

Response to “Two Cheers for Corporate Experimentation”

Jeffrey Fossett

In her article *Two Cheers for Corporate Experimentation: The A/B Illusion and the Virtues of Data-Driven Innovation*¹, Michelle Meyer defends the ethical status of corporate experimentation (or A/B testing), claiming that this practice is not only morally *permissible*, but in fact *preferable* to the alternative (in at least some cases) on account of the welfare gains it may confer for users.

In this essay, I argue that while Meyer is correct that A/B testing is morally permissible in some cases (contrary to wholesale critics), she overstates the moral praiseworthiness of the practice. I claim, instead, that A/B testing in and of itself should be viewed as morally neutral, and that corporations should be ethically evaluated (and perhaps regulated) based on *how* and not *whether* they deploy this practice.

Meyer outlines her ethical defense of A/B testing in the opening sections of *Two Cheers for Corporate Experimentation*. She explains that when corporate changes are A/B tested, this allows them a “far greater chance of being discovered to be unsafe or ineffective [for users]”. If corporations act appropriately on such findings (or share them publicly), A/B testing can thus lead to “substantial welfare gains” for users. This is in contrast to conventional approaches to corporate innovation, in which corporations simply implement a preferred change “by fiat”, retaining little ability to understand the causal impact of the change².

In Meyer’s view, the practitioner who A/B tests is thus better positioned for what she calls “responsible innovation”³. Since A/B testing can deliver a precise understanding of the impact that changes have on users, companies that experiment can better respect their users’ best interests (or at least share findings so that users can make informed decisions). Through this lens, it is in fact the practitioner who implements changes *without* testing who is comparatively unethical: exploiting her power over users to unilaterally implement changes that are believed to be in the company’s best interests, while undertaking no effort to rigorously understand the true impact of the change on user welfare.

Meyer believes that critics of corporate experimentation have been misled by what she calls the “A/B testing illusion” or “the tendency to focus on the experiment in the foreground rather than the ongoing practice that exists in the background”⁴. In short, critics have overreacted to the “risk, uncertainty and power asymmetries” involved in companies experimentally comparing *A* to *B*, while ignoring that the alternative is either (i) the unilateral implementation of *B* without evidence, or (ii) the continued adherence to the baseline *A*, which was itself unilaterally implemented without evidence in the past, and thus is not morally privileged⁵.

¹ Michelle Meyer, *Two Cheers for Corporate Experimentation: The A/B Illusion and the Virtues of Data-Driven Innovation*, (Mar. 7, 2017, 12:22 PM), <http://ctlj.colorado.edu/wp-content/uploads/2015/08/Meyer-final.pdf>

² *Id.* at 277.

³ *Id.* at 310.

⁴ *Id.* at 327.

⁵ Michelle Meyer, *Please, Corporations, Experiment on Us*, (Mar. 7, 2017, 12:22 PM), https://www.nytimes.com/2015/06/21/opinion/sunday/please-corporations-experiment-on-us.html?fb_ref=Default&_r=0

Ultimately, although Meyer does not defend A/B testing in all cases, she does conclude based on the above considerations that it has much to recommend it, at least when certain criteria are met (e.g. risks to users are minimal, welfare gains are localized to those exposed to test). As she puts in the conclusion of a New York Times op-ed on the subject⁶:

... as long as we permit those in power to make unilateral choices that affect us, we shouldn't thwart low-risk efforts, like those of Facebook and OkCupid, to rigorously determine the effects of those choices. Instead, we should cast off the A/B illusion and applaud them.

Meyer is right to highlight the A/B testing illusion, and is correct in her defense of corporate experimentation against those who would reject it based on the view that all human experimentation is “inherently dangerous ... [and that its pursuit] without informed consent is absolutely unethical”⁷. Indeed, as Meyer points out, insofar as we are comfortable with companies unilaterally implementing some change *B* without evidence, there is little ground to criticize the firm that prefers to test it against the baseline *A* first. As Meyer points out, there is no *additional* risk or exploitation involved in the latter case; thus, insofar as we are comfortable with the former, we ought to be comfortable with the latter as well.

However, Meyer also makes a stronger claim: namely, that A/B testing is actually a *more responsible* approach to corporate innovation, and is thus morally-preferable and worthy of applause. As noted above, this claim is based on the premise that A/B testing leads to a “far greater chance of [changes] being discovered to be unsafe or ineffective [for users]” and can thus yield “substantial welfare gains” for those users if results are shared or acted on. But does A/B testing really lead to welfare gains for users? And if it does, does this necessarily make it worthy of moral applause? We investigate these questions in the remainder of this essay.

When corporate and user interests are aligned, users will generally benefit from A/B testing. For example, when Airbnb builds a new search algorithm, the algorithm is valuable to Airbnb only insofar as it does a better job of matching guests with homes that are a good fit for their travel plans. Since users come to Airbnb with precisely this goal in mind, corporate and user interests are fully aligned. Thus, if Airbnb uses A/B testing to optimize its search algorithm and maximize the chance of a match occurring, users benefit (on average) by having an easier time planning a trip. Moreover, since A/B testing allows Airbnb to do this optimization in a more rigorous and data-driven way, users actually benefit *more* in the world Airbnb experiments than in the world where Airbnb does not. In short, A/B testing leads to an increase in user welfare in cases like this where corporate and user interests are aligned. In *Two Cheers*, Meyer explores a similar example that involves OkCupid experimenting to understand the effectiveness of its matching algorithm, a question that likewise aligns corporate and user interests⁸.

Fortunately for users, corporate and user interests typically *are* aligned, and companies generally succeed and profit to the extent that they reward and satisfy their users. Supposing we concede that A/B testing is beneficial to users in such cases, the next question is whether this confers any moral praiseworthiness on

⁶ *Id.*

⁷ Meyer, *supra* note 1

⁸ Meyer, *supra* note 1, at 312

the company that chooses to experiment. Here, the answer seems to be no. The choice to A/B test neither implies nor requires that a company attend to any aspects of user welfare beyond what they would already do by pursuing their own self-interest. The decision to A/B test simply allows a company to better optimize those aspects of user experience that are aligned with corporate interests. Such self-interested pursuit does not seem particularly praiseworthy, except in perhaps the narrowest consequentialist sense. We don't, for example, believe that Apple is worthy of moral praise if it builds a more appealing laptop than Dell. They are just a successful company doing a good job at pursuing their interests in the marketplace. Likewise, a firm like Airbnb that chooses to experiment is not morally preferable to a hypothetical competitor that does not--it's simply a firm that will build a better product.

Although corporate and user interests *are* generally aligned, this is of course not always the case. For example, as a marketplace, Airbnb sometimes has an economic incentive to favor certain market participants (e.g. guests) over others (e.g. hosts) in order to create a more efficient marketplace (which benefits Airbnb). Likewise, Facebook has an economic incentive to surface "fake news" or other clickbait content, even if such content is not ultimately beneficial to users or the broader political dialogue. In cases like this, experiment-driven optimization will typically lead to *worse* outcomes for users (in comparison with the non-experimental world). This is possible because, as noted above, the choice to experiment does not imply any *additional* attention to user welfare beyond what's aligned with corporate interests. Thus, when the corporate & user interests are misaligned, experimentation in and of itself will do little to protect user interests, as Meyer's "responsible innovation" frame might suggest. Now, the objection might be raised that the moral problem here is not with experimentation per se, but with the company that decides to pursue self-interest contra the interests of their users. But this is precisely the point: it's the choice about whether to be sensitive to user interests in the first place--not the choice about whether to experiment--that is morally relevant.

To be fair to Meyer, the main case of corporate experimentation addressed in her article (and much of the broader public dialogue on corporate experimentation) is the well-known Facebook emotion contagion experiment⁹. This experiment is unusual in that it is a case of corporate experimentation in which the interests of the corporation are not strongly represented in the research question ("Does emotion spread in social networks?")¹⁰. Instead, it seems clear for a number of reasons (publication in PNAS, involvement of Cornell researchers etc.) that the study was undertaken in order to address an open (and otherwise difficult to answer) academic question with clear relevance to the general welfare of Facebook users (as Meyer points out).

I agree with Meyer that this sort of research is praiseworthy and in general ought to be encouraged. However, I retain that experimentation per se has little to do with this conclusion. The source of moral praise here is the fact that Facebook chose to contribute time and resources to address an important question of interest to their users and the broader public (which they were uniquely positioned to answer), despite its being of limited relevance to their narrow self-interest. That these resources went towards the

⁹ Adam D.I. Kramer, Jamie E. Guillory & Jeffrey T. Hancock, *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*, 11 PROC. NAT'L ACAD. SCI. 8788, 8788 (2014).

¹⁰ Given the fact that the Facebook experiment is so unrepresentative of the typical corporate A/B test, I have found it somewhat misleading that it has been the primary battleground for so much of the public (and academic) discourse around corporate experimentation.

execution, analysis and publication of a randomized experiment (as opposed to, say, some other sort of analysis) is not morally relevant.

Ultimately, A/B testing is simply a tool that allows companies to more rigorously and effectively pursue the ends that they are interested in. As Meyer's arguments around the A/B illusion show, A/B testing is not inherently morally problematic or objectionable. However, as this essay has demonstrated, it also does not imply any particular moral worth or corporate responsibility. As with many new technologies, the key moral question is not whether companies use this new tool, but to what end. As such, ethical and regulatory debates around corporate experimentation should focus less on *whether* companies experiment and more on *how* to ensure that they experiment responsibly¹¹.

¹¹ For example, we should encourage (or perhaps even require) companies to undertake studies like the Facebook emotion contagion study that help us understand the broader social impact of new technologies. We should also encourage companies to run product experiments in a way that represents the interests of their users (including atypical or minority guests), even when these interests are orthogonal to corporate interests. These are issues for another essay.