11/9/2018

Richard Rennehan

isec3050-701

The Ethics of A/B Testing in the Tech Industry

Case 1

Contents

[Introduction to A/B Testing 2](#_Toc529309540)

[Perspectives 4](#_Toc529309541)

[Example 6](#_Toc529309542)

[Conclusions 9](#_Toc529309543)

[References 10](#_Toc529309544)

# Introduction to A/B Testing

Who doesn’t like to experiment? A/B testing, also known as split testing, is to perform an experiment comparing the results between multiple scenarios. A/B testing is a tried and true method for businesses to figure out how to make their websites, designs, and campaigns more effective. A/B testing helps businesses understand customer behaviors and helps the business optimize its services. However, if A/B testing is treated like a human experiment, it can get into deep water.

The most well-known form of A/B testing is testing multiple website designs to see which is most effective. This form is A/B testing can be completely innocent. An example revolves around Netflix. Netflix once wanted to know whether showing a list of available movies and shows will result in more people signing up for the free 1-month trial ([Manifesto], 2016). The test involved four variations of their site: the original site without the list plus three variations showing the list ([Manifesto], 2016). Users were randomly assigned a version of the site ([Manifesto], 2016). The site that got the largest percentage of users to sign up was declared the winner ([Manifesto], 2016). In a surprise to Netflix, the original site without the movies and shows listed converted the most people over ([Manifesto], 2016). The test was performed 5 times, with the winner moving on to the next round and the losers replaced with new designs ([Manifesto], 2016). The original won every time ([Manifesto], 2016). From this, Netflix concluded that showing the lists distracts and slows down customers, and that it was not necessary to show it when the trial was free. Netflix learned that the experience of Netflix is what sold it to customers, and that users do not always know what they want ([Manifesto], 2016).

Most A/B tests are harmless, especially when they are business-oriented. However, there have been a growing amount of cases where A/B tests are leading to ethical issues, especially when they are human-oriented. John Constine (2014) writes that the biggest issues to tackle are: making riskier experiments opt-in, security audits on tech companies, and educating data scientists on more ethical A/B testing practices.

For science organizations, the laws are pretty clear. Human experiments must be approved by the Institutional Review Board: an ethics committee that requires scientific experiments to meet stern safety and consent standards to ensure the welfare of their subjects (John Constine, 2014, para. 8). However, legislation on A/B testing in the tech industry is light in comparison. Are tighter regulations needed for experiments conducted by tech companies? What affect will new A/B testing legislation have on tech companies?

# Perspectives

The most brought up point is that users often are subjected to an A/B test without consenting to it. It is almost guaranteed that you have been a participant in an A/B test. It can be argued that an example like Netflix’s case, where the results were business-oriented, are harmless. However, many human-oriented A/B experiments can have a negative impact on the user. For example, these experiments may play with the individual’s emotions and well-being. Perhaps if the user knew about the experiment, they never would have agreed to it. Some state their support for opt-in stating that by making riskier experiments opt-in, it will protect consumers from harm and form better relations between the business and consumer. On the opposing side, others argue that it will remove the randomness factor of the experiment, thus, preventing progress from being made.

My idea is that large-scale A/B tests from tech companies should be sent in for review by an ethics board. The board can give three base results: approved, requires user opt-in, and unapproved. Punishments can be set when a company that performed a risky experiment without approval is caught. This will help protect individuals from unlawful testing and provide a consistent standard for businesses to follow. This also answers the randomness argument, as user opt-in requirements are set only when necessary.

There currently exists the Institutional Review Board (IRB): an ethics committee that requires scientific experiments to meet stern safety and consent standards to ensure the welfare of their subjects (John Constine, 2014, para. 8). A possibility is to also require tech companies to comply with IRB submissions and standards. Another possibility is to create a whole new ethics board dedicated to technology companies, if one does not currently exist. Either way, the legislation needs to be stricter. In addition, there should be more policing to catch companies performing harmful studies without permission and to dissuade companies from doing so in the first place.

To conclude, the laws around A/B testing in the tech industry should be more defined to create a standard of ethical practice and so users are more confident they are safe online. Users should know of any potentially harmful experiment if necessary.

# Example

The most cited example of unethical A/B testing is Facebook’s newsfeed emotion study. In this 2014 study, the newsfeeds of over 689,000 users were manipulated with either more positive posts or more negative posts, all without their permission (Robert Booth, 2014, para. 2). Facebook later studied the posts of the 689,000 users to conclude whether emotions expressed by their friends and the news influenced their own moods.

Facebook was quickly attacked following the publishing of their results. The full study can be found at <http://www.pnas.org/content/111/24/8788.full>. There are many reasons why the study triggered a backlash. Users were completely unnotified of the study and had zero ability to opt-in or opt-out. Facebook also failed to reach out to an ethics board, performing only an internal review before commencing the study (John Constine, 2014, para. 8). In addition, it is possible that the study may have had a long-lasting negative impact on someone, such as going into depression or losing a job over a post.

Many criticized Facebook for performing such a risky study without knowledge and consent from their userbase. Jim Sheridan, a member of the Commons media select committee, stated his opinion:

This is extraordinarily powerful stuff and if there is not already legislation on this, then there should be to protect people. They are manipulating material from people's personal lives and I am worried about the ability of Facebook and others to manipulate people's thoughts in politics or other areas. If people are being thought-controlled in this kind of way there needs to be protection and they at least need to know about it. (Robert Booth, 2014, para. 7)

James Grimmelmann, a professor of law at Maryland University, said the following about Facebook’s actions in relation to ethics around human experiments:

Facebook had failed to gain "informed consent" as defined by the US federal policy for the protection of human subjects, which demands explanation of the purposes of the research and the expected duration of the subject's participation, a description of any reasonably foreseeable risks and a statement that participation is voluntary. "This study is a scandal because it brought Facebook's troubling practices into a realm – academia – where we still have standards of treating people with dignity and serving the common good (Robert Booth, 2014, para. 16).

A Facebook spokeswoman said the following about the research following the backlash:

The research was committed to improving our services and to make the content people see on Facebook as relevant and engaging as possible. A big part of this is understanding how people respond to different types of content, whether it’s positive or negative in tone, news from friends, or information from pages they follow (Robert Booth, 2014, para. 8)

It’s important to note that Facebook technically has consent. Facebook’s data use policy which people automatically agree to when signing up states “we may use the information we receive about you…for…data analysis, testing, research and service improvement” (John Constine, 2014, para. 9). However, many still believe Facebook has breached ethical and legal guidelines by failing to inform its users they were being manipulated in the experiment. I feel Facebook did wrong for not notifying users. The sweet spot would have been to notify users that they were selected for an experiment about newsfeed manipulation and to give them the option to opt-out if desired.

Robert Blackie, director of digital at Ogilvy One Marketing Agency, concludes with a refreshing point of view. He states:

To guarantee continued public acceptance they will have to discuss this more openly in the future. There will have to be either independent reviewers of what they do or government regulation. If they don't get the value exchange right then people will be reluctant to use their services, which is potentially a big business problem (Robert Booth, 2014, para. 20).

# Conclusion

A/B testing is a fine example of the phrase: “With great power comes great responsibility.” As the case with Facebook’s emotion study reveals, there should be more regulation around A/B testing and experiments in the tech industry. Tech companies can perform potentially risky experiments without fines or punishment too easily. As expectations of privacy and safety online continue to heighten, so should the relating laws and regulations. Tighter and more defined regulations around the tech industry for experiments will create a clearer standard and protect people from potentially harmful experiments without their permission.

# References

Booth, R. (2014, June 30). Facebook reveals news feed experiment to control emotions. Retrieved from <https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds>

Constine, J. (2014, June 30). The Morality Of A/B Testing. Retrieved from <https://techcrunch.com/2014/06/29/ethics-in-a-data-driven-world/>

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014, June 17). Experimental evidence of massive-scale emotional contagion through social networks. Retrieved from <http://www.pnas.org/content/111/24/8788.full>

[Manifesto]. (2016, Jan 11). Design Testing at Netflix [Video File]. Retrieved from <https://www.youtube.com/watch?v=-Gy8TnoXZf8>