# Thinking about interracial interactions

William J Gehring, Andrew Karpinski & James L Hilton

**White people who have difficulty implicitly pairing black names with positive words also tend to be impaired on tasks requiring cognitive control after interacting with a black experimenter. A new functional imaging study finds that such subjects also show more activity in brain regions associated with cognitive control when looking at black faces that are irrelevant to their task.**

What are the thoughts and feelings of a person interacting with someone of a different race? The question is important because understanding how people act toward others depends on understanding their thoughts and emotions. It is difficult to answer, however, because individuals who view people of other races negatively are not likely to be honest in reporting such attitudes.

In this issue, Richeson and colleagues[1] test the hypothesis that white individuals with biased attitudes toward blacks attempt to control their thoughts and actions during an interracial interaction, presumably in order to appear non-prejudiced. People can exert 'cognitive control' to monitor and control information processing in the brain[2], but using cognitive control for one purpose may limit one's ability to use it for another purpose. (The authors[1] use the equivalent term 'executive control'.) According to this hypothesis, racially biased white people are likely to use cognitive control during an interaction with a black person to suppress the expression of stereotypes and negative attitudes, which depletes their ability to use cognitive control on subsequent tasks. The new study[1] is an ambitious example of social neuroscience, which seeks to use the techniques and theories of neuroscience to understand social behavior[3], by bringing together cognitive neuroscience, cognitive psychology and social psychology.

In the study[1], white participants first completed an implicit association test (IAT; **Fig. 1**), intended to measure their unconscious or automatic racial attitudes. Participants were not told the purpose of the test. Next, a black experimenter asked them to discuss the college fraternity system and racial profiling. Then participants performed a Stroop color-naming task (**Fig. 2**), in which participants must indicate the ink color in which color words (such as 'red' or 'green') are printed. This task requires a great deal of cognitive control[2]. When the word "red" is printed in green ink, participants automatically read the word, and they must override this initial reaction to make the correct response, "green."

Consistent with the authors' hypothesis, the more racial bias participants showed on the IAT before the interaction with the black experimenter, the greater were the decrements in their cognitive control capabilities afterward (measured as slower Stroop task performance). In a control experiment, participants who interacted with a white experimenter showed no relationship between IAT score and Stroop task performance, suggesting that this interaction required little cognitive control. Within two weeks, the same participants were recruited for a seemingly unrelated experiment, in which brain activity was measured by functional magnetic resonance imaging (fMRI) while they viewed photographs of black and white faces. Participants with high race-bias IAT scores showed more activity in brain areas associated with cognitive control (lateral prefrontal cortex, anterior cingulate cortex)[2,4] when presented with black faces.

One must be cautious, however, about claims that any measure provides a direct window into racially biased behavior. This caution is particularly warranted for the IAT: researchers in social psychology disagree about the meaning of IAT scores, yet much research is reported in academic journals and the popular press as if its validity as a measure of racial bias and prejudice were well-established.

The IAT attempts to measure the differential strength of associations between positive or negative words and the concepts of black

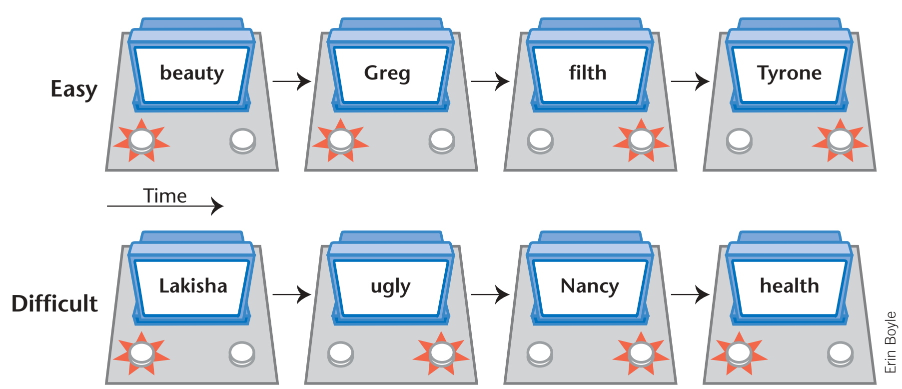


**Figure 1** The Implicit Association Test (IAT). Top panel: the participant presses the left button if the word is a white name or a positive item and the right button if the word is a black name or a negative item. Bottom panel: the participant presses the left button if the word is a black name or a positive word and the right button if the word is a white name or a negative word. The IAT effect is the difference in average reaction time between the two conditions.

William Gehring and James Hilton are in the Department of Psychology, University of Michigan, Ann Arbor, Michigan, USA. Andrew Karpinski is in the Department of Psychology, Temple University, Philadelphia, Pennsylvania, USA.
e-mail: wgehring@umich.edu

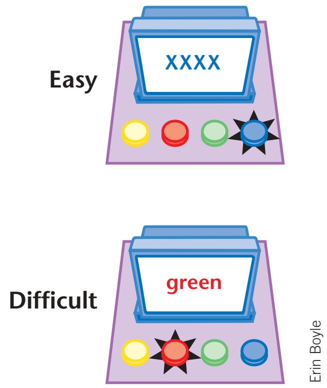


**Figure 2** The Stroop task. The participant presses the button corresponding to the color in which the letters on the screen are printed. In the top panel, the correct button is blue. In the bottom panel, the correct button is red. The Stroop effect is the difference in average reaction time between the two conditions.



and white (**Fig. 1**). Participants view positive words, negative words, black names and white names presented on a computer screen one at a time in a random order. The task is simply to categorize each word as positive, negative, black or white. In one stage, white names and positive words share one response button, while black names and negative words share the other response button. In the second stage, white names and negative words share one button, and black names and positive words share the other button. Consistent with previous IAT studies, Richeson and colleagues[1] found that, on average, white people were slower and less accurate when black was paired with positive (and white with negative) than when black was paired with negative (and white with positive). The authors used the difference in reaction time between these conditions—the 'IAT effect'—to measure racial bias, with a large difficulty pairing black with positive (and white with negative) corresponding to a strong racial bias. The IAT effect was interpreted as reflecting the individual's unconscious attitudes or preferences for whites over blacks[5].

However, there are several alternative explanations for IAT effects that do not involve unconscious evaluations or prejudice[6]. First, even if we believe that the IAT measures racial bias, a large effect could be observed if a person has positive associations with blacks, but has more positive associations with whites. Even if all interactions were positive, simply having a majority of interactions with whites could produce this IAT bias.

A second possibility is that the IAT instead measures environmental associations—a person's exposure to racially biased information—independent of one's endorsement of that information[7]. Nearly all whites and most blacks show an IAT effect for whites compared to blacks[8]. Although it is possible that all these people are biased against blacks, it is also plausible that they live in a culture that is biased against blacks and that the IAT assesses knowledge of these cultural stereotypes.

A third alternative is that IAT effects may be caused by differential familiarity or salience of blacks and whites. According to the 'figure-ground asymmetry' account, any factor that causes one category to stand out more than another category may lead to IAT effects[9]. For example, unpleasant items are more salient than pleasant ones[10]. White people are more familiar with people of their own race than with people of another race, so black names may attract attention as any unfamiliar item does against a background of more familiar items[11]. Thus, when black names and negative words are categorized together, the salience of both types of target words may facilitate responses. Consistent with this perspective, IAT effects occur when insect names and nonsense letter strings are substituted for white and black names, which is difficult to attribute to a preference for insects over nonwords[12]. Other explanations of IAT effects have also been suggested[6,13].

A final possibility is that IAT effects may be influenced by basic cognitive processes, independent of racial biases, preferences and beliefs. The IAT itself requires cognitive control: performance depends on how well participants switch from one task (classifying names as black or white) to another (classifying words as positive or negative). In addition, midway through the experiment, participants must switch the hands they use for one category. The IAT also requires detecting conflicting responses when black is paired with positive, inhibiting errors and overriding the conflict. All these processes depend on frontal lobe systems for cognitive control[2,4].

These plausible, alternative interpretations suggest that we should be very hesitant to view IAT scores as a measure of racial attitudes, particularly as the self-reported prejudice of individuals does not consistently correlate with their IAT scores[6]. Of course, if the IAT measured attitudes that an individual is unwilling to report, such a discrepancy would be understandable. However, even for attitudes and behaviors that should not be influenced by social desirability concerns, the relationship between IAT scores, attitudes and behavior is disappointing[7].

The alternative interpretations of the IAT are compatible with other accounts of these new results[1]. The psychological demands of the study could have led white individuals, regardless of their racial bias, to exercise cognitive control during an interracial interaction. Although it is difficult to control one's IAT responses, participants can tell if they are showing a racial bias or not (as readers can see at https://implicit.harvard.edu/implicit), and so people with higher IAT scores will be more aware that their performance may be interpreted as evidence of prejudice against blacks.

Such concerns could carry over to the interaction with a black (but not white) confederate. High-scoring participants may be especially concerned that their recently revealed racial bias may be evident somehow, and so may take extra steps to control their thoughts and behaviors, placing high demands on brain systems for cognitive control. The demands would continue after the interaction, and Stroop performance would deteriorate because of the additional cognitive load during the Stroop task itself, not because of the residual effects of prior control[14]. Similarly, IAT and fMRI results could be correlated if the same participants were most concerned with expressing racial bias in the fMRI test.

Thus, the authors may be correct that frontal-lobe cognitive control processes prompted by the interracial interaction mediate the relationship between IAT scores and Stroop performance. However, this control may be needed, not to suppress prejudiced thoughts and actions, but because participants are aware that the experiment concerns race, they are concerned that the IAT has shown them to be biased, and they are monitoring their thoughts and actions accordingly during the interaction and during the Stroop task itself. Such concern is just as likely (if not more likely) in an unbiased person as in a biased person.

These alternative interpretations should not detract from the importance of the new study's findings[1] or the promise of the approach. It is indisputable that prejudice exists, and the scientific study of its cognitive and neural underpinnings is exceedingly important. The findings of Richeson and colleagues raise critical issues that future studies of interracial social behavior should address, and we look forward to the innovative approaches that social neuroscience will continue to bring to these problems.


1. Richeson, J.A. *et al. Nat. Neurosci.* **6**, 1323–1328 (2003).
2. Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S. & Cohen, J.D. *Psychol. Rev.* **108**, 624–652 (2001).
3. Cacioppo, J.T. *et al. J. Pers. Soc. Psychol.* **85**, 650–661 (2003).
4. Gehring, W.J. & Knight, R.T. *Nat. Neurosci.* **3**, 516–520 (2000).
5. Greenwald, A.G., McGhee, D.E. & Schwartz, J.L.K. *J. Pers. Soc. Psychol.* **74**, 1464–1480 (1998).
6. Fazio, R.H., & Olson, M.A. *Annu. Rev. Psychol.* **54**, 297–327 (2003).
7. Karpinski, A. & Hilton, J.L. *J. Pers. Soc. Psychol.* **81**, 774–788 (2001).
8. Nosek, B.A., Banaji, M.R. & Greenwald, A.G. *Group Dynam.* **6**, 101–115 (2002).
9. Rothermund, A.W., & Wentura, D. *Z. Exp. Psychol.* **48,** 94–106 (2001).
10. Pratto, F. & John, O.P. *J. Pers. Soc. Psychol.* **61**, 380–391 (1991).
11. Ito, T.A. & Urland, G.R. *J. Pers. Soc. Psychol.* **85**, 616–626 (2003).
12. Brendl, C.M., Markman, A.B. & Messner, C. *J. Pers. Soc. Psychol.* **81**, 760–773 (2001).
13. De Houwer, J. *J. Exp. Soc. Psychol.* **37**, 443–451 (2001).
14. Meyer, D.E. & Kieras, D.E. in *Attention and Performance XVII. Cognitive Regulation of Performance: Interaction of Theory and Application* (eds. Gopher, D. & Koriat, A.) 17–88 (MIT Press, Cambridge, Massachusetts, 1999).




# Sensory-motor control: a long-awaited behavioral correlate of presynaptic inhibition

P Ken Rose & Stephen H Scott

**Presynaptic inhibition of cutaneous afferents influences sensory-motor responses in the spinal cord. *In-vivo* recordings in monkeys now show that this process suppresses the transmission of cutaneous signals generated during volitional movement.**

A survey of popular neuroscience textbooks suggests that presynaptic inhibition is either rare or a figment of the imagination of physiologists and anatomists of the 1960s. To the contrary, considerable progress has been made in unraveling the cellular and molecular details of this neural process[1,2], but little is known about its role in behavior[3,4]. A report in this issue by Seki *et al.*[5] provides an important functional face to presynaptic inhibition by showing how it modulates cutaneous afferent input to spinal neurons during behavior.

In its most conventional form, presynaptic inhibition involves axo-axonic synapses made by GABAergic interneurons[2]. Although the precise actions of GABA remain open to debate, the end result is clear: GABAergic interneuron activity reduces neurotransmitter release from the postsynaptic axon. As compared to inhibitory synapses directly on the postsynaptic neuron, these axo-axonic synapses selectively reduce input from a particular presynaptic neuron without influencing other inputs to the same postsynaptic neuron.

Seki *et al.*[5] studied the behavioral features of presynaptic inhibition by applying procedures pioneered in the lumbar spinal cord of anesthetized cats to the cervical spinal cord of monkeys trained to execute flexion and extension movements of the wrist. This is no small feat and is arguably one of the most challenging experimental preparations in neuroscience today. The cells that the authors studied receive monosynaptic connections from large-diameter cutaneous afferents. These cells occupy a strategic location (**Fig. 1**); they are the first cells in the spinal cord to relay information from mechanoreceptors to the brain, and they are also the first cells in a circuit that terminates on spinal motor neurons. By regulating the flow of cutaneous signals at this location, descending commands can simultaneously influence motor behavior and the perception of somatosensory stimuli. The key question is under what circumstances do descending commands use presynaptic inhibition via GABAergic interneurons to regulate this flow of information?

The authors[5] used several approaches to show that presynaptic inhibition influences the transmission of cutaneous input to spinal interneurons. First, the effect of stimulating a cutaneous nerve on the activity of spinal neurons was task dependent (**Fig. 1**). When the monkey actively flexed or extended its wrist, cell discharge increased up to 8-fold, but the influence of simultaneous nerve stimulation decreased by 50% from the rest condition. In contrast, comparable passive wrist movements did not affect the influence of nerve stimulation. Because cell discharge was similar during active and passive movements, changes in cutaneous input cannot be simply due to increased refractoriness caused by high-frequency discharge during active movements. Rather, the results suggest that cutaneous input to the spinal neuron was suppressed during volitional movement. This leads to two important questions: what is the source of this suppression and how does it occur?

Seki *et al.*[5] answer the first question by showing that descending commands to the spinal cord are at least partially responsible for suppressing cutaneous input to the spinal cord. The GABAergic interneurons shown in **Figure 1** are known to receive connections from descending systems and primary afferents supplying the skin[2]. Thus, activity in either of these connections could inhibit the responses of first-order interneurons during wrist movement. The transmission of cutaneous signals was not reduced during passive movements, suggesting that descending commands were responsible for presynaptic inhibition during active movement. However, the key observation is that the effect of nerve stimulation was reduced by 20% even before the onset of movement. This pre-movement modulation could not be generated by peripheral afferents and provides definitive proof that descending commands are at least partially responsible for modulating cutaneous input to the spinal cord during movement.

The trickiest part of the study was to demonstrate that modulation of cutaneous input during movement is due to presynaptic rather than postsynaptic inhibition. The finding that cutaneous responses are suppressed at the very time that the interneurons are highly active during extension and flexion suggests that postsynaptic inhibition is not responsible for modulation of cutaneous nerve input. If postsynaptic inhibition is not

Ken Rose is in the Department of Physiology and Stephen Scott is in the Department of Anatomy and Cell Biology, Centre for Neuroscience Studies, Queen's University, Kingston, Ontario K7L 3N6, Canada.
e-mail: ken@biomed.queensu.ca