Semi-Supervised Stereotyping: Untested Stereotypic Assumptions Reinforce Stereotypes as Much as Externally Confirmed Stereotypic Assumptions

William T. L. Cox and Patricia G. Devine

BACKGROUND

THE PARADOX

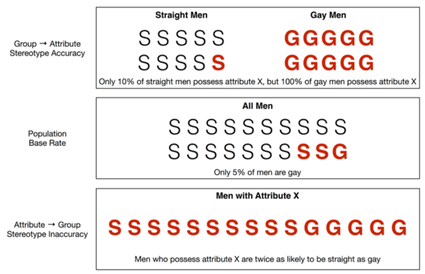
The human brain is remarkably proficient at learning the statistics of its environment.

Social cognition is heavily biased by stereotypes, which often do not reflect reality.

EXAMPLE

Cox, Devine, Bischmann, & Hyde (in-press, *J of Sex Research*)

Because the defining features of gay male group membership are non-visible, people often rely on stereotypic cues, like fashion or body movement, to make stereotypic assumptions about whether a man is gay or straight. Because gay men, however, are such a numerically small minority (~5% of the adult male population), it is nearly impossible for these stereotypic judgments to be pragmatically accurate. Even if the stereotypic attribute (e.g., fashionability, feminine body movement) were present in 100% of gay men, it would have to be present in fewer than 5% of straight men to yield more correct than incorrect conclusions. No such large between-group differences have ever been demonstrated. (See also Cox & Devine, in press, *PLOS ONE*.)



SEMI-SUPERVISED STEREOTYPING?

One possible solution to the persistence of stereotypes and stereotyping in the face of their inaccuracy comes from a concept in the machine learning literature, called *semi-supervised learning* (Gibson, Rogers, & Zhu, 2013; Kalish, Rogers, Lang, & Zhu, 2011).

Semi-supervised learning occurs when prior expectations, provided in this case by culturally-learned stereotypes, lead to judgments that are externally neither confirmed nor disconfirmed. The brain treats these untested assumptions as valid data, thereby reinforcing the preexisting associations. It seems to us that this scenario reflects how most stereotypic judgments typically occur — it is rare that a stereotypic judgment is truly put to the test one way or the other. And if the brain indeed treats untested stereotypic assumptions as though they were confirmed, it would be a powerful mechanism to explain how and why stereotypes can persist in the absence of evidence for them.

THE PRESENT WORK

To test this semi-supervised learning account of stereotyping, we set up a task in which participants made a series of stereotypic judgments and assessed how participants' reliance on stereotypes changed over time. Participants were randomly assigned to receive No Feedback, Stereotype-Confirming Feedback, or Stereotype-Disconfirming Feedback about their judgments.

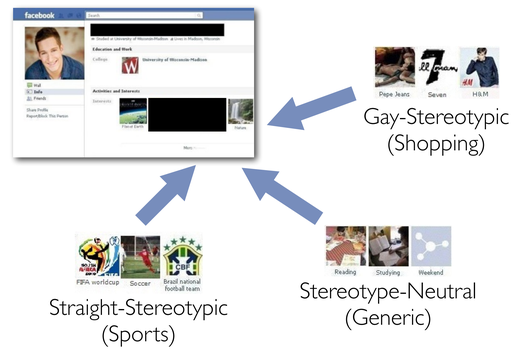
The Stereotype-Confirming and Stereotype-Disconfirming Feedback conditions each created a fabricated reality — one in which stereotypes were largely accurate, and one in which they were largely inaccurate. Participants in the No Feedback condition were presented with an experimental reality that could be consistent with either extreme — absent feedback, these participants had no objective evidence of whether the stereotypes were accurate or inaccurate. From a purely rational perspective (in which stereotypes exist because they are largely accurate), then, rates of stereotyping in the No Feedback condition should remain constant over time. If, however, participants' untested assumptions are processed and encoded as stereotype-supporting evidence, we would expect rates of stereotyping in the No Feedback condition to increase, just as in the Stereotype-Confirming Feedback condition.

METHOD

STEREOTYPING CATEGORIZATION TASK

Participants (N = 469) made gay-or-straight judgments about a series of fictitious men's Facebook profiles, which had been manipulated to be gay-stereotypic, straight-stereotypic, or stereotype-neutral.

(Before we added the Shopping, Sports, or Neutral interests, the profiles were extensively pretested to insure that they did not differ in the extent to which the men seemed gay or straight.)



Participants judged 30 profiles total, split into 20 Early (or Training) trials, and 10 Late (or Test) trials. On the Early trials, participants were randomly assigned to received one of three patterns of feedback.

Feedback Manipulation

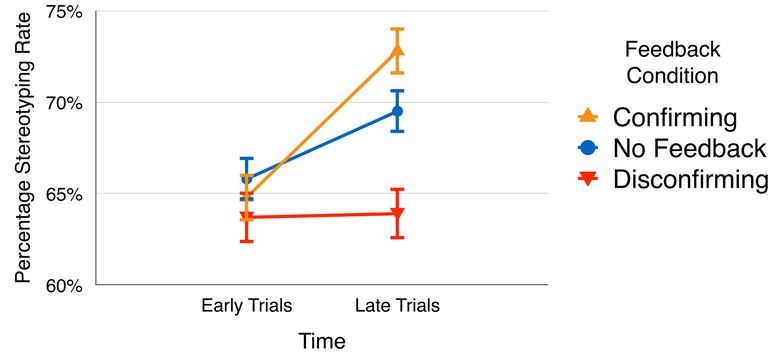
- **Stereotype-Confirming Feedback**: 40-60% of Early trials confirmed stereotypes (e.g., Sports Profiles belonged to Straight men, Shopping profiles belonged to Gay men), only 20% disconfirmed stereotypes, and 20-40% were neutral.
- **No Feedback**: No feedback was given, therefore no information was provided about the stereotypes or their accuracy.
- **Stereotype-Disconfirming Feedback**: 40-60% of Early trials disconfirmed stereotypes (e.g., Sports Profiles belonged to Gay men, Shopping profiles belonged to Straight men), only 20% confirmed stereotypes, and 20-40% were neutral.

The dependent variable was participants' rates of stereotyping, calculated as a percentage of profiles on which participants made stereotype-congruent judgments (i.e., the percentage of Sports profiles judged as Straight and Shopping Profiles judged as Gay). These stereotyping percentages were compared across the early and late trials to evaluate change over time.

RESULTS AND DISCUSSION

RESULTS

A 3 (Feedback Condition: No Feedback vs. Stereotype-Confirming Feedback vs. Stereotype-Disconfirming Feedback) x 2 (Time: Early Trials vs. Late Trials) mixed ANOVA revealed the predicted interaction, F (2, 466) = 5.010, p = 0.007.



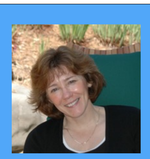
Matching hypotheses, participants who received No Feedback stereotyped at higher rates over time, increasing just as much as participants in the Stereotype-Confirming Feedback condition. This supports a semi-supervised learning account of stereotyping, in which untested stereotypic judgments are treated as valid data that confirm stereotypes.

Strikingly, the Stereotype-Disconfirming Feedback did not reduce stereotyping over time. This may indicate that the stereotype is equally affected by the initial activation (+1) and the disconfirmation (-1), such that they neutralize each other (cf. Devine, 1989).

GENERAL DISCUSSION

The present work supports the conclusion that untested stereotypic judgments are (consciously or non-consciously) treated as valid data supporting the stereotype and influence future reliance on the stereotypes just as much as when the judgments are externally, objectively verified.

Cumulatively, these findings highlight the difficulties inherent to changing preexisting associations. Although the present study explored this process in the context of stereotyping to infer orientation (see also Cox et al., in press, *J of Sex Research*), it seems likely that this mechanism should operate in most cases of stereotyping. Whether the stereotyping process relates to race, gender, or other features, stereotypic assumptions are most often left untested, and the present data indicate that these untested assumptions will nonetheless reinforce stereotypes and increase reliance on those stereotypes over time.



William T. L. Cox and Patricia G. Devine
Department of Psychology,
University of Wisconsin - Madison

will.cox@me.com
www.sciencecox.com
Twitter: @ScienceCox

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON