Peer Review Comments

Responses for Version 1

Reviewer A: Etienne LeBel
Review Completed: 3 December 2013

This data paper describes data from the Race Implicit Association Test (IAT) from the Project Implicit website from 2002 to 2012 (self-report attitudes and personality measures and demographic questions also included). The paper is well written and clearly and correctly describes the datasets, which have been properly archived and are openly accessible on the Open Science Framework. The reuse potential for these data are clear and enormous.

The only major issue I noticed is that the datasets are only available in proprietary SPSS ".sav" files, which means they are not accessible to those without an SPSS license.

Some minor issues:
(1) It would be more user-friendly/accessible if the dataset and codebook files on the OSF were sorted chronologically by year. They are currently in a seemingly random order.
(2) On the OSF page (https://openscienceframework.org/project/52qxL/), it says that one can download "4) Skeleton syntax for SPSS to start your own data analysis". I could not find such syntax.
(3) Footnote 1 in the manuscript contains a few typos and missing words:
"...but was then separately into an ethnicity question after a change (in Sep 2006) in the way the U.S. Office of Management and Budget tracked race and ethnicity of the U.S. population. ...or an ethnicity divided by number of sessions reported race? over all years."

Reviewer B: Kevin van Kalkeren
Review Completed: 14 January 2014

In this manuscript entitled 'Psychology data from the Race Implicit Association Test on the Project Implicit Demo website' , the authors provide and describe the data from the on-going Race Implicit Association Test (Race IAT) study. The data include scores of the 2,355,303 Internet volunteers who completed the test between 2002 en 2012. The study focused on measuring the relations between demographic factors and implicit and explicit racial attitudes (IAT scores). Because of the value of this extensive study for social psychology, but also for the meticulous and systematic way the data were collected, I recommend that paper should be published in the Journal of Open Psychology Data, if following proposed adjustments will be taken into account.

First of all, in the Excel codebooks, the variable with the name 'session_id', is probably not supposed to have been labeled 'Seesion' under the 'Label' column. In the SPSS files this Label was named 'Session'. Furthermore, it is not clear to me where and how the .txt files, mentioned in the labeling-information in the codebooks, were obtained in the trial. Thirdly, perhaps the paper could benefit from emphasizing that it was only after 2005 that the self-report constructs were extensively studied.

Lastly, a minor point: in section 4, Reuse potential, I would recommend the changing the sentence to: "...and regional comparisons as in the Nosek et al (2007) paper.".