

# If Nazi = Red, and Canadian = Red, does Red = Good or Bad? Testing the limits of implicit attitude change using the Implicit Association Test

Michael McCarthy & Dr. Andrea Hughes  
Department of Psychology, University of the Fraser Valley

## Introduction

It is difficult to imagine a psychological construct that has led to as much misunderstanding, misappraisal, misuse, and general confusion, as that of the *implicit attitude construct*—particularly because of the magnitude of its popularity, and the amount of empirical and theoretical scrutiny it has received (Greenwald & Banaji, 1995; cf. Nosek, Greenwald, & Banaji, 2007; cf. Mitchell & Tetlock, 2017). Yet, more than two decades after its conception:

- The basic meaning of the term *implicit attitude* remains unresolved (Conrey et al., 2005; Greenwald & Banaji, 2017).
- The construct validity of implicit attitude measures, such as the *Implicit Association Test* (IAT; Greenwald, McGhee, & Schwartz, 1998), remains unestablished (Bar-Anan & Vianello, 2018); although it should be noted that implicit attitude measures that rely on reaction time do seem to share psychometric properties (Bar-Anan & Nosek, 2014).
- The memetic spread of the implicit attitude construct—and the dual-process typology it is founded on—has led to vast swathes of the population (psychologists included) purporting a number of empirically unsupported, internally inconsistent, and perhaps untestable, claims about human social cognition and behaviour (for reviews, see Melnikoff & Bargh, 2018; Mitchell & Tetlock, 2017).

With the above issues in mind, the purpose of the present experiment was to identify the effects of evaluative conditioning and context cues on IAT performance; building upon evidence from prior experiments by exploring the theoretical limits of these effects.

## The Present Experiment

### Methodology

#### Participants

120 (27 male, 93 female) students participated in exchange for course credit and to enter a raffle for \$100. Data for five of these participants was excluded due to meeting predetermined exclusion criteria. All participants were randomly assigned to a condition using a double-blind procedure.

#### Materials

- *Implicit Association Test.* The attribute items on the IAT consisted of five good words (*freedom, peace, honest, paradise, love*), and five bad words (*murder, death, evil, hatred, filth*). The concept items consisted of five red-coloured shapes and five blue-coloured shapes. Congruency order for the IATs was counterbalanced across participants.
- *Free-Association Priming Task.* A set of 38 images depicting nationalistic displays in Nazi Germany or Canada (experimental conditions), or of people gathered in public places (control). Each image was displayed for two-seconds, followed by a good word and a bad word; participants were asked to select which word best reflected their attitude towards each image.
- *Explicit Measures.* Demographic questions, suspicion probes, and measures of participant's explicit attitudes towards the colours red and blue, and the images they saw in the priming task.

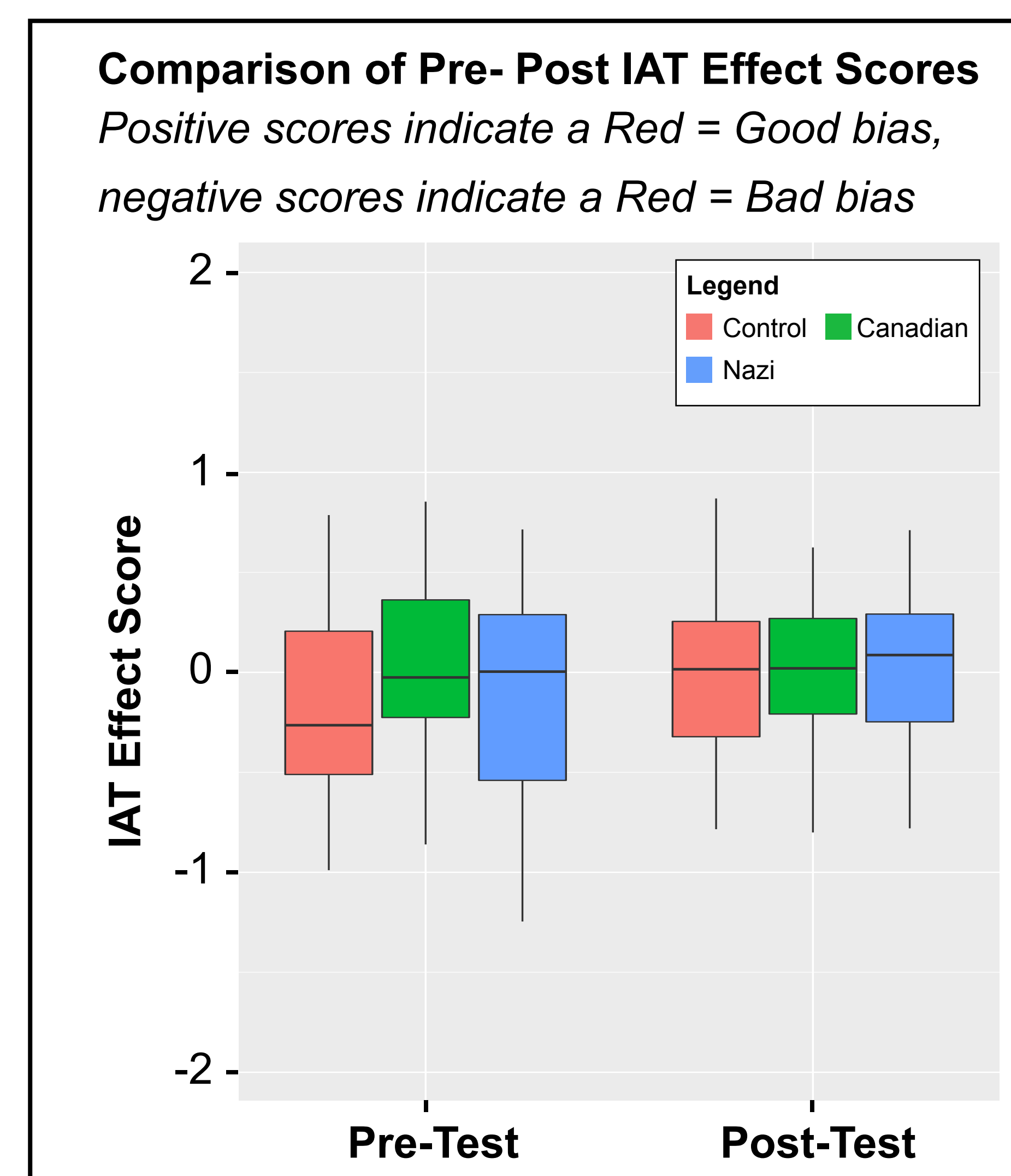
#### Procedure

The experiment was completed on a computer in the following order: 1) Letter of informed consent, 2) Pre-test Red-Blue IAT, 3) Priming task, 4) Post-test Red-Blue IAT, 5) Explicit measures.

### Preliminary Results

#### Did the Prime Affect IAT Effect Scores?

Data analysis was conducted in *RStudio* to ensure all analyses were reproducible. A 3 (condition: Nazi, Canadian, Control) X 2 (time: pre-test vs. post-test) ANOVA with repeated measures on time was conducted on IAT Effect scores. This analysis revealed no significant differences between condition,  $F(2, 224) = 1.17, p = 0.31$ , time,  $F(1, 224) = 2.32, p = 0.13$ , or the Condition X Time interaction,  $F(2, 224) = 1.28, p = 0.28$ .



### Implicit-Explicit Attitude Correlations

Pre-test IAT	1.00			
Post-test IAT	0.43*	1.00		
Colour Attitude	0.22*	0.29*	1.00	
Prime Attitude	0.06	-0.15	-0.14	1.00
	Pre-	Post-	Colour	Prime

## Discussion

Although the theoretical assumptions thought necessary for the experimental manipulation to have an effect were met—namely that participants felt negatively towards Nazis or positively towards Canadians, respectively—it is clear that participant's explicit attitudes towards the priming images did not influence their performance on the post-test IAT.

While this result is not surprising, it is of theoretical interest in light of other implicit attitude change research which has demonstrated that performance on the IAT can be influenced through evaluative conditioning and changes in context cues (for a review, see Gawronski & Bodenhausen, 2006). The present experiment utilized both these mechanisms in its priming task, but set itself apart from prior experiments in the relationship that existed between IAT and prime items.

Where prior experiments utilized a prime that mirrored the IAT content exactly (e.g., showing images of admired Black people and disliked White people prior to a Race-IAT), the present experiment utilized a holistically represented prime (Nazi, Canadian) and an IAT that measured a singular component of that prime (Red); the point of interest here being that the IAT may be insensitive to complex relational information (i.e., Nazi = Red, therefore Red = Bad).

Such a proposition seems reasonable when considering other research that has demonstrated participants respond to the IAT at the categorical level rather than the stimuli level (De Houwer, 2001), in which case it would be expected that IAT performance would only change under conditions where perceptions of the categories themselves change (e.g., if the Red category was relabelled as Nazi on the post-test IAT).