

Mind Over Prejudice: An Implicit Bias Training in Medical Education Using Cognitive Bias Modification

Che-Wei Hsu , PhD

Health inequity is defined as the systematic differences in health outcomes and health care opportunities.¹ It is a significant issue for marginalized populations who often experience discrimination and exclusion from mainstream health resources.² Implicit (unconscious) bias, the tendency for individuals to think, feel, or respond in a particular manner that is outside of their awareness, is one contributor to health inequity.³ Implicit bias can hinder patient-practitioner relationships and communication,⁴ prevent access to mainstream treatment, and impact medical decision-making and treatment outcomes,⁵ thereby contributing to increased risk of developing physical and mental health issues.⁶ Presently, only a few bias training methods have demonstrated efficacy.⁷ Among those, training often requires effortful introspection and self-reflection of one's biases, which can be cognitively and temporally demanding.⁸ Here, in the context of existing bias training methods, we introduce a new perspective in bias training for medical learners, known as Cognitive Bias Modification for Stereotype (CBM-S).

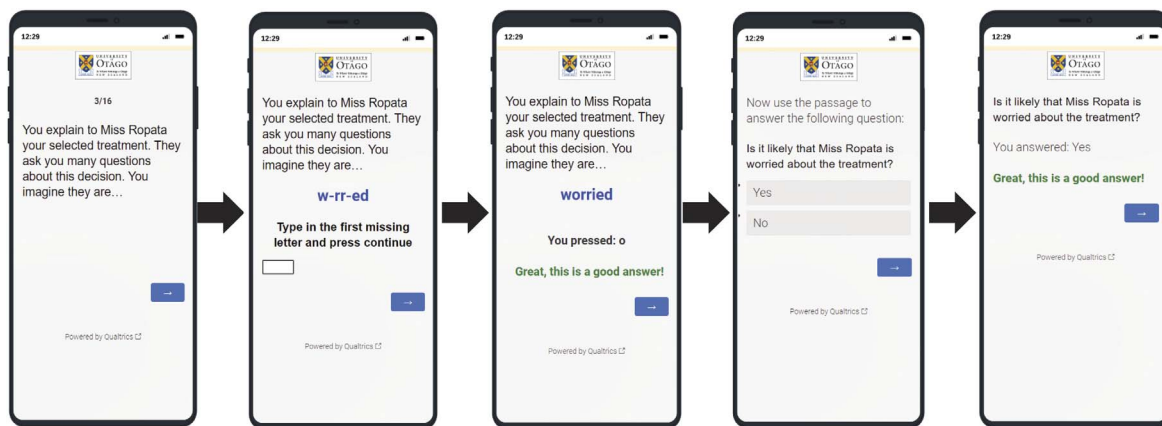
One existing bias training—called metacognition—aims to draw out implicit bias to a conscious level by *explicitly* increasing one's awareness through (1) *naming* bias, (2) *examining* its sources and expression in society, and (3) *working actively* to dismantle it.⁹ By way of example, students and trainees are encouraged to reflect on their own thinking processes through self-questioning (eg, is my clinical decision influenced by specific biases?), critically evaluating evidence and considering alternatives (eg, could there be other possibilities?), and implementing evidence-based decision-making tools to guide their practice (eg, mnemonic checklists). Metacognition is an effective method in reducing bias¹⁰; however, it requires constant monitoring of one's biases with the prerequisite of having good insight and communication.¹¹ Furthermore, some researchers have speculated that metacognitive strategies may increase personal evaluation of one's abilities,¹² which may

lower self-esteem and self-efficacy, strengthening stereotypical associations to protect in-group identity.¹³

Other common methods used to address implicit bias include fact provision,¹⁴ group discussions,¹⁴ and presenting counterstereotypical exemplars (eg, *framing Black people as superheroes and White people as villains*).¹⁵ There are several issues with these methods. First, implicit biases are often developed early in life,¹⁶ making them more resistant to modification.¹⁷ Furthermore, researchers have demonstrated that presenting people with countering information to their original beliefs likely instigates resistance in the form of a rebuttal explanation.¹⁸ This likely explains the small effect size typically found with this class of training, both in medical students,¹⁴ and in the population in general.⁷ Finally, presenting counterstereotypical exemplars may have negative implications, in that people are encouraged to transfer their beliefs from one group to another,¹⁵ which is unethical.

Unlike existing bias training, cognitive bias modification (CBM) methods focus on relatively amendable, punitive cognitive bias processes at an *implicit* level.¹⁹ CBM is a class of bias training based on the premise that social situations are often ambiguous, and that people tend to consistently interpret ambiguity in a particular manner.²⁰ CBM was originally designed to address cognitive bias in mental health. Through (implicit) inferential learning using a prompted word task, CBM trains patients to interpret texts depicting ambiguous everyday events in a more constructive manner (see the [FIGURE](#) for an example of CBM in medical education). In learning theories, prompting users for a response in a forced direction may promote better learning outcomes compared to presenting open information.^{21,22}

Many studies have demonstrated that CBM is effective in modifying intrinsic thinking patterns, with one study showing a reduction in interpretation bias of paranoid-relevant threat (and symptom reduction) following one month of CBM training, with improvements that remained at a 6-week follow-up.²³ CBM has yet to be examined in implicit bias training toward



FIGURE

An Example CBM Training Item for Ethnic Bias in Health Care Delivered via Mobile Phone

Abbreviation: CBM, cognitive bias modification.

marginalized groups in health care. Given that a similar cognitive biased process also exists during social encounters with underrepresented groups,²⁴ there are potential benefits of translating CBM to address unhelpful biases held by students and trainees. To effectively achieve this translation, it is important to take a codesign approach and work collaboratively with relevant contributors to develop materials specific to the targeted bias (eg, bias toward marginalized groups). Researchers have demonstrated the importance of content specificity in capturing social bias (eg, bias toward different ethnicities and genders),²⁵ and in developing an effective CBM intervention.²⁶

At the University of Otago, New Zealand, we created content for and obtained preliminary data for CBM in a health education context with medical students, which we called CBM-S. Our goal was to address biases toward Māori patients—an indigenous group in Aotearoa, New Zealand. Early data revealed promising results, with CBM-S content development and efficacy study results to be reported separately. The FIGURE shows an example of a CBM-S training item that provokes different interpretations of a common ambiguous scenario in health care that graduate medical trainees could encounter: “You explain to Miss Ropata your selected treatment. They ask you many questions about this decision. You imagine they are...” This scenario could be interpreted in various ways, including (1) Miss Ropata is uneducated about western medicine, or (2) Miss Ropata is worried about her health condition and related treatment. A health care professional who holds stereotypical biased beliefs that Māori patients are generally from a low socioeconomic background may be more likely to endorse the former interpretation.

In CBM-S, medical students are presented with a series of common health care scenarios involving

Māori. Similar to a CAPTCHA-type mechanism used to distinguish human from computer responses, the final word of each scenario is presented with missing letters (“...w-r-r-ed”), and users are prompted to enter the first letter of that word to complete the word task. The final word invites a constructive resolution of the ambiguity (“You imagine they are worried.”), which is followed by a yes/no question to reinforce this constructive interpretation (“Is it likely that Miss Ropata is worried about the treatment?” YES). As described above, using a prompted word task is an effective way to navigate the training task, but more importantly, to implicitly force a less biased response.

CBM-S has the advantage over existing bias training in that it is an alternative low-cost training that can be self-administrated and delivered via a digital platform to improve accessibility and scalability. We envisage that CBM training will be used as an adjunct to existing training methods such as metacognition as a part of a weekly independent learning approach in medical education and professional development. Based on previous dose-response studies of CBM in mental health,²³ we hope to see positive and prolonged results in social bias modification after one month of independent training.

In conclusion, health care practitioners’ implicit bias influences health care opportunities, quality, and access, which prevents equitable patient care. We reason that a well-researched and effective clinical bias training called CBM may also have the potential to reduce social bias in a health care context. Preliminary data of this application, which we called CBM-S, have revealed promising results in reducing medical students’ biases toward an indigenous population in New Zealand. CBM-S prompts learners to complete the training task through a word exercise, which implicitly encourages a less biased response. CBM is relatively simple to design

and may be extended to bias training for other marginalized communities (eg, LGBTQIA+, elderly) to close the inequity gap in health care.

References

1. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Population Health and Public Health Practice; Committee on Informing the Selection of Leading Health Indicators for Healthy People 2030. *Criteria for Selecting the Leading Health Indicators for Healthy People 2030*. The National Academies Press; 2019.
2. Baah FO, Teitelman AM, Riegel B. Marginalization: conceptualizing patient vulnerabilities in the framework of social determinants of health—an integrative review. *Nurs Inq*. 2019;26(1):e12268. doi:10.1111/nin.12268
3. Gonzalez CM, Kim MY, Marantz PR. Implicit bias and its relation to health disparities: a teaching program and survey of medical students. *Teach Learn Med*. 2014;26(1):64-71. doi:10.1080/10401334.2013.857341
4. van Ryn M, Fu SS. Paved with good intentions: do public health and human service providers contribute to racial/ethnic disparities in health? *Am J Public Health*. 2003;93(2):248-255. doi:10.2105/ajph.93.2.248
5. Harris R, Cormack D, Tobias M, Yeh LC, Talamaiavao N, Timutimu R. Self-report experience of racial discrimination and health care use in New Zealand: results from 2007/07 New Zealand health survey. *Am J Public Health*. 2012;102(5):1012-1019. doi:10.2105/ajph.2011.300626
6. Paradies Y. A systematic review of empirical research on self-reported racism and health. *Int J Epidemiol*. 2006;35(4):888-901. doi:10.1093/ije/dyl056
7. Lai CK, Marini M, Lehr SA, et al. Reducing implicit racial preferences: I. a comparative investigation of 17 interventions. *J Exp Psychol Gen*. 2014;143(4):1765-1785. doi:10.1037/a0036260
8. Cohen AD. *Strategies in Learning and Using a Second Language*. Routledge; 2014.
9. Sabin J, Guenther G, Ornelas IJ, et al. Brief online implicit bias education increases bias awareness among clinical teaching faculty. *Med Educ Online*. 2022;27(1):2025307. doi:10.1080/10872981.2021.2025307
10. Siegesmund A. Increasing student metacognition and learning through classroom-based learning communities and self-assessment. *J Microbiol Biol Educ*. 2016;17(2):204-214. doi:10.1128/jmbe.v17i2.954
11. Oxford RL. *Teaching & Researching: Language Learning Strategies*. Routledge; 2013.
12. Tarricone P. *The Taxonomy of Metacognition*. Psychology Press; 2011.
13. Rubin M, Hewstone M. Social identity theory's self-esteem hypothesis: a review and some suggestions for clarification. *Pers Soc Psychol Rev*. 1998;2(1):40-62. doi:10.1207/s15327957pspr0201_3
14. Ruben M, Saks NS. Addressing implicit bias in first-year medical students: a longitudinal, multidisciplinary training program. *Med Sci Educ*. 2020;30(4):1419-1426. doi:10.1007/s40670-020-01047-3
15. Joy-Gaba JA, Nosek BA. The surprisingly limited malleability of implicit racial evaluations. *Soc Psychol*. 2010;41(3):137-146. doi:10.1027/1864-9335/a000020
16. Kelly DJ, Quinne PC, Slater AM, et al. Three-month-olds, but not new-borns, prefer own-race faces. *Dev Sci*. 2005;8(6):31-36. doi:10.1111/j.1467-7687.2005.0434a.x
17. Haselton MG, Nettle D, Andrews PW. The Evolution of Cognitive Bias. In: Buss DM, ed. *Handbook of Evolutionary Psychology*. Wiley; 2005:724-746.
18. Hastie R. Causes and effects of causal attribution. *J Pers Soc Psychol*. 1984;46(1):44-56. doi:10.1037/0022-3514.46.1.44
19. Beard C. Cognitive bias modification for anxiety: current evidence and future directions. *Expert Rev Neurother*. 2011;11(2):299-311. doi:10.1586/ern.10.194
20. Beck AT. Thinking and depression. I. idiosyncratic content and cognitive distortions. *Arch Gen Psychiatry*. 1963;9:324-333. doi:10.1001/archpsyc.1963.01720160014002
21. Rowland CA. The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol Bull*. 2014;140(6):1432-1463. doi:10.1037/a0037559
22. Adesope OO, Trevisan DA, Sundararajan N. Rethinking the use of tests: a meta-analysis of practice testing. *Rev Educ Res*. 2017;87(3):659-701. doi:10.3102/0034654316689306
23. Yiend J, Lam C, Schmidt N, et al. Cognitive bias modification for paranoia (CBM-pa): a randomised controlled feasibility study in patients with distressing paranoid beliefs. *Psychol Med*. 2022;14:1-13. doi:10.1017/S0033291722001520
24. Birtel MD, Crisp RJ. Psychotherapy and social change: utilizing principles of cognitive-behavioral therapy to help develop new prejudice-reduction interventions. *Front Psychol*. 2015;6:1771. doi:10.3389/fpsyg.2015.01771
25. von Hippel W, Sekaquaptewa D, Vargas P. The Linguistic Intergroup Bias as an implicit indicator of prejudice. *J Exp Soc Psychol*. 1997;33(5):490-509. doi:10.1006/jesp.1997.1332
26. Vancleef LMG, Peters ML. Examining content specificity of negative interpretation biases with the Body Sensations Interpretation Questionnaire (BSIQ). *J Anxiety Disord*. 2008;22(3):401-415. doi:10.1016/j.janxdis.2007.05.006



Che-Wei Hsu, PhD, is a Lecturer, Department of Psychological Medicine, University of Otago, Dunedin School of Medicine, Dunedin, New Zealand.

Corresponding author: Che-Wei Hsu, PhD, University of Otago, Dunedin, New Zealand, jerry.hsu@otago.ac.nz