

Issues of Research Ethics in the Facebook Google Mood Manipulation Study: The Importance of Multiple Perspectives (full text)

By: [Michelle Broaddus, Ph.D.](#)

[A recent paper published in the Proceedings of the National Academy of Sciences](#) describes a mood manipulation experiment conducted by Facebook scientists during one week in 2012 that suggests evidence of “emotional contagion,” or the spread of positive and negative affect between people. The backlash to this publication has been significant. As two examples, Slate.com published a piece entitled [“Facebook’s Unethical Experiment: It intentionally manipulated users’ emotions without their knowledge”](#) and The Atlantic’s piece, [“Even the Editor of Facebook’s Mood Study Thought It Was Creepy.”](#)

In the interest of full disclosure, I have a personal but not close acquaintance with the lead author of the study, through conferences, and of course, Facebook. I have not been in direct contact with the lead author since the publication of the study.

So, was it unethical? One of the pillars of ethically conducted research is balancing the risks to the individual participants against the potential benefits to society or scientific knowledge.

So first, what *were* the benefits? What did we learn? Previous research (some of it using Facebook) has suggested an effect of emotional contagion, but these previous studies used observational data. In other words, there was no “manipulation.” Therefore, the researchers could not conclude a causal effect of emotional contagion, given the possibility of several other variables that could have contributed to the spreading of emotions. The only way to conclude the possibility of a causal effect of one person’s mood on another is to randomly assign participants to experience different stimuli, or “manipulate” their exposure to stimuli.

Facebook researchers did just that. Using the News Feed algorithm, on some users’ Feeds, their friends’ posts using positive words were reduced between 10% and 90%. These users were compared to a control group of users for whom their friends’ posts were reduced at random (without regard to emotional words). At the same time, a parallel experiment was conducted reducing negative words compared to a corresponding control group. The researchers found a *tiny* effect suggesting emotional contagion, with decreased subsequent posts containing positive words by 0.1% and subsequent posts containing negative words by 0.04% compared to the appropriate control groups.

So the benefit in terms of scientific knowledge is the demonstration of emotional contagion, wherein it is reasonable to conclude that being exposed to fewer positive words within Feeds *caused* a decrease in use of positive words in subsequent posts. The existence of this effect could have wide ranging contributions to scientific understanding of how other’s emotions affects us beyond Facebook to all forms of social media, and indeed all forms of media.

I will return to a discussion of the balance of risks and benefits, but first want to explore more fully the inherent risks, and the basis of charges of unethical conduct. The main source of backlash seems to be driven by the fact that this experiment was conducted without informed consent, another pillar of ethically sound research. Facebook's terms of service include the statement that "in addition to helping people see and find things that you do and share, we may use the information we receive about you ... for internal operations, including troubleshooting, data analysis, testing, research and service improvement." This is referenced in the paper itself as constituting informed consent, yet much of the backlash calls this into question.

Indeed, informed consent as defined in the [Code of Regulations for the Protection of Human Subjects](#) published by the federal government should include explanations and descriptions of the purposes, procedures, risks, and benefits of the research, as well as assurance that participation is voluntary, and who to contact in the event of harm or questions regarding the research. Obviously, this information is not included in the terms of service, therefore would not constitute informed consent under this definition. However, Facebook is a private company, and the research was not funded by a federal agency. Therefore, they are not subject to these regulations.

So should they be? Every time a grocery store chain wants to research how store configurations may influence buying decisions by randomly assigning some stores to the new configuration and comparing to stores still using the old configuration, do they need to station research assistants by the grocery carts to obtain everyone's permission before they enter the store? This is a "manipulation" every bit as much as altering what users are exposed to on Facebook (I'll return to the use of the word "manipulation" later). Such a manipulation would cause to buy more products than they "normally" would. Should

Facebook be required to obtain informed consent for any “manipulation” of News Feeds they do? That would mean that Facebook is no longer in control of their own site.

While informed consent is one of the pillars of ethically conducted research, non-researchers may be surprised to know that it is not actually always necessary for research, even when funded by federal agencies. Let’s imagine that this research had gone through a traditional, university-based institutional review board (IRB). The researchers could have realistically applied for a waiver of informed consent. The [Code of Regulations for the Protection of Human Subjects](#) states that:

“An IRB may approve a consent procedure which does not include, or which alters, some or all of the elements of informed consent set forth in this section, or waive the requirements to obtain informed consent provided the IRB finds and documents that:

- (1) The research involves no more than minimal risk to the subjects;
- (2) The waiver or alteration will not adversely affect the rights and welfare of the subjects;
- (3) The research could not practicably be carried out without the waiver or alteration; and
- (4) Whenever appropriate, the subjects will be provided with additional pertinent information after participation.”

So, as a thought experiment, what would an IRB have decided in the Facebook study? Would they have allowed a waiver of consent? Let’s dissect these criteria in turn. First, minimal risk, as

defined within the federal code of regulations, “means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.”

Note that this language does not indicate that minimal risk is not dependent on an individual’s own level of harm or discomfort. Therefore, use of Facebook, regardless of individual people’s expectations of their user experience, is reasonably defined as minimal risk. Additionally, mood manipulations are used in countless studies conducted in psychology. In the decades of ethically-sound psychological research conducted, university-based researchers have also “intentionally made thousands upon thousands of people sad” (to borrow the language of Slate’s contributor cited above) with no outcry.

Note also the use of the word “anticipated.” Would a university-based IRB have been able to anticipate this level of discomfort? Importantly, the discomfort being felt in the backlash is not a result of the experimental procedures themselves, but that the research was carried out at all, which falls outside the purview of an IRB’s decision-making ability. IRBs are not used to determine what research should or should not be conducted, only that the research is conducted under ethical regulations.

Next our hypothetical IRB must determine that the research does not adversely affect the rights and welfare of the subjects. There are no risks to privacy or confidentiality, considering that not even the researchers themselves accessed individual-level data. There were no legal or financial implications for the participants, and no one’s reputation was harmed. One could argue that the research damaged people’s overall trust in Facebook, yet they still have the right to terminate their profile.

The third criterion our hypothetical IRB must consider is whether the research could practicably be carried out if informed consent was required. The fact that it would be inconvenient to obtain consent is not enough to justify that it is not practicable. Therefore, it could be practicable for Facebook to have conducted an online informed consent process, although probably incredibly inconvenient and potentially taxing on staff and resources. However, even if informed consent itself were practicably able to be obtained, that does not mean the research itself would be practicable.

“Traditional” mood manipulation studies are often mildly deceptive, as consent forms could not include description of the specific manipulation to be used (“what you experience may make you sad”) without invalidating the procedures themselves. If people know someone is trying to make them sad, they will guard against it and this will bias the results. Therefore, the component of informed consent requiring explanation of the procedures may be altered (although admittedly not usually waived). Although the use of deception is not explicitly discussed in the federal regulations, the American Psychological Association (APA) encourages “debriefing,” a common practice in psychological research wherein researchers discuss the experiment and any deception with participants after all procedures are finished to mitigate potential negative effects. This process leads us to the final criterion to consider for a waiver of consent.

Our hypothetical IRB would need to consider if it would be appropriate for participants to be provided with additional pertinent information after participation. This criterion is often important with studies that include aspects of deception, as mentioned above. Considering the amount of backlash, participants *are* being provided with additional information from Facebook’s representatives, but only after the data were already

published. The APA suggests debriefing procedures should occur “as early as is feasible, preferably at the conclusion of [participants’] participation, but no later than at the conclusion of the data collection,” and researchers should “permit participants to withdraw their data.” Would this have been appropriate?

Should the 689,003 participants been sent a message after the conclusion of the study to let them know that if they had been feeling unusually “down” in the previous week, this may have been because of their News Feeds? Perhaps this wouldn’t have decreased the backlash, only quickened it, but quite possibly a large number of participants would not have been bothered about their involvement in the experiment, not chosen to withdraw their data, and the research could still have been practicably conducted. Indeed, considering the extremely small effect size, most participants probably would have considered it likely that they were in a control group. The fact that no one knows if they were in fact a participant or not means that participants cannot come forward to indicate whether they were consciously affected at all or were not subject to harm or discomfort because of it. Often, people enjoy participating in research studies, even when there is deception involved (Smith and Richardson, 1983).

What would our hypothetical IRB have decided? Of course there can be no way of knowing. IRBs consist of human beings, much in the way that juries are (although of course with a high degree of training and education). Two IRBs could come to different conclusions presented with the same facts in the same way that two juries could. To extend the analogy a bit further, IRBs decisions can be affected by the ability of the researchers to argue their “case.” Yet I believe an IRB could reasonably have approved a waiver of consent for this study. Other researchers, completely independently, [have come to similar conclusions.](#)

Facebook obviously conducts research on the user experience, including evaluating different manipulations of the News Feed aside from those reported in the most recent study. Yet most of this research is never published, as it is used for internal business decisions instead of attempting to contribute to scientific body of knowledge. Could these other activities not also be considered “research?” Did this study only become research once it was published?

My general sense is that “everybody knows” that Facebook manipulates different aspects of the user experience “all the time.” We as a society seem to be fine with that idea when these aspects are aimed at tailoring advertising in order to affect people’s behavior in terms of increasing product sales. Why does affecting mood seem to be qualitatively different and more upsetting? Is it because people still feel a sense of control over their purchasing behaviors that no amount of targeting advertising could overcome?

So much of advertising is itself based on manipulating our moods. Are people most uncomfortable about the idea that our moods can be affected without our conscious awareness and completely outside our control? Psychology often shows how we are unaware of the myriad ways our environment affects us. Perhaps evoking “psychology” brings to mind Orwellian mind control fantasies in a way that “market research” doesn’t, even though the underlying mechanisms and manipulations used are often indistinguishable. Or, speaking of Orwellian fantasies, is the anxiety not so much what Facebook researchers did, but how it demonstrates what Facebook in general *could* do? What is potentially comforting, and important to note, is that Facebook was under no obligation to publish their findings. The fact that they did demonstrates a level of transparency, and perhaps responsibility, rarely found in corporate research.

The reporting on this study could have exacerbated these kinds of perhaps paranoid lines of thinking. [The Atlantic piece](#) opens by calling Facebook staff “puppet masters who play with the data trails we leave online.” [The Slate piece](#) states that “nothing in the data use policy suggests that Facebook reserves the right to seriously bum you out by cutting all that is positive and beautiful from your news feed. Emotional manipulation is a serious matter, and the barriers to experimental approval are typically high.”

Considering that the study neither seriously bummed people out, nor cut out all that is positive and beautiful from people’s News Feed, this language is exaggerated and inflammatory. The effect size of the manipulation was so small that only the fact that Facebook was able to access thousands upon thousands of participants allowed for it to be detected statistically. Users posted about one fewer positive word in the subsequent week. This hardly constitutes being seriously bummed out, or would justify calling Facebook’s scientists as “puppet masters.” In fact, some researchers could contest that this effect even constitutes an effect on mood, as use of positively or negatively valenced affective words certainly may not correlate with actual mood state. However, reporting on the Facebook study in this way leads to one Twitter user I saw wondering how many negative effects Facebook had on people’s lives, from relationships dissolved to jobs quit. In reality, it was simply one less use of a word like nice, sweet, happy, pretty, or good over a week.

Additionally, even if we assume people’s moods *were* affected, in psychology, emotional manipulation is actually quite common. A cursory Google Scholar search for “mood manipulation” indicates a wide range of these types of studies to examine the effects on outcomes such as cigarette craving (Willner and Jones, 1996), overeating behaviors (Bongers et al., 2013), and dehumanization of outgroup members (Buckels and Trapnel 2013). These are all

arguably much more serious outcomes than the number of happy or sad words you post in your Facebook posts. Both of these points also illustrate the danger of using “jargon” of social science in science reporting.

You might have seen (perhaps posted on Facebook), [“10 Scientific Ideas that Scientists Wish You Would Stop Misusing”](#). One of these included the term “statistically significant.” In common language, “significant” reflects a level of importance, whereas statistical significance only reflects a very specific conclusion indicating the likelihood that the null hypothesis is true. While the Facebook mood results were “statistically significant,” I think given the modest effect size it is still debatable how “important” it is. Similarly, “manipulation” in psychology could refer to any differential stimuli researchers expose their participants to, including changing grocery store configurations. “Manipulation” in common language is much more charged, evoking a level of control over people’s minds and behaviors that is not accurate.

[The Atlantic piece cited above](#) reported on an interview with the very respected editor of the prestigious Proceedings of the National Academy of Sciences, Susan Fiske. She stated, “So, I think it’s an open ethical question. It’s ethically okay from the regulations perspective, but ethics are kind of social decisions. There’s not an absolute answer. And so the level of outrage that appears to be happening suggests that maybe it shouldn’t have been done...I’m still thinking about it and I’m a little creeped out, too.” There is something remarkable about an editor referring to research published in her journal as “creepy.”

The Atlantic piece details the murkiness of communication during the review process regarding the use of an “official” IRB. This detail underlines the importance for consideration to be given to ethical issues at all levels of research, from the design of

the research, the use (or waiver) of informed consent, and also at the point of peer review. Editors or reviewers who experience reservations about the ethical nature of research could invite the authors for further discussion of the ethical issues within the paper itself. If journals have a policy that these discussions do not impact authors' page or word limits, that could encourage authors to be more thoughtful. Perhaps including this information in the original paper could have forestalled some of the ensuing backlash. Professor Fiske's point that ethical questions are never fully resolved also should not be lost in this discussion, or overshadowed by the phrase "creeped out." These issues should involve constant questioning and discussions, especially as research evolves in response to extremely rapidly changing technologies. As this study and its associated backlash have demonstrated, these discussions should involve perspectives from participants, researchers, editors, reviewers, and science journalists. Evoking this part of Professor Fiske's interview would also have been a more appropriate source for the title of the piece.

So to return to our original pillar of ethically sound research, was it all worth it? Did the benefits outweigh the risks? As mentioned several times, the effect size was tiny. However, to quote the original paper: "These effects nonetheless matter given that the manipulation of the independent variable (presence of emotion in the News Feed) was minimal whereas the dependent variable (people's emotional expressions) is difficult to influence given the range of daily experiences that influence mood... More importantly, given the massive scale of social networks such as Facebook, even small effects can have large aggregated consequences... suggest[ing] the importance of these findings for public health."

However, the lead author has recently posted on his own Facebook feed that "[i]n hindsight, the research benefits of the

paper may not have justified all of this anxiety." Perhaps the researchers could have still practicably conducted the research with other safeguards in place to reduce this anxiety, including yes, incorporating informed consent that more specifically discusses research, perhaps as a second level of "terms of service" that are not mandatory to access Facebook, disseminating "debriefing" information to participants, and better explanations of published research's motivations and contributions to social science. However, the kind of inflammatory and misleading language used to describe these kinds of studies has ethical implications as well. I imagine this backlash will serve only to chill the efforts of Facebook's data scientists to use their considerable user base to contribute to scientific knowledge. However, it could be argued that with access to such a diverse user base it would be unethical NOT to harness that power to contribute to science.

Michelle Broaddus holds a Ph.D. in Social Psychology, and is an Assistant Professor of Psychiatry & Behavioral Medicine. [She completed a two-year fellowship with the Fordham University HIV and Drug Abuse Prevention Research Ethics Training Institute.](#) The views expressed are her own and do not necessarily represent the views of her institution, Fordham University, the Research Ethics Training Institute, the National Institutes of Health, or the United States Government.

The author wishes to thank Jennifer Kubota, Ph.D., for her feedback on an original draft of this post.

References:

Bongers, P., Jansen, A., Havermans, R., Roefs, A., & Nederkoorn, C. (2013). Happy eating: The underestimated role of overeating in a positive mood. *Appetite*, 67, 74-80.

Buckels, E. E., & Trapnell, P. D. (2013). Disgust facilitates outgroup dehumanization. *Group Processes & Intergroup Relations*, 16, 771-780.

Smith, S. S., & Richardson, D. (1983). Amelioration of deception and harm in psychological research: The important role of debriefing. *Journal of Personality and Social Psychology*, 44(5), 1075.

Willner, P., & Jones, C. (1996). Effects of mood manipulation on subjective and behavioural measures of cigarette craving. *Behavioural Pharmacology*, 7(4), 355-363.