions in each block).¹⁰ Second, instead of the validity manipulation, we used a relational information manipulation. Specifically, after the exposure to the descriptions, participants in the opposition condition were informed that Lisa had mixed up the words “less than” and “more than” in her descriptions of the Niffites and Luupites. In other words, all the traits that Niffites were described as having less than Lisa, they actually have more than Lisa. All the traits that Niffites were described as having more than Lisa, they actually have less than Lisa. All the traits that Luupites were described as having less than Lisa, they actually have more than Lisa. And, all the traits that Luupites were described as having more than Lisa, they actually have less than Lisa. Participants in the control condition did not read this information. Third, we counterbalanced the order of the automatic and self-reported stereotyping measures, as well as the order of the automatic and self-reported evaluation measures (the stereotyping measures were always administered first). Fourth, opposite to Experiment 1, to keep the structure of the descriptions similar, the trait “intelligent” was used in the learning (“more intelligent” instead of “more smart”) and the trait “smart” was used in the IAT. Fifth, because the response options of the comprehension measure in Experiment 1 did not fit the descriptions used in Experiment 2, we used different response options (i.e., “They are warm”, “They are competent”, “They are tall” and “I don’t remember”). Finally, to simplify the design of the experiment, for the evaluation IAT we used only one block order (Luupites-Positive/Niffites-Negative first). Automatic and self-reported stereotyping scores were calculated such that positive scores indicated stereotypes in line with the initial information (IAT: $r = .65$, 95% CI [.56, .72], ratings: $\alpha = .94$). Automatic and self-reported evaluation scores were computed to reflect a preference for

Niffites over Luupites (IAT: $r = .65$, 95% CI [.57, .72], ratings: $\alpha = .88$).

Results

For both the IAT and self-reported stereotyping, we tested if the stereotyping scores were predicted by the relational information condition (opposition vs. control). For the IAT analysis, we controlled for group-trait assignment, stereotyping measures order, stereotyping IAT block order, and their interactions (with each other and with the relational information condition) and for IAT evaluation scores.¹¹ The full results of the analysis are detailed in Table 4 (see the end of the manuscript). Importantly, the IAT stereotyping scores were predicted by the relational information condition, $F(1, 174) = 29.26$, $p < .001$, $\eta_p^2 = .14$, 90% CI [.07, .21], $BF_{10} > 1000$. As predicted, the effect reflected a larger stereotyping score (reflecting stereotypes in line with the initial information) in the control condition ($M = 0.35$, $SD = 0.43$, difference from zero: $t(93) = 7.78$, $p < .001$, *Cohen’s d* = 0.80, $BF_{10} > 1000$) than in the opposition condition ($M = -0.01$, $SD = 0.47$, difference from zero: $t(96) = -0.27$, $p = .79$, *Cohen’s d* = 0.03, $BF_{10} = 0.12$, see Figure 2, panel A). For the self-reported stereotyping analysis, we controlled for group-trait assignment, stereotyping measures order, their interactions, and for the self-reported evaluation of the groups. The self-reported stereotyping scores were also predicted by the relational information condition, $F(1, 182) = 480.74$, $p < .001$, $\eta_p^2 = .73$, 90% CI [.66, .76], $BF_{10} > 1000$. As with the IAT, self-reported stereotyping scores were larger in the control condition ($M = 2.35$, $SD = 0.96$ difference from zero: $t(93) = 23.72$, $p < .001$, *Cohen’s d* = 2.45, $BF_{10} > 1000$) than in the opposition condition ($M = -1.81$, $SD = 1.55$ difference from zero: $t(96) = -11.53$, $p < .001$, *Cohen’s d* = 1.17, $BF_{10} > 1000$, see Figure 2, panel B).¹²

¹⁰ We conducted an additional study prior to this experiment which tested the effect of these descriptions on automatic and self-reported stereotyping, and found that both measures were sensitive to the information. That is, a stereotyping score indicating a “Niffites are warm/Luupites are competent” stereotype was larger when Niffites were described as warmer and less competent than Lisa than when Luupites were described as warmer and less competent than Lisa. A full report of this study can be found in the online supplement (osf.io/6stnx/).

¹¹ In the preregistration of Experiment 2, we planned to control only for the main effects of all the factors. However, following a reviewer’s recommendation, we report the results of an ANOVA that include, besides the relational information condition the group-trait assignment, stereotyping measures order, stereotyping IAT block order (only for the IAT analysis), and their interactions with each

other and with the relational information condition, and the IAT/self-report evaluation score as a covariate.

¹² The results without exclusions based on the comprehension questions revealed a similar pattern of results. Specifically, the IAT stereotyping scores were predicted by relational information condition, $F(1, 234) = 12.60$, $p < .001$, $\eta_p^2 = .05$. The effect reflected a larger IAT stereotyping score in the control condition ($M = 0.26$, $SD = 0.46$, difference from zero: $t(121) = 6.22$, $p < .001$, *Cohen’s d* = 0.56) than in the opposition condition ($M = 0.04$, $SD = 0.48$, difference from zero: $t(129) = 0.99$, $p = .32$, *Cohen’s d* = 0.13). The self-reported stereotyping scores were predicted by validity condition, $F(1, 242) = 168.05$, $p < .001$, $\eta_p^2 = .41$. The effect reflected a larger self-reported stereotyping score in the control condition ($M =$

Table 6*Descriptions presented in Experiment 2*

More warm (less competent) group	More competent (less warm) group
Niffites[Luupites] are more friendly than Lisa	Luupites[Niffites] are less friendly than Lisa
Niffites[Luupites] are more helpful than Lisa	Luupites[Niffites] are less helpful than Lisa
Niffites[Luupites] are more kind than Lisa	Luupites[Niffites] are less kind than Lisa
Niffites[Luupites] are more good-natured than Lisa	Luupites[Niffites] are less good-natured than Lisa
Niffites[Luupites] are less intelligent than Lisa	Luupites[Niffites] are more intelligent than Lisa
Niffites[Luupites] are less scientific than Lisa	Luupites[Niffites] are more scientific than Lisa
Niffites[Luupites] are less confident than Lisa	Luupites[Niffites] are more confident than Lisa
Niffites[Luupites] are less determined than Lisa	Luupites[Niffites] are more determined than Lisa

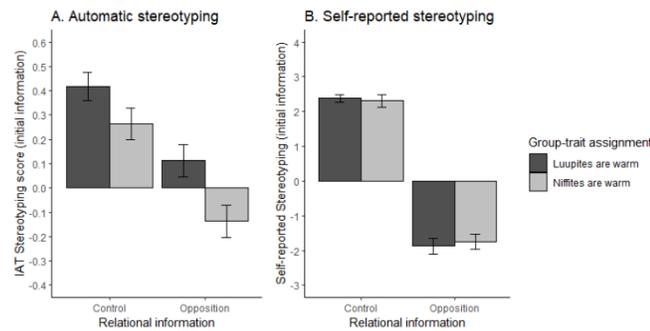


Figure 2. Experiment 2: Stereotyping scores (reflecting stereotypes in line with the initial information) as a function of relational information condition (control versus opposition) and group-trait assignment (Niffites are warm versus Luupites are warm). Graph A presents the result for automatic stereotyping (measured with the IAT). Graph B presents the result for self-reported stereotyping. Error-bars represent standard errors.

Direct comparison between self-reported and IAT stereotyping effects. Standardized scores were computed as in Experiment 1. We submitted these standardized scores to a 2 (*Measure type*) x 2 (*relational information condition*) x 2 (*group-trait assignment*)

1.67, $SD = 1.76$ difference from zero: $t(121) = 10.51, p < .001$, *Cohen's d* = 0.95) than in the mixed-up condition ($M = -1.36, SD = 1.87$ difference from zero: $t(128) = -8.23, p < .001$, *Cohen's d* = -0.72).

¹³ The measure type x validity interaction was significant also in the sample without exclusions based on the comprehension questions, $F(1, 235) = 33.73, p < .001, \eta_p^2 = .13$, indicating that the effect of

ment) x 2 (*stereotyping measures order*) x 2 (*stereotyping IAT block order*) mixed ANOVA with measure type as the only within-participants factor. Table 5 (see the end of the manuscript) summarize the results of the ANOVA. Importantly, the interaction between relational information and measure type was significant, $(1, 175) = 38.26, p < .001, \eta_p^2 = .18$, 90% CI [.10, .26], $BF_{10} > 1000$, indicating that the effect of relational information on self-report stereotyping, $F(1, 175) = 450.68, p < .001, \eta_p^2 = .72$, was stronger than the effect on automatic stereotyping, $F(1, 175) = 29.50, p < .001, \eta_p^2 = .14$.¹³

Discussion

The results of Experiment 2 exhibited an impact of relational information on automatic stereotyping. Automatic stereotyping scores, reflecting stereotypes in line with the initial information provided to participants, were larger in the control condition than in the opposition condition. As with Experiment 1, we did not observe a full reversal in the opposition condition, but relational information still exerted a clear influence over responding, leading to IAT effects not significantly different from zero when participants learned that the words “less than” and “more than” had been mixed up. A direct comparison between the two stereotyping measures, which were counterbalanced in Experiment 2, revealed a greater effect of validity on self-reported than on automatic stereotyping.

General Discussion

In two experiments, we tested predictions derived from a propositional perspective on automatic stereotyping. We tested if automatic stereotyping was sensitive to two components that are characteristic of propositional representations: truth value (validity) and relational information. Experiment 1 found that automatic (and self-reported) stereotyping of novel groups was influenced by validity information. Two novel groups were paired with warm or competent traits. Then, some participants learned that the pairings did not provide valid information for inferring that the groups possess the traits that were paired with them.

validity on self-report stereotyping, $F(1, 235) = 166.43, p < .001, \eta_p^2 = .41$, was stronger than the effect on automatic stereotyping, $F(1, 235) = 12.66, p < .001, \eta_p^2 = .05$.

The results revealed less stereotyping in line with the pairing when the pairing was described as invalid. In addition, the validity manipulation had a stronger effect on self-reported than on automatic stereotyping.

Experiment 2 found that automatic (and self-reported) stereotyping was also sensitive to relational information. In this experiment, we described the two groups in relational terms. One group was described as warmer and less competent than Lisa (a very warm and competent person). The other group was described as less warm and more competent than Lisa. Then, we added a manipulation that required participants to use the encoded relational information about the tribes to update the stereotypes about them. Specifically, we informed half of the participants that the words “more than” and “less than” were mixed-up. The results revealed less stereotyping in line with the initial information when the words “more than” and “less than” were mixed-up. As in Experiment 1, the relational information manipulation had a stronger effect on self-reported than on automatic stereotyping.

Importantly, in all experiments, we controlled for evaluation by (a) using only positive traits in our manipulations and measures of stereotyping, and (b) measuring the evaluations of the groups and controlling for their influence in the analyses. Because of this, we can safely argue that the present results speak to automatic *stereotyping* rather than automatic *evaluation*.

Implications for Research on Automatic Stereotyping

The present research results have three levels of implications: adding new knowledge about moderators of automatic stereotyping, providing constraints on cognitive theories of automatic stereotyping, and providing new ideas for applied interventions used to fight the negative consequences of automatic stereotyping. First, the present results contribute to our understanding of factors that determine how and when exposure to information affects automatic stereotyping. Overall, our results indicate that automatic stereotyping takes into account the validity and relational nature of information that pairs social groups and traits. In both experiments the invalid/opposition relation conditions did not lead to a full reversal in stereotyping, but rather led to the absence of statistical difference of the stereotyping effects from zero. The lack of full reversal is compatible with what had been usually found for the effects of manipulations of validity and relational information on automatic evaluation

(e.g., Moran et al., 2015; Moran & Bar-Anan, 2020; Wyer, 2016; Zanon et al., 2012). Despite the lack of full reversal, the results from both of our experiments indicated that automatic stereotyping can be impacted by relational and validity information.

Second, although the present research was not designed to prove that a specific model of automatic stereotyping is correct, our results can help to constrain cognitive theories of automatic stereotyping. Research on automatic stereotyping has flourished in the last two decades (e.g., Blair et al., 2001; Nosek et al., 2009). However, the vast majority of this research has relied on an associative theoretical perspective. In the present research, we adopted an alternative propositional perspective. Importantly, a propositional perspective of automatic stereotyping has merit because it allows researchers to conceptualize new questions about the fundamental nature of automatic stereotyping (see Kuhn, 1970, for a treatment on how theories may constrain the questions scientists ask). In our paper, we focused on questions about validity and relational information. Our results demonstrate that validity and relational information can in fact have a substantial impact of automatic stereotyping.

It is important to note that none of the existing models for automatic stereotyping can easily explain the whole set of the current research findings. The propositional perspective clearly predicts the sensitivity of automatic stereotyping to validity and relational information. However, it is less clear what the propositional perspective predicts with regard to the relative impact of that information on automatic versus non-automatic stereotyping. It can provide a post-hoc explanation for the findings that validity and relational information had a weaker influence on automatic than on non-automatic stereotyping by assuming that a) both automatic and self-reported stereotyping are influenced by propositions, and b) under sub-optimal conditions (i.e., automatic stereotyping) there is less control regarding which propositions influence stereotyping. For example, the learning in Experiment 1 could have led to the formation of the propositions “Niffites were paired with warm traits” and “Niffites are really competent”. When people have ample time and opportunity to consider which proposition to use in their response, it is plausible to assume that they will give more weight to the latter proposition. However, when processing conditions are suboptimal there is less control of which of the propositions will influence performance. Importantly, however, the propositional perspective does not provide clear predictions

regarding the exact conditions under which different propositions will influence automatic stereotyping.

Similarly, models that rely on an associative theoretical perspective also cannot fully explain the current set of results. For example, the Associative-Propositional Evaluation (APE) model (Gawronski & Bodenhausen, 2006) assumes that automatic evaluation is mainly influenced by the activation of associations between an object and valence, regardless of the validity of these associations. However, this model also assumes that validity can sometimes have an indirect effect on automatic evaluation when the inference from the validity information affects self-reported evaluation, which may then lead to the creation of new associations that reflect the inferred valence. Applying the assumptions of the APE model to automatic stereotyping, the results of Experiment 1 (a lack of full reversal in the invalid condition for automatic stereotyping, and a full reversal for self-reported stereotyping) can be interpreted as evidence for an indirect effect of validity information on automatic stereotyping. Importantly, however, the APE model does not provide clear predictions regarding the exact conditions under which validity will have an indirect effect on automatic stereotyping.

The full set of findings of the current research not only raise questions about current accounts of automatic stereotyping but also inspire future research and theories on this topic. Any future model of stereotyping should be able to account for (a) why automatic stereotyping is sensitive to relational and validity information, (b) why automatic stereotyping in some cases is not completely corrected by relational and validity information, and (c) why validity and relational information have a stronger effect on self-reported than on automatic stereotyping.

Third, we believe that adopting a propositional perspective of automatic stereotyping could help to improve current interventions used to fight the negative consequences of automatic stereotyping. Most current interventions in this regard are conceptualized from an associative perspective (e.g., Dasgupta & Asgari, 2004; Johnson et al., 2018). However, the propositional perspective uniquely predicts new and important ways in which we can maximize the effectiveness of these interventions. For example, a popular intervention to reduce automatic stereotyping is exposure to a counter-stereotypic individual (e.g., a successful female scientist; Dasgupta & Asgari, 2004; Stout et al., 2011). The idea is that this exposure creates new associations linking the target group with the

counter-stereotypic attributes. However, a propositional perspective can inform us that simply presenting such counter-stereotypic individual may not necessarily be sufficient to achieve this change: factors like validity (e.g., the extent to which one believes that the counter-stereotypic example is valid) and relational information will have a critical impact on what effect this presentation might have. These factors should therefore be considered in-depth in order to optimize the effectiveness of such an intervention.

One limitation of the present research is that we used a limited set of stimuli (i.e., fictitious social groups as targets) and traits (i.e., warm and competent), and a specific sample (Prolific participants). Future research should extend the present research by examining the sensitivity of automatic stereotyping to validity and relational information with different stimuli, contexts, and samples. Moreover, the present research used only one type of automatic stereotyping measure (the IAT). It is possible that the observed sensitivity of automatic stereotyping to validity and relational information is due to specific properties of the IAT and does not transfer to other automatic stereotyping measures. Future research could extend this line of research by using different automatic measures such as the Semantic Misattribution Procedure (SMP; Ye & Gawronski, 2018).

In addition to extending the present research results, future research could also use the assumptions of the propositional perspective of automatic stereotyping to generate new research questions. For example, future research could test if, as in the case of automatic evaluation, verbal instructions are sufficient to form and change automatic stereotyping (Kurdi & Banaji, 2017), and if automatic stereotyping is also sensitive to diagnosticity (Cone & Ferguson, 2015) and believability (Cone et al., 2019). Finally, as we already highlighted, the present research was not designed to prove that a specific model of automatic stereotyping is correct. Future research could be designed to directly compare competing predictions of different theoretical models. Overall, the present research highlights the value of considering a propositional perspective in the context of automatic stereotyping.

References

- Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*(4), 652-661.

- Amodio, D. M., & Ratner, K. G. (2011). A memory systems model of implicit social cognition. *Current Directions in Psychological Science*, 20(3), 143-148.
- Andreychik, M. R., & Gill, M. J. (2012). Do negative implicit associations indicate negative attitudes? Social explanations moderate whether ostensible "negative" associations are prejudice-based or empathy-based. *Journal of Experimental Social Psychology*, 48(5), 1082-1093.
- Ashmore, R. D., & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 1-35). Hillsdale, NJ: Erlbaum.
- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: the moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, 81(5), 828-841.
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, public policy, and law*, 5(3), 665-692.
- Carlsson, R., & Björklund, F. (2010). Implicit stereotype content: Mixed stereotypes can be measured with the implicit association test. *Social Psychology*, 41(4), 213-222.
- Cone, J., & Ferguson, M. J. (2015). He did what? The role of diagnosticity in revising implicit evaluations. *Journal of Personality and Social Psychology*, 108(1), 37-57.
- Cone, J., Flaharty, K., & Ferguson, M. J. (2019). Believability of evidence matters for correcting social impressions. *Proceedings of the National Academy of Sciences*, 116(20), 9802-9807.
- Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology*, 83(5), 1076-1094.
- Cummins, J., & De Houwer, J. (2019). An inkblot for beliefs: the truth misattribution procedure. *PloS One*, 14(6): e0218661.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, 15(12), 806-813.
- Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, 40(5), 642-658.
- De Houwer, J. (2014). A Propositional Model of Implicit Evaluation. *Social and Personality Psychology Compass*, 8, 342-353.
- De Houwer, J. (2019). Implicit Bias Is Behavior: A Functional-Cognitive Perspective on Implicit Bias. *Perspectives on Psychological Science*, 14(5), 835-840.
- De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin and Review*, 20(4), 631-642.
- De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013). A functional-cognitive framework for attitude research. *European Review of Social Psychology*, 24(1), 252-287.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, 135(3), 347-368.
- De Houwer, J., Van Dessel, P., & Moran, T. (2020). Attitudes beyond associations: On the role of propositional representations in stimulus evaluation. *Advances in Experimental Social Psychology*, 61, 127-183.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5-18.
- Dotsch, R. O. N., Wigboldus, D. H., & Van Knippenberg, A. D. (2013). Behavioral information biases the expected facial appearance of members of novel groups. *European Journal of Social Psychology*, 43(1), 116-125.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. *The Handbook of Social Psychology*, 2(4), 357-411.
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77-83.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902.
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692-731.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 109(1), 3-25.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464-1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197-216.
- Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, 90(1), 1-20.
- Hilton, J. L., & Von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47(1), 237-271.
- Jabold, P. (2020). Stereotype Formation but No Dissociation: Contradicting Statistics Reduce Explicit and Implicit Stereotypes Engendered by Disproportional Crime Reporting. *Social Psychological and Personality Science*, 11(3), 364-373.

- Johnson, I. R., Kopp, B. M., & Petty, R. E. (2018). Just say no!(and mean it): Meaningful negation as a tool to modify automatic racial attitudes. *Group Processes & Inter-group Relations*, 21(1), 88-110.
- Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. *Advances in Experimental Social Psychology*, 55, 1-80.
- Koch, A. J., D'Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology*, 100(1), 128-161.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed., enlarged). Chicago, IL: University of Chicago Press.
- Kurdi, B., & Banaji, M. R. (2017). Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes?. *Journal of Experimental Psychology*, 146(2), 194-213.
- Kurdi, B., Mann, T. C., Charlesworth, T. E., & Banaji, M. R. (2019). The relationship between implicit intergroup attitudes and beliefs. *Proceedings of the National Academy of Sciences*, 116(13), 5862-5871.
- Martin, D., Hutchison, J., Slessor, G., Urquhart, J., Cunningham, S. J., & Smith, K. (2014). The spontaneous formation of stereotypes via cumulative cultural evolution. *Psychological Science*, 25(9), 1777-1786.
- Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology*, 67, 263-287.
- Moran, T., & Bar-Anan, Y. (2020). The effect of co-occurrence and relational information on speeded evaluation. *Cognition and Emotion*, 34(1), 144-155.
- Moran, T., Bar-Anan, Y., & Nosek, B. A. (2015). Processing goals moderate the effect of co-occurrence on automatic evaluation. *Journal of Experimental Social Psychology*, 60, 157-162.
- Moran, T., Bar-Anan, Y., & Nosek, B. A. (2017). The effect of the validity of co-occurrence on automatic and deliberate evaluations. *European Journal of Social Psychology*, 47(6), 708-723.
- Moskowitz, G. B., Stone, J., & Childs, A. (2012). Implicit stereotyping and medical decisions: unconscious stereotype activation in practitioners' thoughts about African Americans. *American Journal of Public Health*, 102(5), 996-1001.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, 31(2), 166-180.
- Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., ... & Kesebir, S. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, 106(26), 10593-10597.
- Peters, K. R., & Gawronski, B. (2011). Are we puppets on a string? Comparing the impact of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin*, 37(4), 557-569.
- Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The Meta-Cognitive Model (MCM) of attitudes: Implications for attitude measurement, change, and strength. *Social Cognition*, 25(5), 657-686.
- Phills, C. E., Hahn, A., & Gawronski, B. (2020). The bidirectional causal relation between implicit stereotypes and implicit prejudice. *Personality and Social Psychology Bulletin*, 46(9), 1318-1330
- Rubinstein, R. S., Jussim, L., & Stevens, S. T. (2018). Reliance on individuating information and stereotypes in implicit and explicit person perception. *Journal of Experimental Social Psychology*, 75, 54-70.
- Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men?. *Journal of Personality and Social Psychology*, 87(4), 494-509.
- Sabin, J. A., Rivara, F. P., & Greenwald, A. G. (2008). Physician implicit attitudes and stereotypes about race and quality of medical care. *Medical Care*, 678-685.
- Sherman, J. W. (1996). Development and mental representation of stereotypes. *Journal of Personality and Social Psychology*, 70, 1126-1141.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *Dialogue: The Official Newsletter of the Society for Personality and Social Psychology*, 26(2), 4-7.
- Smith, E. R., & Zarate, M. A. (1990). Exemplar and prototype use in social categorization. *Social Cognition*, 8(3), 243-262.
- Spencer, K. B., Charbonneau, A. K., & Glaser, J. (2016). Implicit bias and policing. *Social and Personality Psychology Compass*, 10(1), 50-63.
- Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, 100, 255-270.
- Wyer, N. A. (2016). Easier done than undone... by some of the people, some of the time: The role of elaboration in explicit and implicit group preferences. *Journal of Experimental Social Psychology*, 63, 77-85.
- Ye, Y., & Gawronski, B. (2018). Validating the semantic misattribution procedure as an implicit measure of gender stereotyping. *European Journal of Social Psychology*, 48(3), 348-364.
- Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning and Motivation*, 43(3), 155-165.

Table 4
 Full ANOVA results for IAT and self-reported stereotyping scores in Experiments 1 and 2

Experiment 1		
Factors	Results -IAT	Results – self-report
Validity	$F(1, 191) = 34.35, p < .001, \eta_p^2 = .15^3$	$F(1, 195) = 409.33, p < .001, \eta_p^2 = .68^3$
Group-trait assignment	$F(1, 191) = 9.92, p = .002, \eta_p^2 = .05^4$	$F(1, 195) = 2.78, p = .10, \eta_p^2 = .01$
sIAT block order¹	$F(1, 191) = 0.02, p = .88, \eta_p^2 < .01$	
IAT/self-report evaluation score	$F(1, 191) = 0.38, p = .54, \eta_p^2 < .01$	$F(1, 195) = 2.28, p = .13, \eta_p^2 = .01$
Validity × Group-trait assignment	$F(1, 191) = 0.01, p = .94, \eta_p^2 < .01$	$F(1, 195) = 0.29, p = .59, \eta_p^2 < .01$
Validity × sIAT block order	$F(1, 191) = 2.73, p = .10, \eta_p^2 = .01$	
Group-trait assignment × sIAT block order	$F(1, 191) = 0.42, p = .52, \eta_p^2 < .01$	
Validity × Group-trait assignment × sIAT block order	$F(1, 191) = 3.59, p = .06, \eta_p^2 = .02$	
Experiment 2		
Relational information	$F(1, 174) = 29.26, p < .001, \eta_p^2 = .14^5$	$F(1, 182) = 480.74, p < .001, \eta_p^2 = .73^5$
Group-trait assignment	$F(1, 174) = 7.96, p = .005, \eta_p^2 = .04^6$	$F(1, 182) = 0.04, p = .83, \eta_p^2 < .01$
sIAT block order	$F(1, 174) = 0.09, p = .76, \eta_p^2 < .01$	
Measures order²	$F(1, 174) = 0.75, p = .39, \eta_p^2 < .01$	$F(1, 182) = 1.09, p = .30, \eta_p^2 < .01$
IAT/self-report evaluation score	$F(1, 174) = 0.24, p = .63, \eta_p^2 < .01$	$F(1, 182) = 0.25, p = .62, \eta_p^2 < .01$
Relational information × Group-trait assignment	$F(1, 174) = 0.39, p = .53, \eta_p^2 < .01$	$F(1, 182) = 0.70, p = .41, \eta_p^2 < .01$
Relational information × sIAT block order	$F(1, 174) = 4.48, p = .04, \eta_p^2 = .03^7$	
Relational information × Measures order	$F(1, 174) = 0.01, p = .92, \eta_p^2 < .01$	$F(1, 182) = 7.25, p = .008, \eta_p^2 = .04^8$
Group-trait assignment × sIAT block order	$F(1, 174) = 1.63, p = .20, \eta_p^2 < .01$	
Group-trait assignment × Measures order	$F(1, 174) = 1.66, p = .20, \eta_p^2 < .01$	$F(1, 182) = 0.30, p = .59, \eta_p^2 < .01$
sIAT block order × Measures order	$F(1, 174) = 0.25, p = .62, \eta_p^2 < .01$	
Relational information × Group-trait assignment × sIAT block order	$F(1, 174) = 0.57, p = .45, \eta_p^2 < .01$	
Relational information × Group-trait assignment × Measures order	$F(1, 174) = 0.98, p = .32, \eta_p^2 < .01$	$F(1, 182) = 0.07, p = .78, \eta_p^2 < .01$
Relational information × sIAT block order × Measures order	$F(1, 174) = 0.07, p = .79, \eta_p^2 < .01$	
Group-trait assignment × sIAT block order × Measures order	$F(1, 174) = 0.02, p = .89, \eta_p^2 < .01$	
Relational information × Group-trait assignment × sIAT block order × Measures order	$F(1, 174) = 0.36, p = .55, \eta_p^2 < .01$	

Note. ¹sIAT block order = stereotyping IAT block order. ²Measures order = stereotyping measure order. ³The main effect of validity condition (Experiment 1) indicated a larger IAT/self-reported stereotyping score (reflecting stereotypes in line with the initial information) in the control condition than in the mixed-up condition. ⁴The main effect of group-trait assignment (Experiment 1) indicated a larger IAT stereotyping score when Luupites were initially described as warm than when Niffites were initially described as warm. ⁵The main effect of relational information condition (Experiment 2) indicated a larger IAT/self-reported stereotyping score (reflecting stereotypes in line with the initial information) in the control condition than in the opposition condition. ⁶The main effect of group-trait assignment (Experiment 2) indicated a larger IAT stereotyping score when Luupites were initially described as warm than when Niffites were initially described as warm. ⁷The relational information × sIAT block order interaction (Experiment 2) indicated a stronger effect of relational information when Niffites first shared response key with “warm” than when Luupites first shared response key with “warm”. ⁸The relational information × stereotyping measures order interaction (Experiment 2) indicated a stronger effect of relational information when the IAT was first than when the self-report measure was first.

Table 5

Full ANOVA results for the direct comparison of IAT and self-reported stereotyping scores in Experiment 1 and 2

Factors	Results: Experiment 1	Results: Experiment 2
Measure type	$F(1, 192) = 11.58, p = .001, \eta_p^2 = .06^4$	$F(1, 175) = 7.15, p = .008, \eta_p^2 = .04^4$
Condition ¹	$F(1, 192) = 230.61, p < .001, \eta_p^2 = .55^5$	$F(1, 175) = 232.18, p < .001, \eta_p^2 = .57^8$
Group-trait assignment	$F(1, 192) = 0.72, p = .40, \eta_p^2 < .01$	$F(1, 175) = 0.32, p = .57, \eta_p^2 < .01$
sIAT block order ²	$F(1, 192) = 0.16, p = .69, \eta_p^2 < .01$	$F(1, 175) = 0.03, p = .86, \eta_p^2 < .01$
Measures order ³		$F(1, 175) = 2.10, p = .15, \eta_p^2 = .01$
Measure type × Condition	$F(1, 192) = 46.02, p < .001, \eta_p^2 = .19^6$	$F(1, 175) = 38.26, p < .001, \eta_p^2 = .18^9$
Measure type × Group-trait assignment	$F(1, 192) = 0.26, p = .61, \eta_p^2 < .01$	$F(1, 175) = 0.05, p = .82, \eta_p^2 < .01$
Measure type × sIAT block order	$F(1, 192) = 0.05, p = .82, \eta_p^2 < .01$	$F(1, 175) = 0.57, p = .45, \eta_p^2 < .01$
Measure type × Measures order		$F(1, 175) = 0.00, p = .99, \eta_p^2 < .01$
Condition × Group-trait assignment	$F(1, 192) = 0.87, p = .35, \eta_p^2 < .01$	$F(1, 175) = 0.01, p = .93, \eta_p^2 < .01$
Condition × sIAT block order	$F(1, 192) = 2.74, p = .10, \eta_p^2 = .01$	$F(1, 175) = 5.30, p = .022, \eta_p^2 = .03^{10}$
Condition × Measures order		$F(1, 175) = 2.28, p = .13, \eta_p^2 = .01$
Group-trait assignment × sIAT block order	$F(1, 192) = 1.45, p = .23, \eta_p^2 = .01$	$F(1, 175) = 3.43, p = .066, \eta_p^2 = .02$
Group-trait assignment × Measures order		$F(1, 175) = 2.25, p = .14, \eta_p^2 = .01$
Stereotyping IAT block order × Measures order		$F(1, 175) = 0.09, p = .77, \eta_p^2 < .01$
Measure type × Condition × Group-trait assignment	$F(1, 192) = 0.52, p = .47, \eta_p^2 < .01$	$F(1, 175) = 0.97, p = .33, \eta_p^2 = .01$
Measure type × Condition × sIAT block order	$F(1, 192) = 1.54, p = .22, \eta_p^2 = .01$	$F(1, 175) = 1.97, p = .16, \eta_p^2 = .01$
Measure type × Condition × Measures order		$F(1, 175) = 2.15, p = .14, \eta_p^2 = .01$
Measure type × Group-trait assignment × sIAT block order	$F(1, 192) = 0.03, p = .86, \eta_p^2 < .01$	$F(1, 175) = 0.10, p = .76, \eta_p^2 < .01$
Measure type × Group-trait assignment × Measures order		$F(1, 175) = 0.37, p = .54, \eta_p^2 < .01$
Measure type × sIAT block × Measures order		$F(1, 175) = 0.29, p = .59, \eta_p^2 < .01$
Condition × Group-trait assignment × sIAT block order	$F(1, 192) = 3.94, p = .049, \eta_p^2 = .02^7$	$F(1, 175) = 0.07, p = .79, \eta_p^2 < .01$
Condition × Group-trait assignment × Measures order		$F(1, 175) = 1.51, p = .22, \eta_p^2 = .01$
Condition × sIAT block order × Measures order		$F(1, 175) = 0.01, p = .92, \eta_p^2 < .01$
Group-trait assignment × sIAT block order × Measures order		$F(1, 175) = 0.35, p = .56, \eta_p^2 < .01$
Measure type × Condition × Group-trait assignment × sIAT block order	$F(1, 192) = 1.72, p = .19, \eta_p^2 < .01$	$F(1, 175) = 1.37, p = .24, \eta_p^2 = .01$
Measure type × Condition × Group-trait assignment × Measures order		$F(1, 175) = 0.29, p = .59, \eta_p^2 < .01$
Measure type × Condition × sIAT block order × Measures order		$F(1, 175) = 0.32, p = .57, \eta_p^2 < .01$
Measure type × Group-trait assignment × sIAT block order × Measures order		$F(1, 175) = 0.09, p = .77, \eta_p^2 < .01$
Condition × Group-trait assignment × sIAT block order × Measures order		$F(1, 175) = 0.02, p = .88, \eta_p^2 < .01$
Measure type × Condition × Group-trait assignment × sIAT block order × Measures order		$F(1, 175) = 1.34, p = .25, \eta_p^2 = .01$

Note. ¹Condition = validity information in Experiment 1, relational information in Experiment 2. ²sIAT block order = stereotyping IAT block order. ³Measures order = stereotyping measure order. ⁴The main effect of measure type (Experiments 1 and 2) indicated overall stronger stereotyping effects (reflecting stereotypes in line with the initial information) for automatic stereotyping measures than for self-reported stereotyping measures. ⁵The main effect of validity condition (Experiment 1) indicated a larger stereotyping score in the control condition than in the mixed-up condition. ⁶The interaction between measure type and validity condition indicated a larger effect of validity on self-report stereotyping than on automatic stereotyping. ⁷The Validity × Group-trait assignment × Stereotyping IAT block order three-way interaction was not further analyzed due to complexity and lack of theoretical interest. ⁸The main effect of relational information condition (Experiment 2) indicated a larger stereotyping score in the control condition than in the opposition condition. ⁹The interaction between measure type and relational information condition indicated a larger effect of relational information on self-report stereotyping than on automatic stereotyping. ¹⁰The relational information × Stereotyping IAT block order interaction was not further analyzed due to complexity and lack of theoretical interest.