The Effects of Oncologist Implicit Racial Bias in Racially Discordant Oncology Interactions

Penner, et al

Supplemental Material

1. Administration and Scoring of Implicit Association Test

Logic: The Implicit Association Test (IAT) is predicated on the well-established finding that the more strongly a person associates two things or sees them as belonging together, the more rapidly she/he will respond to a representation of them. Thus, if a person strongly associates, say, the word "mother" with "love" the more rapidly she/he will automatically respond (e.g., recognize) this pairing. The speed of this response would be much faster than the response to a pairing of "mother" and, say "hate." In the race IAT, the speed of response to four pairings are compared: Black-Good; Black-Bad; White Good; White-Bad. Again, it is assumed the stronger the association between the words in these pairings, the faster people will respond to them; the weaker the association the slower they will respond.

Administration: Oncologists were presented positively- or negatively-valenced words (e.g., ugly, beautiful) or pictures of Blacks and of Whites screens that were displayed in the center of computer screens and asked to quickly classify these words/pictures into one of four categories. Two categories represented social groups (Black Patient vs. White Patient) and two represented a valence (negative vs. positive). They were presented as a pair (e.g., "Good or Black Patient," "Good or White Patient"). When the word/picture appeared in the middle of the screen, two contrasting pairs appeared in the screen's upper corners. Oncologists classified the word/picture in the middle of the screen by pressing the appropriate key on the computer keyboard. Oncologists sequentially saw many different screens, which contained all combinations of a set of different words and pictures.

Scoring: Response times for the classifications were computed in milliseconds (ms). Times of less than 200ms and greater than 10,000 ms. were considered too fast or too slow to be valid and were eliminated from the scoring. The average response to the each of the four pairs were calculated and the differences between response times for each of matching pairs (e.g., White Good vs. Black Good; White Bad vs. Black Bad) were computed, and divided by the standard deviation for all scores in that pairing. This yielded a D score for each pair, which is the standardized difference between two means. The two D scores are averaged to form an overall difference score (*d*). This difference represent the relative strength of pro-White and Anti-Black implicit bias[1].

Nosek, Brian A., et al. "Understanding and using the brief implicit association test: recommended scoring procedures." *PloS one* 9.12, 2014

2. Small Effect Size with Large Consequences: The Impact of Aspirin on Deaths Due to Myocardial Infraction

Background. In 1982 *Physicians' Health Group Study*[2] began a study on the effects of aspirin on myocardial infraction. At the time it ended the participants included 22,071 physicians almost equally divided into those who had taken aspirin once a day and those who did not. Over the course of the study, using aspirin reduced the risk of first myocardial infarction by an estimated 44%, which was significant at p<.0000l. (The estimated p in large part reflects the large sample size.) Because of the differences found, the study was ended early. Subsequently, it became quite common for physicians to recommend daily use of low doses aspirin for men at risk for MI's, most typically men who were over 50.

Rosenthal[3] was the first to bring attention to the effect size for this study. He used data from the original report of study to estimate the effect size of aspirin usage on the incidence of MI's over 5 years. He cast the data in a 2 by 2 table: Aspirin/placebo MI present/absent and found an effect size of .03 (the phi coefficient for data in a 2 by 2 table), which is the best estimate of effect size for such data[4]. The table below shows the reported numbers of MI's for the 22,071 physicians in final report of the study[2].

	MI Absent	MI Present
Aspirin	10,898	139
Placebo	10,795	239

As already noted, the differences in the table cells are significant. However, as Rosenthal found, these data yield an effect size of .03. This is considered a very small effect size[5] by almost any criteria.

However this very small effect can have very substantial effects on the overall incidence of MI's . According to U.S. census data, there are 46,00000 men over 50 in the U.S. at for MI's. Based on original findings, the incidence of MI's would be about 2.16% per 100,000 annually without aspirin and 1.27 annually per 100,000 with aspirin. Over a five-year period, there would be 993,600 MI's (2.16% x's 46,000,000) without aspirin and 584,200 with aspirin (1.27 x's 46,000,000). This yields a difference of 409,000 MI's over 5 years[6]. Based on the original data from the Physician Health Group Study, about 10% of these would have been fatal. In sum, a very small effect size has dramatic health implications.

References

1 Nosek BA, et al.: Understanding and using the brief implicit association test: recommended scoring procedures." *PloS one* 9:12, e110938. 2014

2 Steering Committee of the Physicians' Health Study Research Group: Final report on the aspirin component of the ongoing Physicians' Health Study. N. Eng. J. Med. 321:129–35, 1989

3 Rosenthal R: How are we doing in soft psychology? Amer Psychol 45:775-77, 1990

4 Fritz CO, Morris PE, Richler JJ: Effect size estimates: current use, calculations, and interpretation. J Exp Psychol: Gen 141:2-18, 2012

5 Cohen J: A power primer. Psych. Bulletin 112:155-59, 1992

6 Greenwald AG, Banaji MR, Nosek BA: Statistically small effects of the Implicit Association Test can have societally large effects. J Pers Soc Psychol 108:553-61, 2015