

## Replication of Hoffman et al 2016

**“Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites”**

**PNAS 2016, 113 (16), 4296-4301.**

<http://www.pnas.org/content/113/16/4296>

*The original paper includes several studies but only one on MTurk. In this between-subject experiment, participants report the amount of pain they would feel across 18 scenarios on a four-point Likert-scale and are then randomly assigned to rate the pain of a gender-matched black or white target across the same scenarios. Thereafter, participants rated 15 biological differences between blacks and whites as true or untrue on a six-point Likert-scale. The original authors find that the pain ratings reported for black targets are lower than for white targets and that this effect was qualified by the interaction between target race and false beliefs.*

**Hypothesis to replicate and bet on:** Non-medically trained, white individuals exhibit a racial bias in pain perception and this bias is related to holding false medical beliefs about blacks vs. whites. The authors test this hypothesis by aggregating the Likert scale surveys continuously and regressing pain ratings on target race, false beliefs, and their interaction, controlling for age, gender, and self-ratings of pain. The authors report a significant and negative interaction between target race and false beliefs,  $\beta = -0.07$ ,  $SE = 0.03$ ,  $F(1,85) = 4.36$ ,  $p = 0.040$ ,  $\eta_p^2 = 0.05$ ; p. 4297.

**Criteria for replication:** The criteria for replication are an effect in the same direction as the original study and a  $p$ -value  $< 0.05$  (two-sided  $F$ -test).

**Power analysis:** The original sample size was 92. The standardized effect size (Cohen's  $d$ ) was  $d = 0.871$ . To have 90% power to detect 67% of the original effect size, a sample size of  $n = 499$  is required.

**Sample:** As in the original article, we will exclude participants who were not born in the US as well as participants who are not native English speakers. We will as in the original article also exclude participants who are nonwhite and participants who do not complete the survey. We will make sure that participants can only participate once from the same account in this specific study, and we will only recruit participants with a HIT approval rate of 95% or higher. We will also check all IP addresses via <https://www.ipqualityscore.com/>; and we will remove any participants where one or more of the following is true: fraud score  $\geq 85$ ; TOR = True; VPN = True; Bot = True; abuse velocity = high. The replication sample size is the sample size after any exclusions of participants.

**Materials:** We will use the same material as in the original study, provided in an OSF repository (<https://osf.io/i4qk5/>) by the original authors. The replication team will re-program the survey in Qualtrics based on the available information.

**Procedure:** We will closely follow the procedure of the original experiment. The following summary of the experimental procedure is therefore largely based on the description of the experiment in the article (pp. 4297–4298):

Participants will first be shown a Captcha, and will thereafter provide informed consent. After this we will include an attention check that participants will need to pass to continue to the study. This attention check is in addition to any other potential attention check(s) used in the original study. Thereafter, participants will be asked to provide their age and gender so the survey program can route them to a gender-matched target.

Then, participants will fill in the scales to be used for the analyses. Participants will be asked to report the amount of pain they would feel across 18 scenarios (e.g., “I slam my hand in a car door”; scale: 1 = not painful, 2 = somewhat painful, 3 = moderately painful, 4 = extremely painful) and will then be randomly assigned to rate the pain of a gender-matched black or white target across the same scenarios. Participants will also be asked to rate the extent to which 15 biological differences between blacks and whites are true or untrue on a six-point scale (1 = definitely untrue, 2 = probably untrue, 3 = possibly untrue, 4 = possibly true, 5 = probably true, 6 = definitely true)”.

After this measure, participants will be asked to provide demographic information, including their race/ethnicity, nationality, and primary language.

**Analysis:** The analysis will be performed as in the original article, by using regression analysis: Pain ratings will be regressed on target race, false beliefs, and their interaction, controlling for age, gender, and self-ratings of pain. The interaction between target race and false beliefs will be tested using an F-test.

**Subject payments:** We are standardizing payments across all replications so that studies have a certain show-up fee depending on the expected length of the study, with an hourly wage from the show-up fee of \$8 and a minimum payment of \$1 (for studies with incentive payment we use the same incentive payment as in the original study; and this payment is paid in addition to the show-up fee). If we have problems recruiting, we will increase the show-up fee.