Alex Madva March 30, 2013

**Implicit Bias, Moods, and Moral Responsibility**

Background research:

Dovidio et al. (2002): participants had anti-racist explicit attitudes but racially biased implicit attitudes.

* Explicit questionnaire predicted the friendliness of what they said to a black interlocutor
* Indirect measure predicted the *un*friendliness of nonverbal “microbehaviors” (eye contact, seating distance)
* The white participants formed positive impressions of the exchange, while the black interlocutors believed that the participants were consciously prejudiced against them.

Cortina et al. (2011): women, esp. women of color, report experiencing more workplace *incivility* than men.

* City government, law enforcement agency, US military
* Often not overt harassment but generic (and ambiguous) forms of rudeness, e.g. condescension, interruption.
* Those who reported more incivility were more likely to report intention to quit.

🡪 Widespread and frequent repetitions of implicit discriminatory behavior constitute “microinequities,” which stack up to reinforce macro-level disparities.

**Are individuals morally responsible and blameworthy for these implicit discriminatory behaviors?**

Saul (forthcoming): “A person should not be blamed for an implicit bias that they are completely unaware of, which results solely from the fact that they live in a sexist culture. Even once they become aware that they are likely to have implicit biases, they do not instantly become able to control their biases, and so they should not be blamed for them. (They may, however, be blamed if they fail to act properly on the knowledge that they are likely to be biased— e.g. by investigating and implementing remedies to deal with their biases.)”

2 necessary conditions for moral responsibility: Awareness and control. This talk is about awareness.

Individuals might not be aware:

* that implicit biases exist at all,
* that they themselves harbor them,
* that they are expressing them at a given time, or
* that they are able to do anything about them.

Empirical considerations

Implicit biases were originally thought to be unconscious because participants did not report them.

* Greenwald and Banaji (1995, 4-6): implicit biases are “unavailable to self-report or introspection”
* Influential studies (like Dovidio’s above) found dissociations between self-reports and indirect measures.

Recent evidence for awareness:

Experimental manipulations can reduce or remove dissociations

🡪 primarily moving self-reports and explicit measures in the direction of indirect measures.

* “Focus on your feelings” (LeBel and Gawronski 2006).
* Describe indirect measure as “accurate measure of racial attitudes” and “the closest thing to a lie detector that social psychologists can use to determine your true beliefs about race” (Nier 2005, 43).
* Simply tellparticipants that their “gut feelings” reflect their “genuine attitudes” (Payne April 2012).

Hall and Payne (2010, 222): “hard evidence that people have attitudes and beliefs that they don’t know about, or can’t know about when they try, is difficult to find.”

Working hypothesis: affective elements of implicit biases contribute, or are available, to conscious experience.

* In many instances, implicit biases are not objects of explicit attention.
* Implicit biases are *phenomenally conscious* if not (always) *access conscious* 🡪 *felt* if not *noticed*

Evidence outside the laboratory? *Gentleman’s Agreement*: a reporter pretends to be Jewish for six weeks in order to study anti-Semitism, encountering repeated expressions of implicit bias. The director bluntly described the film’s message this way: “You are an average American and you are anti-Semitic. Anti-Semitism is in you.”

* Film came out in 1947: Academy Awards for Best Picture, Best Director, and Best Supporting Actress.

🡪 Social psychology is expanding our appreciation of phenomena that have been part of our collective awareness for some time.

**Does this partial awareness satisfy the awareness condition for moral responsibility?**

Argument by analogy: implicit biases and moods

1. The degree of awareness individuals have of their mood-influenced behavior satisfies awareness.

 🡪 Appeal to commonsense practices of responsibility and “reactive attitudes” surrounding moods.

2. The awareness that individuals have of their implicit discriminatory behavior is importantly similar.

Affective phenomenal experience in ordinary practices of responsibility

**Reactive attitudes** are not “cold, cognitive” evaluations, but automatic, affect-laden reactions.

* Agents *can* become reflectively aware of them, but they need not in order for those reactions to constitute tacit forms of approval or disapproval, and to be potential candidates of praise or blame.
* E.g., I feel resentful toward you for being short with me.

Commonsense understanding of moods, awareness, and responsibility:

I might not notice that I’m in a grumpy mood, but I may be in one just the same (feeling it all the while). My failure to *notice* my mood would not simply exonerate my rude behavior. We routinely hold others and ourselves responsible for the things said and done because of bad moods.

2 sorts of pernicious effects of moods on behavior:

negative affect: Moods lead individuals to express unwarranted negative affect toward others.

biased cognition: Moods bias cognition: what we notice, how we perceive/interpret, how we weight info

Moods are mitigating, not exonerating:

Citing a bad mood as a (partial) explanation can make the behavior appear less objectionable, mitigating the severity of the offence – but not by transporting it outside the realm of blame.

* How moods mitigate can be illuminated by reference to apologies.

Apologies in a completely exonerated case: Nagel’s extremely safe driver who gets into an accident because her brakes suddenly stop working. The driver might feel awful and apologize profusely. Appropriate replies?

* “Apology accepted!”
* “You can’t beat yourself up about this. You did nothing wrong.”

Apologies and moods: Does being “stressed out” or “in a bad mood” have the same blame-defeating force?

An individual snaps at a friend, and subsequently apologizes. (“I’m sorry for being obnoxious. I’ve just been under a lot of pressure lately,” or, “I just woke up on the wrong side of the bed today.”) Appropriate replies?

* “Come now, you have *nothing* to apologize for. You didn’t do anything wrong.”
* “It’s okay. Don’t worry about it. I know things have been stressful for you.”

🡪 Citing a bad mood often puts the person on the receiving end of the rudeness in a position to *accept* the apology, or acknowledge it some way, rather than deny the need for it altogether.

Complicating factors:

* the sort of behavior makes a difference (rudeness vs. under-evaluating a CV or punching someone)
* the more serious the consequences, the less mitigating a bad mood will be (Schlenker and Darby 1981)
* *if* this is ever justified, perhaps it’s because people are better able to control themselves when the stakes are raised (an armed mugger vs. a close friend, a boss vs. a subordinate)

🡪 Individuals are responsible regardless of these mitigating factors.

Implicit discriminatory behavior:

Stephanie wrinkles her nose and shows other signs of automatic disgust (rather than sympathy) at her friend Dennis’ choice of attire (or lifestyle).

* Would Dennis be warranted in resenting her and holding it against her?

If Stephanie were to apologize for the disgusted reaction and explain that it did not represent her beliefs, it would not simply undo the harm or exonerate the reaction. The presence of mitigating factors (such as growing up in a prejudiced society) would not make it the case that Stephanie had *nothing to apologize for*.

🡪 Less blameworthy than reflectively endorsed actions, but more blameworthy than mere reflexes. We should not conclude of her, or each other more generally, that we are all bad prejudiced people, but it is fair to conclude that she could be, in an important sense, better than she is.

Valian (1998, 146):

A storm has damaged a large tree in the back yard, and a tree surgeon has come to look at it… As I ask the tree expert various questions about the damage and what needs to be done, I feel there is something a little odd about his responses. Finally, I realize that I am looking at him when I ask my questions, but that he is looking at J when he answers them. For his part, J is mostly looking abstractly out into space, reflecting his lack of interest in the proceedings. For the entire consultation, in fact, J is silent. I continue with my questions, and the surgeon continues to direct his answers to J. Perhaps he is riveted by J’s virtuosic ventriloquism. I got the information I wanted, but I don’t know what modifications I might have made—speaking louder? asking longer questions? being more assertive?—to get the tree surgeon to talk to me instead of J. I can imagine the surgeon saying to his crew afterward, “Did you see that woman? She didn’t let that guy get a word in edgewise.” J himself has noticed nothing, because he has been thinking about something else the whole time.

**A Few Further Issues…**

Disclaimer on “individual” vs. “social-institutional” approaches to implicit bias: The harms/inequities suffered on the basis of race and gender depend to a great extent on social-institutional forces. Institutional change is necessary, but institutions are composed both of a set of rules and laws as well as a set of individuals, and we have to understand the roles that each plays in contributing to large-scale harms. This talk focuses on the role of individuals and, in particular, the responsibility they bear for their own implicit discriminatory behaviors.

Counterproductive for activism? Attributing responsibility for implicit bias might *backfire*, e.g., it might contribute to the *inefficacy* of diversity training. “Indeed, management consultants and researchers find mixed reactions to diversity management among white males, who report that they are “tired of being made to feel guilty in every discussion of diversity… being cast as oppressors”.” (Hemphill and Haines 1997, cited in Kalev et al. 2006)

Saul: “What we need is an acknowledgement that we are all likely to be implicitly biased—only this can provide the motivation for what needs to be done. If acknowledging that one is biased means declaring oneself to be one of those bad racist or sexist people, we cannot realistically expect the widespread acknowledgement that is required. Instead, we’ll get defensiveness and hostility.”

* I take this strategic concern seriously 🡪 what is the *use* of responsibility and blame in our social practices? Can these notions be productively employed in the broader struggle against prejudice and discrimination?
* Notice that Saul is making falsifiable empirical claims about what can(not) motivate change.
* The empirical evidence is not obviously in her favor (see Holroyd 2012 for discussion).
* Countenancing implicit discriminatory behavior as in “the realm of responsibility” ≠ “playing the blame game” and calling everybody racist and sexist
	+ There’s a lot of room for nuance, e.g., the distinction between “bad behavior” and “bad people.”

Moral relevance without moral responsibility? Kelly and Roedder (2008): implicit biases might be “morally wrong” or even “condemnable” without being “morally blameworthy.”

Alternative distinction: responsibility is *graded* 🡪 comes in degrees

Compare control over addictive behavior and young children’s impulse control. See Coates & Swenson (2012) for an extensive account of responsibility as graded.

Empirical inquiry into “commonsense” or “folk” understandings of moods? See Beedie et al. (2005).

**Bibliography**

Beedie, C., Terry, P., & Lane, A. (2005). Distinctions between emotion and mood. *Cognition & Emotion*, *19*(6), 847-878.

Coates, D.J. and Swenson, P. 2012: Reasons-responsiveness and degrees of responsibility. *Philosophical Studies*.

Cortina, L.M., Kabat Farr, D., Leskinen, E., Huerta, M. and Magley, V.J. 2011: Selective incivility as modern discrimination in

organizations: Evidence and impact. *Journal of Management*.

Dovidio, J.F., Kawakami, K., and Gaertner, S.L. 2002: Implicit and explicit prejudice and interracial interaction. *Journal of*

*Personality and Social Psychology* 82, 62-68.

Greenwald, A.G., and Banaji, M.R. 1995: Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review* 102.

Hall, D., and Payne, B. K. 2010: Unconscious attitudes, unconscious influence, and challenges to self-control. In *Self-control in*

*Society, Mind, and Brain*, edited by Y. Trope, K. Ochsner, and R. Hassin, 221-242*.* New York, NY: Oxford University Press.

Holroyd, J. 2012: Taking Responsibility for Bias. In Special Edition of *Journal of Social Philosophy*, M. Crouch & L. Schwartzman.

Kelly, D., and Roedder, E. 2008: Racial Cognition and The Ethics of Implicit Bias. *Philosophy Compass*, 3 (3), 522-540,

LeBel, E.P., and Gawronski, B. 2006 January: Reasons versus feelings: Introspection and the relation between explicit and implicit

attitudes. Poster presented for the *7th annual Meeting of the Society for Personality and Social Psychology*.

Levy, E. *Gentleman’s Agreement (1947)*. URL = <http://www.emanuellevy.com/review/gentlemans-agreement-1947-6/>

Levy, N. 2012: Consciousness, Implicit Attitudes, and Moral Responsibility. *Noûs*

Nagel, T. 1970: *The Possibility of Altruism*. Princeton, NJ: Princeton University Press.

Nier, J.A. 2005: How dissociated are implicit and explicit racial attitudes? A bogus pipeline approach. *Group Processes and*

*Intergroup Relations* 8, 39-52.

Payne, B.K., Cooley, E., Lei, F. 2012 April: Who owns implicit attitudes? Testing a meta-cognitive perspective. Presentation for *The*

*Implicit Bias & Philosophy Workshop*. University of Sheffield, UK.

Saul, J. Forthcoming: Unconscious Influences and Women in Philosophy. In *Women in Philosophy: What Needs to Change?*

Schlenker, B.R., and Darby, B.W. 1981: The use of apologies in social predicaments. *Social Psychology Quarterly* 4, 271-278.

Strawson, P.F. 1974: *Freedom and resentment and other essays*. London: Methuen.

Valian, V. 1998: *Why so slow? The advancement of women*. Cambridge, MA: M.I.T. Press.