I. Summary

Accuracy-affecting injustice:
· Biased (inaccurate) performance reviews, used as data for prediction.
· Over-policing and using past-arrests for prediction.

Non-accuracy-affecting injustice:
· E.g. using educational attainment in employment decisions, or income in lending decisions.

The cases of compounding injustice she’s focusing on are non-accuracy-affecting. Examples:
· Amy, a victim of domestic abuse, and life-insurance premiums.
· Aliyah, a sufferer of unjust educational-resource distribution, and hiring decisions.
· Darnell, a low-income worker due to racial discrimination, and lending decisions.

Hellman argues that when (you have evidence that) someone has been the victim there is additional moral reason to avoid doing things which would compound that injustice.

Anti-compounding injustice principle (ACI): When an actor takes the fact of prior injustice or its immediate sequalae¹ as a reason for action, “the actor involves itself sufficiently in the prior injustice to bear some responsibility for any extension of the harm of the injustice that its actions give rise to” (6).

What’s wrong with compounding injustice?
· Basically: it seems bad, since it makes a bad situation worse (9).

Why think it poses a special problem for big data?
· Basically, because most real-world data is a consequence of past injustice, so the more we rely on it, the more we’ll compound (even if our algorithms are accurate!).

II. Questions

Q1: Modest vs. radical version of claim.
· **Modest:** Increasing marginal disutility of harms. People who’ve previously suffered harms are more likely to suffer more from further harms.

· **Radical:** there’s a special reason to avoid compounding injustice over and above the extra harm.

· Cases to think about:
  - Two people in exact same situation, one as a result of injustice other as a result of misfortune.
  - A has faced prior injustice, B hasn’t. You can only give one a loan. B will be harmed more by not getting a loan (he’ll lose his job) than A (she’ll lose her car). Should you give to A or B?

**Q2:** Does this just apply to actions, or also to predictions?

· Is this meant to be an objection to the approaches to algorithmic fairness that Hedden and others are talking about, or just pointing to another issue that that discussion is distracting from?

· Should we think of this as a question for the epistemology or for ethics of algorithmic bias?

  Standard reasons for separation between belief and values: the value of evidence; we don’t know what decisions we’ll face, and we should expect more accurate beliefs to help us navigate them better.

· E.g. Tamar Gendler’s case: at a ritzy club, you see a Black man. Given what you know about the base rates, 9 in 10 Black men in the club work there. But you don’t want to offend/insult/compound injustice by presuming that he does. Consider two options:
  
  $O_b$, change beliefs: modestly disvalue mistaken categorization,\(^2\)
  
  but have a middling credence he’s working: $C(\text{works}) = 50\%$.

  $O_v$, change values: strongly disvalue mistakenly categorizing him,\(^3\) but have a high credence he’s working: $C(\text{works}) = 90\%$.

  Which to do? Reason to do $O_v$ is that if you face a new decision, with unknown potential (dis)values, you’ll be best able to navigate it if you have accurate beliefs.

  Back to Hellman and Hedden (and Siegel, etc.). Do we want to (think we can) separate the epistemology from the ethics cleanly?

**Q3:** Why think this is going to get worse with big data?

· Sure, past data is the product of injustice. But so are the factors that go into non-big-data approaches to hiring/lending/etc.!

**Q4:** Quibble: why does ACI hinge on whether we take those things as reasons? What about causal antecedents to injustice that are nevertheless correlated?