INTERNET POLICY REVIEW
Journal on internet regulation

Volume 3 | Issue 4

Staking out the unclear ethical terrain of online social experiments

Cornelius Puschmann
Alexander von Humboldt Institute for Internet and Society, Berlin, Germany, cornelius.puschmann@hiig.de

Engin Bozdag
Delft University of Technology, Netherlands, v.e.bozdag@tudelft.nl

Published on 26 Nov 2014 | DOI: 10.14763/2014.4.338

Abstract: In this article, we discuss the ethical issues raised by large-scale online social experiments using the controversy surrounding the so-called Facebook emotional contagion study as our prime example (Kramer, Guillory, & Hancock, 2014). We describe how different parties approach the issues raised by the study and which aspects they highlight, discerning how data science advocates and data science critics use different sets of analogies to strategically support their claims. Through a qualitative and non-representative discourse analysis we find that proponents weigh the arguments for and against online social experiments with each other, while critics question the legitimacy of the implicit assignment of different roles to scientists and subjects in such studies. We conclude that rather than the effects of the research itself, the asymmetrical nature of the relationship between these actors and the present status of data science as a (to the wider public) black box is at the heart of the controversy that followed the Facebook study, and that this perceived asymmetry is likely to lead to future conflicts.

Keywords: Research ethics, Online social experiments, Data science, Transparency, Algorithmic curation

Article information

Received: 07 Aug 2014 **Reviewed:** 06 Oct 2014 **Published:** 26 Nov 2014
Licence: Creative Commons Attribution 3.0 Germany
Competing interests: The author has declared that no competing interests exist that have influenced the text.

URL:
http://policyreview.info/articles/analysis/staking-out-unclear-ethical-terrain-online-social-experiments

Citation: Puschmann, C. & Bozdag, E. (2014). Staking out the unclear ethical terrain of online social experiments. *Internet Policy Review*, *3*(4). https://doi.org/10.14763/2014.4.338

1. THE FACEBOOK EMOTIONAL CONTAGION EXPERIMENT

The article "Experimental evidence of massive-scale emotional contagion through social networks" by Adam D.I. Kramer (Facebook), Jamie E. Guillory (University of California) and Jeffrey T. Hancock (Cornell University) was published on 17 June 2014 in *Proceedings of the National Academy of Sciences of the United States of America* (PNAS), a highly competitive interdisciplinary science journal (cf. Kramer, Guillory, & Hancock, 2014). The paper tested the assumption that basic emotions, positive and negative, are contagious, that is, that they spread from person to person by exposure. This had been previously tested for face-to-face communication in laboratory settings, but not online, and not using a large random sample of subjects. The authors studied roughly three million English language posts written by approximately 700,000 users in January 2012. The experimental design consisted of an adjustment of the Facebook News Feed of these users to randomly filter out specific posts with positive and negative emotion words to which they would normally have been exposed. A subsequent analysis of the emotional content of the subjects' posts in the following period was then conducted to determine whether exposure to emotional content would affect the subjects. Kramer and colleagues stressed that no content was added to the subjects' News Feed, and that the percentage of posts filtered out in this way from the News Feed was very small. The basis for the filtering decision was the Linguistic Inquiry and Word Count (LIWC) software package, developed by James Pennebaker and colleagues, which is used to correlate word usage with physical well-being (Pennebaker, Booth, & Francis, 2007). LIWC's origins lie in clinical environments and originally the approach was tested using diaries and other traditional written genres, rather than short Facebook status updates (Grohol, 2014). The study found that basic emotions are in fact contagious, though the effect that the researchers measured was quite small. The authors noted that given the large sample, the global effect was still notable, and argued that emotional contagion had not been observed before in a computer-mediated setting based purely on textual content.

The article provoked some very strong reactions both in the international news media (e.g. *The Atlantic*, *Forbes*, *Venture Beat*, *The Independent*, *The New York Times*) and among scholars (James Grimmelmann, John Grohol, Tal Yarkoni, Zeynep Tufekci, Michelle N. Meyer — see Grimmelmann, 2014b, for a detailed collection of responses). *The New York Times*' Vindu Goel surmised that "to Facebook, we are all lab rats" and *The Atlantic*'s Robinson Meyer called the study a "secret mood manipulation experiment" (Goel, 2014, M.N. Meyer, 2014). Responses from scholars were more mixed: a group of ethicists reacted with skepticism to the many critical media reports, arguing that they overplayed the danger of the experiment and warning that the severe attacks could have a chilling effect on research (M.N. Meyer, 2014). Several critics noted that the research design and the magnitude of the experiment were poorly represented by the media, while others claimed that a significant breach of research ethics had occurred, with potential legal implications (Tufekci, 2014; Grimmelmann, 2014a). First author Adam D.I. Kramer responded to the criticism with a Facebook post in which he explained the team's aims and apologised for the distress that the study has caused (Kramer, 2014).

The strong reactions provoked by the paper, especially in the media, seem related to the large scale of the study and its widespread characterisation as "a mood-altering experiment" (Lorenz, 2014). Furthermore, the 689,003 users whose News Feeds were changed between 11 and 18 January 2012 were not aware of their participation in the experiment and had no way of knowing how exactly their News Feeds were adjusted. In their defense, Kramer and colleagues

pointed out that: (1) the content omitted from the News Feed as part of the experiment was still available by going directly to the user's Wall; (2) the percentage of omitted content was very small; (3) the content of the News Feed is generally the product of algorithmic filtering rather than a verbatim reproduction of everything posted by one's contacts; and (4) no content was examined manually, that is, read by a human researcher, but that the classification was determined by LIWC automatically. Some of these aspects were misrepresented in the media reactions to the study, but more basic considerations such as how the study had been institutionally handled by Facebook, Cornell, and *PNAS*, and whether agreement to the terms of service constituted informed consent to participation in an experiment were also raised in the debate that followed.

2. THE UNCLEAR ETHICAL TERRAIN OF ONLINE SOCIAL EXPERIMENTS

How can the extremely divergent characterisations of the same event be explained, and what do such conflicting perspectives spell out for the ethics of large-scale online social experiments? In what follows, we will discuss these questions, drawing on multiple examples of similar studies. Researchers at Facebook have conducted other experiments, for instance studying forms of self-censorship by tracking what users type into a comment box without sending it (Das & Kramer, 2013); displaying products that users have claimed through Facebook offers to their friends in order to see whether a buying impulse is activated by peer behaviour (Taylor et al., 2013); showing users a picture of a friend next to an advertisement without the friend's consent (Bakshy et al., 2013); hiding content from certain users to measure the influence peers exert on information sharing (Bakshy et al., 2012b); and offering users an 'I Voted' button at the top of their News Feeds in order to nudge family members and friends to vote and at the same time assess the influence of peer pressure on voting behaviour (Bond et al., 2012).

While the Facebook emotional contagion study caused the largest controversy, other companies actively conduct very similar experiments. OkCupid, an online dating company, undertook an experiment that consisted of displaying an incorrect matching score to a pair of users in order to assess the effect that an artificially inflated or reduced score would have on user behaviour. A couple that was shown a 90% preferential match was an actual 20% match according to the OkCupid algorithm and an actual 90% match was shown as a 20% score (Rudder, 2014). According to the results, the recommendation was sufficient to inspire bad matches to exchange nearly as many messages as good matches typically do (Paumgarten, 2014), calling the effectiveness of the algorithm into question. Co-founder and president of OkCupid Christian Rudder responded to this criticism by claiming that: "when we tell people they are a good match, they act as if they are [..] even when they should be wrong for each other" (Rudder, 2014). OkCupid also removed text from users' profiles and hid photos for certain experiments in order to gauge the effect that this would have on user behaviour (BBC, 2014). Similar experiments are conducted by companies such as Google, Yahoo, Amazon, Ebay and Twitter, all of which have access to large volumes of user data and increasingly employ interdisciplinary teams of research scientists that approach problems beyond the scope of traditional computer science. Such teams consist of mathematicians, psychologists, sociologists and ethnographers who analyse data from user transactions, interviews, surveys and ethnographic studies in order to optimise company services (Ungerleider, 2014). Very often (as in the Facebook case) the results of their research is presented at international conferences or published in academic journals in order to stimulate discourse with the academic community. Frequently multi-

authored papers bring together company researchers and scientists at academic institutions, particularly in the Unites States. Therefore the questions of whether something constitutes industry research or academic research is much harder to answer than it may seem at the onset, with the lines deliberately being blurred by the quasi-academic environment cultivated at major internet companies.

3. ARGUMENTS FOR AND AGAINST ONLINE SOCIAL EXPERIMENTS

In the debate that followed the publication of the study, different stances were assumed by a range of actors including journalists, user rights advocates, government officials, company representatives, and academics from a variety of fields, a small and nonrepresentative selection of which is presented in the following (see table 1 for a summary). Our sample is based on a list compiled by legal scholar James Grimmelmann (2014b), who collected sources and called for references from social media users in the period after the study had been widely publicised. Grimmelmann does not specify exact criteria for the items on his list, simply referring to them as "major primary sources", but we believe that it provides a valuable overview of the types of arguments made in favour of and in opposition to the study. Many commentators reacted critically to the research, but some also expressed concerns in relation to how the study had been handled, blaming media hype and misrepresentation of the experiment for some of the negative responses. Our aim is to characterise these reactions through their implicit conceptualisations by identifying a set of recurring arguments provided in defense of the experiment. Our intent is furthermore to categorise and contrast different arguments, and to point out how they relate to the actors who benefit most from what they imply. By categorising actors along with arguments, we show that the discussion around online experiments is strongly shaped by different and at times conflicting epistemological frameworks that implicitly privilege certain viewpoints over others to attain legitimacy.

3.1 BENEFITS OF ONLINE EXPERIMENTS FOR THE INDIVIDUAL

A number of media reports stated that as part of the experiment, the News Feed had been "manipulated" (Arthur, 2014; BBC, 2014; Hill, 2014; Lennard, 2014; R. Meyer, 2014), a wording that appeared problematic to some commentators, as the News Feed is generally filtered to represent a selection of status updates curated according to algorithmic criteria (Bozdag, 2013; Gillespie, 2014). Since the News Feed is algorithmically personalised to foster user engagement in Facebook, it is difficult to judge what kind of modifications qualify as manipulations and which constitute website optimisation. Gillespie (2014) points out that Facebook's curation of user data in the News Feed is already part of the site's terms of service and its data use policy. Sandvig (2014) in turn offers a list of examples outside the News Feed in which pieces of personal communication are effectively recontextualised, for example to be used as advertisements. Facebook has stated that out of an average of 1,500 updates, the News Feed algorithm selects approximately 300 items for each user with each update (Backstrom, 2012). According to Facebook, in an unfiltered stream of information, people would be missing "something they wanted to see" (Backstrom, 2012). Since the selection of items is achieved through constant testing of alternative site designs, content selection is the product of constant experimentation. As platforms such as Facebook are generally subject to some sort of algorithmic filtering, some commentators have argued that we are ultimately faced with "a problem with the ethics of there *being* an algorithm in the first place." (Robbins, 2014)

On the other hand, research shows that most Facebook users have no precise idea about how the News Feed algorithm works, or that there is a filtering process at all (Sandvig, Karahalios, & Langbort, 2014). Contrary to intuition, an average Facebook post reaches only 12% of a user's followers (Constine, 2012). This curation is assumed to add value, and given the amount of content that is published on Facebook, it reduces clutter. But the filtering criteria cannot be controlled by users (in contrast to, for example, privacy settings), and the precise set of criteria is not transparent. Sandvig (2014) refers to the dangers of a curation that results in a distorted sense of the social context as "corrupt personalization" which he characterises as "the process by which your attention is drawn to interests that are not your own". He acknowledges that it is difficult to pinpoint inauthentic personal interests, but argues convincingly that a commercialisation of communication through algorithmic curation may conflict with user interests without the subject noticing that this is the case. Sandvig categorically differentiates between tailoring content to a user in her best interest and deriving a profit from it, and prioritising commercial content over non-commercial content in a non-transparent fashion. He interprets the latter not merely as an ethical issue to be resolved, but also as a waste of the potential of algorithmic curation.

3.2 INFORMED CONSENT AND ITS MANY INTERPRETATIONS

A second point of contention is whether or not agreeing to the Facebook terms of service constitutes informed consent to an experiment in which the News Feed is manipulated in the described way. This question has narrower legal and broader ethical implications. A clause in the terms of service covers research to improve the site and make it more attractive to users, but experts disagree on whether this covers an experimental design as the one chosen by Facebook (cf. Grimmelmann, 2014a; M.N. Meyer, 2014). The Facebook study provoked a discussion among legal scholars about the responsibility of institutional review boards (IRBs) that is still ongoing, demonstrating that massive online experiments represent unchartered territory not just from the perspective of internet companies, but also for academic regulatory bodies, who are likely to approach such experiments in markedly different ways. Grimmelmann (2014a) argues that "informed consent, at a minimum, includes providing a description of the research to participants, disclosing any reasonably foreseeable risks or discomforts, providing a point of contact for questions, and giving participants the ability to opt out with no penalty or loss of benefits to which the subject is otherwise entitled", which in his view the Facebook study did not do effectively. Taking on a similar perspective, Gray (2014) points out that Facebook could have notified the participants in a follow up email, sharing the results with them and offering them a link to the happy and sad moments that they missed in their News Feed while the experiment was underway. Facebook could also have given participants the option of deleting their data after the research was concluded, which the company did not. Jeffrey Hancock, a co-author of the study, also argued for such a "notify after" approach as a response to criticism. Hancock claimed opt-in procedures to be unrealistic for online experiments due to their ubiquity. Instead, he argued in favour of retroactively informing users after an experiment has taken place, including more information about the study, and contact information for the researchers or an ombudsman (LaFrance, 2014). Of course, user data samples based on prior consent may be less attractive to scientists than random samples (cf. Bernstein, 2014). But while the risk of influencing results by informing users in advance is acknowledged, legal scholars argue that this cannot be effectively weighed against informed consent, because "if it were, informed consent would never be viable" (Grimmelmann, 2014c).

Beyond the question of what kind of provisions are covered by the terms of service in this concrete case, informed consent more generally is seen by some experts as being in need of reform. Erika C. Hayden refers to informed consent as "a broken contract" (2014) and Mary

DeRosa describes it as being "overdue for a wake-up call" (2014, para 2). In the context of the reactions to Facebook's study, DeRosa discusses the difference between what may constitute legal agreement and ethical behaviour, asking: "Would anyone seriously argue that Facebook users *expected* this kind of manipulation of their News Feed or examination of their data for this purpose? Some consumers would knowingly consent to research like this, but it is unlikely that a single one actually did" (para 6). As DeRosa points out, a key problem is that the expectations of users are violated, rather than that consent with online experiments is necessarily *per se* rare.

Van de Poel (2011) argues that applying the principle of informed consent to social experiments in technology raises the question of whether it makes sense to ask people to consent to unknown hazards. As accepting to be a part of an experiment with unknown consequences seems to entail accepting all negative consequences emerging from the experiment, it is difficult to see how people could rationally agree to such an approach. However, Van de Poel argues, any social experiment involving ignorance and a lack of mutual understanding is unacceptable. Instead of directly trying to apply the principle of informed consent, it might be better to focus on the underlying moral concern on which consent is based. Instead of blindly accepting an agreement, the emphasis could rest on informing users about the experiment as such and the risks it entails, providing the option to stop participating if desired, and notifying participants once the experiment is stopped.

3.3 THE UBIQUITY OF ONLINE SOCIAL EXPERIMENTS

Some proponents of the study claim that online experiments should be accepted as a fact of life, since every social media company conducts them and they are without any feasible alternative (Andreessen in Sullivan, 2014). Furthermore, some researchers argue that online experiments should not be regulated by the same ethical guidelines that are applied to offline laboratory experiments as they are unique, novel and provide a great opportunity to discover human behaviour at a large scale (Bernstein, 2014; Watts, 2014). However, experiments do not always occur in a traditional laboratory setting. Van de Poel (2009) shows that certain innovations, such as nanotechnology, cannot be developed in a laboratory setting and it is hardly possible to reliably predict risks of such technologies before they are actually employed in society. It may not be feasible to reliably predict the possible hazards to all potential users of a technology, and even when we can, we may not properly express their likelihood in numbers. Van de Poel (2009, 2011) lists conditions for the acceptability of social experiments: (1) the absence of alternatives, (2) the controllability of the experiment, (3) informed consent, (4) the proportionality of hazards and benefits, (5) the approval by democratically legitimised bodies, (6) the possibility for subjects to influence the set-up, carrying out and stop the experiment if needed, (7) the protection of potentially vulnerable subjects, and (8) careful and proportional scaling of the sample size.

Clearly many online intermediaries do not adhere to these principles, mixing different types of considerations: (1) users are rarely informed before or after an experiment is conducted, (2) experiments are approved from within the company, rather than by independent bodies, (3) the subjects cannot influence or stop the experiment, nor give feedback, (4) vulnerable subjects are not protected, (5) experiments are conducted in large scale from the start, (6) the distribution of potential hazards and benefits are not clearly shown, (7) alternatives to the experiments are not considered, and (8) experiments are not subject to the control of participants in the sense that they are able to revoke or modify their participation after the experiment has started. While the ubiquity of such experiments is a result of the pervasiveness of online platforms in which users are able to interact, this hardly makes the experiments ethically less consequential. All actors involved need to jointly discuss and devise criteria for the ethics of online experiments in

accordance with existing guidelines (see for example Association of Internet Researchers, 2012). This by no means excludes users, who also can better weigh risks and benefits when they are adequately informed. In this vein, arguing for a better understanding of how social media platforms operate, Muench (2014) observes that it is "important for users to be aware of how these sites are designed to engage and reinforce our browsing behavior through evolutionary reward systems".

3.4 DIFFERENT PERCEPTIONS OF RISK IN ONLINE EXPERIMENTS

The authors of the Facebook study claimed that because Facebook did not insert emotional messages into the News Feed, but only hid certain posts for certain users, the experiment did not represent any danger to users. This argument has been opposed on the grounds that if persuasion does not happen voluntarily and if the persuader does not reveal her intentions before the persuading act takes place, this is to be considered manipulative (Smids, 2012; Spahn, 2012), making manipulation as much an issue of *intent* as much as an issue of *effect*. Others argue that involuntary persuasion is acceptable only if there is a very significant benefit for society that would outweigh possible harms (e.g. Berdichevsky & Neuenschwander, 1999). In the case of the Facebook study, it is difficult to adequately judge the benefits of the research at this point, while the harm, if only in terms of public perception, has become quite obvious. Data scientist Duncan Watts optimistically argues in *The Guardian* that online social experiments will usher in "a golden age for research" (2014), but this depends on each actor's perspective. Mary L. Gray (2014) draws a comparison to early nuclear research and experiments on human subjects, and sees data science as undergoing a learning process with regards to research ethics. In reaction to Kramer's response to the criticism, published on his personal Facebook page, individual Facebook users responded with personal accounts of emotional hardship and depression, expressing concern that Facebook would experiment on the content of the News Feed in ways that could adversely affect them. The question of risk beyond individual users seems impossible to answer without precedence, but the lack of transparency towards participants is likely to weigh more strongly in the eyes of many users than the small size of the effect reported in the study - and the details of how the filtering was conducted. Furthermore, as Kramer and colleagues point out, the impact of systematically seeking to influence users may still be strong, even if it is restricted to a small group. In a 61 million user experiment in 2010, Facebook users were shown messages at the top of their News Feeds that encouraged them to vote, pointed to nearby polling places, offered a place to click "I Voted" and displayed images of select friends who had already voted (Bond et al, 2012). The results suggest that the Facebook social message increased turnout by close to 340,000 votes. It has consequently been argued that if Facebook can persuade users to vote, it can also persuade them to vote for a certain candidate, a kind of influence which, while hypothetical, does present obvious risks (Zittrain, 2014).

3.5 BENEFITS OF ONLINE EXPERIMENTATION FOR THE SOCIETY

A popular argument among proponents of online social experiments resides in their potential benefits to society, and associated with these, the danger that negative responses could have a chilling effect on collaborations between industry and academics (Bernstein, 2014; M.N. Meyer, 2014; Tarkoni, 2014; Watts, 2014). Michelle N. Meyer (2014) makes this argument in two parts, stating first that "rigorous science helps to generate information that we need to understand our world, how it affects us and how our activities affect others", and secondly that "permitting Facebook and other companies to mine our data and study our behavior for personal profit, but penalizing it for making its data available for others to see and to learn from makes no one better off". Similar arguments are made by Watts (2014), and also by Tarkoni (2014), who contends:

"Consider: by far the most likely outcome of the backlash Facebook is currently experiencing is that, in future, its leadership will be less likely to allow its data scientists to publish their findings in the scientific literature[..] The fact that Facebook is willing to allow its data science team to spend at least some of its time publishing basic scientific research that draws on Facebook's unparalleled resources is something to be commended, not criticized."

What justifies the risks, if potential, that are incurred by large-scale online social experiments? Watts draws an analogy between the rise of empiricism during the Enlightenment and the current circumstances, arguing that "the arrival of new ways to understand the world can be unsettling". But this analogy is made at least latently problematic by the commercial interests that are at play – the opportunities of learning anything about basic human behaviour are no more pertinent than the opportunities to influence behaviour, for whatever purpose. Muench (2014) compares online social experiments to Skinnerian operant conditioning, in which strategic choices, such as exposing subjects to stimuli in randomised intervals, lead to greater engagement. To make good on the claim of societal benefit, a clearer case needs to be made for the positive impact of online social experiments, a case that is able to transcend the aim of increasing user engagement.

3.6 THE UNAVOIDABILITY OF ONLINE EXPERIMENTS

Advocates of online social experiments, such as OkCupid's CEO Christian Rudder, argue that such experiments are unavoidable, because all aspects of the design of digital platforms are shaped by constant experimentation in order to make improvements:

"OkCupid doesn't really know what it's doing. Neither does any other website. It's not like people have been building these things for very long, or you can go look up a blueprint or something. Most ideas are bad. Even good ideas could be better. Experiments are how you sort all this out." (Rudder, 2014).

He continues to argue that experiments are needed to make sure that the current algorithm works better than a random one, and that there is no alternative to such an incremental approach to optimally address user preferences. He also believes that while experiments presently cause controversies, they will be fully accepted in the future. Critics contend that the potential to innovate via experimentation must still be weighed with possible drawbacks, rather than being accepted as being without an alternative. For instance, Howell (2014) responds to Rudder arguing that he "is clearly acting wrongly, and for (at least) two reasons: 1) He is being dishonest by providing something other than what he says he will provide. Rudder thus provides a system that performs bad matches to see how people will react, instead of their claim "Our matching algorithm helps you find the right people.[1] 2) he subjects his (users) to potential harm that they have actively sought to avoid". Howell (2014) further argues that the defense of the company is disingenuous: "either OkCupid believes its sales pitch or it doesn't. If it doesn't, we already have a moral issue. If it does, then they are doing what they believe will be harmful to their customers". Grimmelmann (2014c) shares this view when proposing that, unless risks are minimal or nonexistent, researchers cannot decide that an experiment is worth a particular risk. That decision should instead be made by users.

Table 1 summarises our observations on the arguments made by the proponents and critics of the Facebook study, and similar online experiments.

Table 1: Arguments for and against online social experiments surrounding the Facebook emotional contagion study.

Argument theme	Pro experiment	Contra experiment
Benefits of online experiments for the individual	• Filtering reduces clutter • Users want filtered, rather than unfiltered content	• Users are not aware of filtering • Filtering cannot be controlled • Filtering mechanisms are not transparent
Informed consent and its many interpretations	• Accepting terms of service is a form of consent • Opt-in is annoying to users • Opt-in influences user behaviour	• Possibility of biased user behaviour does not counter informed consent • Users could be informed post-experiment • Consenting to unknown hazards is problematic
The ubiquity of online social experiments	• Experiments are essential to platform improvement • Differ from offline experiments by being unique and novel • Provide opportunities to study human behaviour at scale	• Same principles that govern offline experiments can be applied • Experiments should not be conducted at large scale when there is no need • Alternatives should be considered • Users should be able to influence or stop the experiments and provide feedback
Different perceptions of risk in online experiments	• Withholding information does not cause danger • In the long term, benefits will outweigh risks	• If participation is not voluntary, it is manipulative • Persuasion is likely to benefit the persuader at least as much as the persuaded
Benefits of online experimentation for the society	• Online experiments create new opportunities for science and society • Constant scrutiny will have a chilling effect on collaboration between industry and academia	• Exact benefits are unclear • We learn less about human interaction than about media effects • It is not sufficient to equate scientific benefit with social benefit
The unavoidability of online experiments	• Online platforms cannot be improved without experimentation • Incremental improvement is the only way to succeed	• Potential risks also need consideration • Judging risks to be minimal without having considered them is premature

4. DISCUSSION

We have aimed to show that the ethical issues raised by social experiments can be described on multiple discursive levels, depending on the roles that the discussants assume. We have shown that the problem is complex and involves interests reflected in different arguments, such as the individual and social benefits of online experiments, their ubiquity and relevance, the fact that consent is provided and that users are not exposed to any significant risks. We have shown that some of these values themselves are dependent on specific frames of reference (e.g., the attainment of status in science) and that further debate is needed to balance their relation to one

another. Perhaps our central observation is that the asymmetrical relationship between data scientists and users of social media platforms is what underpins these conflicting frames of reference. Furthermore, as long as there is no consensus regarding the ethics of online experiments that transcends a single stakeholder group, such conflicts are likely to arise again in the future, rather than abate. In this paper, we have used the Facebook experiment as a use case to discuss a range of arguments provided by different stakeholders to illustrate this conflict.

While the study has provoked strong reactions, it is worth to again point to similar research, both at Facebook and elsewhere, to clarify that this is a broader issue, rather than a singular case. In a 2012 study on information diffusion, Facebook researchers randomly blocked some status updates from the News Feeds of a pool of some 250 million users, many more than in the emotion contagion experiment. Google provides a set of tools to conduct A/B tests for website optimisation, as does Amazon. Beyond A/B testing to improve the quality of search results, issues become yet more complicated when experiments around information exposure are conducted with social improvement in mind, and without explicit consent. In research conducted at Microsoft, researchers Yom-Tov, Dumais, & Guo (2013) changed search engine results in order to promote more balanced civil discourse. In the study, the authors modified results that were displayed when users entered specific political search queries, so that subjects entering the query *obamacare* would be exposed both to liberal and conservative sources, rather than just to content biased into one ideological direction. While the researchers arguably had the best intentions, they did not notify users that their search results were being modified, neither during the experiment nor afterwards. This raises complex questions regarding the ethics of manipulation with the aim of affording social improvement. Some have claimed that when persuasion is conducted for a higher ethical goal, this can be acceptable (Berdichevsky & Neuenschwander, 1999), while others disagree (Smids, 2012; Spahn, 2012). In the light of the discrepancy between the ethical standards of academic research on human subjects and the entirely different requirements of building and optimising social media platforms and search engines, it is tempting and simplistic to single out any particular company for filtering content algorithmically. New collaborative models of joint corporate and academic research are considerably blurring the boundaries between basic and industry research, and complicating the picture of disinterested academia and result-driven commercial research.

The public outcry in reaction to the Facebook study underlines that there is a growing expectation towards more transparency regarding how content is filtered and presented, beyond assuming a 'take it or leave it'-style attitude. A company may have the interests of its users in mind, whether this goal is usability, more relevant search results, happier status updates, or a better match in dating platforms. However, users have to be able to *assess these intentions for themselves*, and evaluate the balance between their personal benefits and the interests of the company. There is a pronounced fear among publicly-funded academics that Facebook and other social media companies might limit the already fairly sparse access to their data, as they clearly see benefits in publishing studies based on unprecedented amounts of data – not solely for science, but also for their own careers. Competition for cutting-edge research results is neither unique to social media data nor surprising, but it spells out a potential conflict of interest between users whose sense of freedom and privacy is in potential conflict with scientists' interest in advancing a nascent field vying for scholarly acceptance through high-profile publications. To users, it remains largely unclear what exactly the benefits of such research may be. The argument made by Meyer, that "rigorous science helps to generate information that *we* need to better understand our world" (our emphasis), is qualified by the highly media-specific nature of such research – we learn much more about how people *react to each other on Facebook* than about human interaction in any broader, more universal sense.

After the controversy had erupted, the editor of the publication, Susan Fiske, noted the complexity of the situation, pointing out that the Institutional Review Board of the authors' institutions had approved the research, and arguing that Facebook could not be held to the same standards as academic institutions. Kramer and colleagues clearly saw their experiment in line with Facebook's continued efforts to optimise the News Feed, yet as we have pointed out, the arguments made in defense of this and similar experiments are strongly coloured by the interests of different parties, with users relatively far removed from the benefits in favour of which the proponents argue. Data science must show more convincingly that it balances the interests of scientists, companies and users to deliver on its many promises. Laboratories, regardless of their size, are governed by rules ensuring that the research conducted under their oversight is not just legal, but also ethical. Legalistic attempts to seek to cover behind the terms of service have failed to achieve this type of broad societal acceptance for what undoubtedly constitutes a new approach to science. While some researchers argue that online social experiments should not be subjected to the same ethical guidelines that are used for offline social experiments, we find the 'newness' of such experiments to lie in their potential scale, rather than in their ethics. The point is not to wring our hands about hypothetical potentials for abuse, but to carefully examine cases such as the Facebook study and ask why the reference points of users and data scientists are as different as they apparently are, and whether these differences can be reconciled in the future. Benefits for science should be balanced with possible hazards that may be caused by experiments, rather than precluding that such benefits outweigh the gains. Transparency towards users is paramount, as is seeking articulated consent for participation.

5. REFERENCES

Arthur, C. (2014). Facebook emotion study breached ethical guidelines, researchers say. *The Guardian*. Retrieved from http://www.theguardian.com/technology/2014/jun/30/facebook-emotion-study-breached-ethical-guidelines-researchers-say

Association of Internet Researchers (2012). Ethical decision-making and internet research: Recommendations from the AoIR Ethics Working Committee (Version 2.0). Retrieved from http://aoir.org/reports/ethics2.pdf

Backstrom, L. (2013). News Feed FYI: a window into News Feed. *Facebook*. Retrieved from https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed

Bakshy E., Eckles D., Yan R., Rosenn I. (2013). Social Influence in Social Advertising: Evidence from Field Experiments. EC'12, June 4–8, 2012, Valencia, Spain.

Bakshy E., Rosenn I., Marlow C., Adamic L. (2012). The Role of Social Networks in Information Diffusion. WWW 2012, April 16–20, 2012, Lyon, France.

BBC. (2014). Official complaint filed over Facebook emotion study. *BBC News Technology*. Retrieved from http://www.bbc.com/news/technology-28157889

BBC (2014). OKCupid experiments with 'bad' dating matches. 29 July 2014. http://www.bbc.com/news/technology-28542642

Berdichevsky, D., & Neuenschwander, E. (1999). Toward an ethics of persuasive technology. *Communications of the ACM*, 42(5), 51–58. doi:10.1145/301353.301410

Bernstein, M. (2014). The destructive silence of social computing researchers. *Medium*. Retrieved from https://medium.com/@msbernst/the-destructive-silence-of-social-computing-researchers-9155cdff659

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415), 295–298. doi:10.1038/nature11421

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209–227. doi:10.1007/s10676-013-9321-6

Carmody T. (2014). The problem with OKCupid is the problem with the social web. August 2014. http://kottke.org/14/08/the-problem-with-okcupid-is-the-problem-with-the-social-web

Carmody T. (2014). Why don't OKCupid's experiments bother us like Facebook's did? July 28 2014. http://kottke.org/14/07/why-dont-okcupids-experiments-bother-us

Constine, J. (2012). Your average Facebook post only reaches 12% of your friends. *TechCrunch*. Retrieved from http://techcrunch.com/2012/02/29/facebook-post-reach-16-friends/

Crawford K. (2014). The Test We Can—and Should—Run on Facebook. How to reclaim power in the era of perpetual experiment engines. *The Atlantic*. July 2, 2014. http://www.theatlantic.com/technology/archive/2014/07/the-test-we-canand-shouldrun-on-facebook/373819/

Das S., Kramer A. (2013). Self-Censorship on Facebook. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media (ICWSM). Palo Alto, CA: The AAAI Press.

DeRosa, M. (2014). How informed consent has failed. *TechCrunch*. Retrieved from http://techcrunch.com/2014/07/26/how-informed-consent-has-failed/

Gillespie, T. (2014). Facebook's algorithm — why our assumptions are wrong, and our concerns are right. *Culture Digitally*. Retrieved from http://culturedigitally.org/2014/07/facebooks-algorithm-why-our-assumptions-are-wrong-and-our-concerns-are-right/

Goel, V. (2014). Facebook tinkers with users' emotions in News Feed experiment, stirring outcry. *The New York Times*. Retrieved from http://www.nytimes.com/2014/06/30/technology/facebook-tinkers-with-users-emotions-in-news-feed-experiment-stirring-outcry.html

Gray, M. L. (2014). When science, customer service, and human subjects research collide. Now what? *Ethnography Matters*. Retrieved from http://ethnographymatters.net/blog/2014/07/07/when-science-customer-service-and-human-subjects-research-collide-now-what/

Grimmelmann, J. (2014a). Letter to Inder M. Verma, editor-in-chief, PNAS. Retrieved from http://james.grimmelmann.net/files/legal/facebook/PNAS.pdf

Grimmelmann, J. (2014b). The Facebook emotional manipulation study: sources. *The Laboratorium*. Retrieved from http://laboratorium.net/archive/2014/06/30/the_facebook_emotional_manipulation_study_source

Grimmelmann, J. (2014c). Illegal, Immoral, and Mood-Altering. How Facebook and OkCupid Broke the Law When They Experimented on Users. September 23 2014. https://medium.com/@JamesGrimmelmann/illegal-unethical-and-mood-altering-8b93af772688

Grohol, J. M. (2014). Emotional contagion on Facebook? More like bad research methods. *PsychCentral*. Retrieved from http://psychcentral.com/blog/archives/2014/06/23/emotional-contagion-on-facebook-more-like-bad-research-methods/

Hayden, E. C. (2012). Informed consent: a broken contract. *Nature, 486*(7403), 312–314. doi:10.1038/486312a

Hill, K. (2014). Facebook manipulated 689,003 users' emotions for science. *Forbes*. Retrieved from http://www.forbes.com/sites/kashmirhill/2014/06/28/facebook-manipulated-689003-users-emotions-for-science/

Howell R. (2014) OK, Stupid–OK Cupid's Ethical Confusion. August 25, 2014. http://rjhjr.com/dailysabbatical/?p=639&fb_action_ids=10100643090683099

Kramer, A. D. I. (2014). Facebook post. Retrieved from

https://www.facebook.com/akramer/posts/10152987150867796

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America, 111*(24), 8788–8790. doi:10.1073/pnas.1320040111

LaFrance, A. (2014). How Much Should You Know About the Way Facebook Works? August 20 2014. The Atlantic.
http://www.theatlantic.com/technology/archive/2014/08/how-much-should-you-know-about-how-facebook-works/378812/

Lennard, N. (2014). The troubling link between Facebook's emotion study and Pentagon research. *Vice*. Retrieved from https://news.vice.com/article/the-troubling-link-between-facebooks-emotion-study-and-pentagon-research

Lorenz, T. (2014). Plugin allows you to recreate Facebook's controversial mood-altering experiment on YOUR News Feed. *The Daily Mail*.

Meyer, M. N. (2014). Misjudgements will drive social trials underground. *Nature, 511*(7509), 265. doi:10.1038/511265a

Meyer, R. (2014). Everything we know about Facebook's secret mood manipulation experiment. *The Atlantic*. Retrieved from
http://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/

Muench, F. (2014). The new Skinner Box: web and mobile analytics. *Psychology Today*. Retrieved from
http://www.psychologytoday.com/blog/more-tech-support/201403/the-new-skinner-box-web-and-mobile-analytics

On the Media (2014). An imperfect match: transcript. July 31, 2014.
http://www.onthemedia.org/story/32-ok-cupid/transcript/

Paumgarten N. Make me a match. New Yorker. Brave New World. August 25, 2014 Issue.

http://www.newyorker.com/magazine/2014/08/25/2710913

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count: LIWC2007*. Austin, TX. Retrieved from
http://homepage.psy.utexas.edu/HomePage/Faculty/Pennebaker/Reprints/LIWC2007_OperatorManual.pdf

Robbins, M. (2014). Does OKCupid need our consent? *The Guardian*. Retrieved from
http://www.theguardian.com/science/the-lay-scientist/2014/jul/30/does-okcupid-need-our-consent

Rudder C. (2014). We Experiment On Human Beings! OkCupid Blog. Retrieved from
http://blog.okcupid.com/index.php/we-experiment-on-human-beings/

Sandvig, C. (2014). Corrupt personalization. *Social Media Collective Research Blog*. Retrieved from http://socialmediacollective.org/2014/06/26/corrupt-personalization/

Sandvig, C., Karahalios, K. G., & Langbort, C. (2014). Christian Sandvig, Karrie G. Karahalios, and Cedric Langbort look inside the Facebook News Feed. *MediaBerkman: Berkman Center for Internet & Society Podcast*. Retrieved from http://blogs.law.harvard.edu/mediaberkman/2014/07/24/christian-sandvig-karrie-g-karahalios-and-cedric-langbort-look-inside-the-facebook-news-feed-audio/

Smids, J. (2012). The voluntariness of persuasive technology. In M. Bang & E. L. Ragnemalm (Eds.), *Persuasive Technology. Design for Health and Safety* (pp. 123–132). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-31037-9_11

Spahn, A. (2012). And lead us (not) into persuasion…? Persuasive technology and the ethics of communication. *Science and Engineering Ethics, 18*(4), 633–650. doi:10.1007/s11948-011-9278-y

Sullivan, M. (2014). Facebook investor Andreessen sarcastically defends Facebook's mood experiment. *Venture Beat*. Retrieved from http://venturebeat.com/2014/06/29/facebook-investor-andreessen-sarcastically-defends-facebooks-mood-experiments/

Taylor S., Bakshy E., Aral S. (2013). Selection Effects in Online Sharing: Consequences for Peer Adoption. ACM Conference on Electronic Commerce (EC). EC'13, June 16–20, 2013, Philadelphia, USA.

Tufekci, Z. (2014). Facebook and engineering the public. *Medium*. Retrieved from https://medium.com/message/engineering-the-public-289c91390225

Ungerleider N. (2014). EBay Is Running Its Own Sociology Experiments. August 2014. http://www.fastcolabs.com/3033885/ebay-is-running-its-own-sociology-experiments

Van de Poel, I. (2009). The introduction of nanotechnology as a societal experiment. In S. Arnaldi, A. Lorenzet, & F. Russo (Eds.), *Technoscience in Progress. Managing the Uncertainty of Nanotechnology* (pp. 129–142). Amsterdam: IOS Press. doi:10.3233/978-1-60750-022-3-129

Van de Poel, I. (2011). Nuclear Energy as a social experiment. *Ethics, Policy & Environment, 14*(3), 285–290. doi:10.1080/21550085.2011.605855

Watts, D. (2014). Stop complaining about the Facebook study. It's a golden age for research. *The Guardian*. Retrieved from http://www.theguardian.com/commentisfree/2014/jul/07/facebook-study-science-experiment-research

Yarkoni, T. (2014). In defense of Facebook. *Citation Needed*. Retrieved from http://www.talyarkoni.org/blog/2014/06/28/in-defense-of-facebook/

Yom-Tov, E., Dumais, S., & Guo, Q. (2013). Promoting civil discourse through search engine diversity. *Social Science Computer Review, 32*(2), 145–154. doi:10.1177/0894439313506838

Zittrain, J. (2014). Engineering an election. *Harvard Law Review Forum*, 127, 335–341.

FOOTNOTES

1. See http://www.okcupid.com for the claim.