



Commentary: Unlearning implicit social biases during sleep

Balazs Aczel^{1*}, Bence Palfi¹, Barnabas Szasz¹, Aba Szollosi¹ and Zoltan Dienes²

¹ Institute of Psychology, Eotvos Lorand University, Budapest, Hungary, ² School of Psychology, University of Sussex, Brighton, UK

Keywords: social biases, IAT, sleep, memory reactivation, system-level consolidation process

A commentary on

Unlearning implicit social biases during sleep

by Hu, X., Antony, J. W., Creery, J. D., Vargas, I. M., Bodenhausen, G. V., and Paller, K. A. (2015). *Science* 348, 1013–1015. doi: 10.1126/science.aaa3841

Without question, reducing social biases is one of the main aims and challenges of modern societies. Implicit biases have been shown to be robust findings in social psychology, popularly measured by the IAT. Recently, Hu et al. (2015) developed a training to counter these biases relying on the system-level consolidation process, in which repeated reactivation of information during sleep is aimed to improve post-sleep memory performance. As a baseline test, participants ($N = 40$) completed two social bias IATs where the face-word assignments were either consistent or inconsistent with the typical bias. In their counter-bias training counter-gender and counter-racial bias trials were paired with different sound cues. All participants took part in a nap session where one of the sound cues was repeatedly presented as they entered slow-wave sleep (SWS). 10 min and 1 week after their nap, the participants completed more IATs to test the endurance of the training for cued and uncued biases. Based on their results, Hu et al. (2015) claimed that (1) implicit social biases were reduced by the counter-stereotypic training; (2) this bias reduction was fortified by memory reactivation during sleep; and (3) the long-term effect of the training is sleep-dependent, as the reduction remained only for the social bias being cued during sleep.

The most important promise of this research is an application of this sleep-training in a potential bias-reducing intervention. The crucial evidence here would be the maintained benefit that the training can produce between the baseline and a later occasion. The relevant analysis, therefore, is where Hu et al. (2015) compared the change between the baseline and the delayed bias assessment, differentiated by the performance on the cued and uncued biases. Specifically, Hu et al. (2015) conducted a Two-Way within-subjects ANOVA (Sleep training \times Time) on IAT scores as dependent variable and they compared the baseline time point with the delayed time point. Neither the interaction (Sleep training \times Time), $F_{(1, 37)} = 0.471$, $P = 0.497$, nor the main effect of Time were significant, $F_{(1, 37)} = 3.095$, $P = 0.087$ (only the interaction was reported). Although the interaction was non-significant, the authors conducted uncorrected *post-hoc t*-tests between the two times finding that the cued biases got significantly weaker, $P = 0.034$, while the uncued biases did not, $P = 0.603$. This simple declaration of the significant and non-significant results reflects a common statistical fallacy by suggesting that the two conditions had different effect (Gelman and Stern, 2006; Nieuwenhuis et al., 2011), whereas the interaction was not significant. Traditionally, the non-significant interaction should disallow the subsequent *t*-tests and non-significant change could not be interpreted as no change in uncued biases. To ascertain whether there is evidence for a null hypothesis, or rather just no evidence either way, we conducted a Bayesian analysis on the original data. Bayes Factor (B) takes into account the effect size obtained, its standard error, and the rough effect size predicted in an optimum way to produce a measure of strength

OPEN ACCESS

Edited by:

Juha Silvanto,
University of Westminster, UK

Reviewed by:

Eric-Jan Wagenmakers,
University of Amsterdam, Netherlands
David Shanks,
University College London, UK

*Correspondence:

Balazs Aczel,
aczel.balazs@ppk.elte.hu

Specialty section:

This article was submitted to
Consciousness Research,
a section of the journal
Frontiers in Psychology

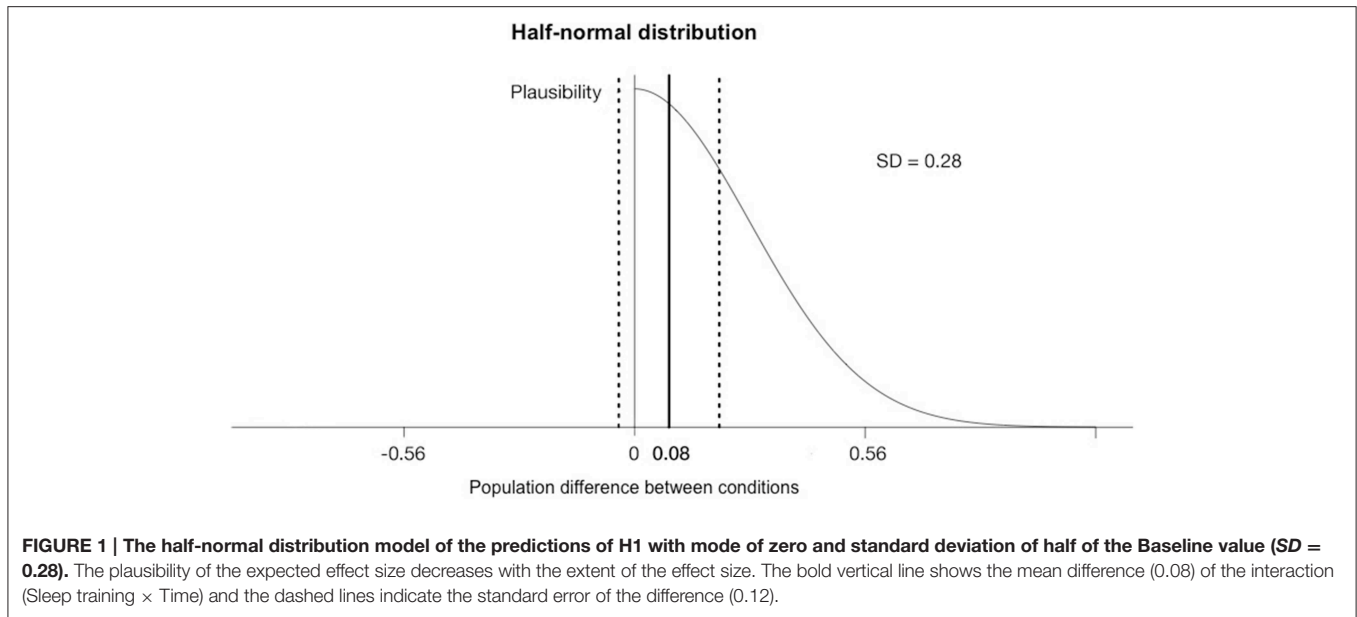
Received: 12 August 2015

Accepted: 07 September 2015

Published: 23 September 2015

Citation:

Aczel B, Palfi B, Szasz B, Szollosi A
and Dienes Z (2015) Commentary:
Unlearning implicit social biases
during sleep. *Front. Psychol.* 6:1428.
doi: 10.3389/fpsyg.2015.01428



of evidence for the alternative hypothesis (H1) rather than the null hypothesis (H0). By convention (Jeffreys, 1961; Dienes, 2014), a B more than 3 is substantial evidence for H1; less than 1/3 is substantial evidence for H0; and in between 1/3 and 3 is not much evidence either way. A p -value by itself does not allow a non-significant result to be interpreted. A Bayes factor allows statements of strength of evidence by modeling the predictions of H1, i.e., taking into account the rough size of effect predicted. Dienes (2014) recommends modeling H1 with a half-normal for a theory making a directional prediction.

Taking a half-normal distribution with a standard deviation of the half of the average range measured in the Baseline condition ($SD = 0.28$) to model the predictions of H1 (see Figure 1), yields $B = 0.72$ for the interaction and $B = 1.56$ for the main effect of time. Both Bayes Factors indicate that we don't have evidence for the efficacy of the training procedure in the long run, because the data are insensitive.

Hu et al.'s study showed convincing evidence that their sleep-training can considerably decrease social biases as measured by the IAT immediately after the participants' nap. We also

agree with the authors that "maintaining the benefits of training is crucial for the ultimate usefulness of potential bias-reducing interventions" (p. 1014) as it would serve great public benefit. Nevertheless, the results of the current experiment are inconclusive as to whether memory reactivation during sleep enhances counter-stereotype training or if bias reduction is sleep-dependent. Bayes Factor calculation could help the authors finalize their analysis by additional data collection. Bayes Factor as a measure of strength of evidence is not sensitive to optional stopping (Rouder, 2014), it enables researchers to collect data until discrimination is reached ($B > 3$ or $< 1/3$). Therefore, the present study could be concluded when the study meets the level of sensitivity sufficient to evaluate the model, but until then we have no evidence that sleep-training can be used as an intervention to reduce social and gender biases in real life.

Acknowledgments

We would like to thank Melissa Wood and Peter Simor for their comments on an earlier draft of this manuscript.

References

- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Front. Psychol.* 5:781. doi: 10.3389/fpsyg.2014.00781
- Gelman, A., and Stern, H. (2006). the difference between "significant" and "not significant" is not itself statistically significant. *Am. Stat.* 60, 328–331. doi: 10.1198/000313006X152649
- Hu, X., Antony, J. W., Creery, J. D., Vargas, I. M., Bodenhausen, G. V., and Paller, K. A. (2015). Unlearning implicit social biases during sleep. *Science* 348, 1013–1015. doi: 10.1126/science.aaa3841
- Jeffreys, H. (1961). *The Theory of Probability*. Oxford, UK: Oxford University Press.
- Nieuwenhuis, S., Forstmann, B. U., and Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat. Neurosci.* 14, 1105–1107. doi: 10.1038/nn.2886

- Rouder, J. N. (2014). Optional stopping: no problem for Bayesians. *Psychon. Bull. Rev.* 21, 301–308. doi: 10.3758/s13423-014-0595-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Aczel, Palfi, Szasz, Szollosi and Dienes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.