The Cambridge Handbook of Consumer Privacy

Edited by

EVAN SELINGER

Rochester Institute of Technology

JULES POLONETSKY

Future of Privacy Forum

OMER TENE

International Association of Privacy Professionals

Pp. 214-227 most relevant

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107181106
DOI: 10.1017/9781316831960

First published 2018

Printed in <country> by <printer>

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
NAMES: Selinger, Evan, 1974- editor. | Polonetsky, Jules, 1965- editor. | Tene, Omer, editor.
TITLE: The Cambridge handbook of consumer privacy / edited by Evan Selinger, Jules Polonetsky, Omer Tene.
DESCRIPTION: Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2018. | Includes bibliographical references and index.
IDENTIFIERS: LCCN 2017054702 | ISBN 9781107181106 (hardback : alk. paper)
SUBJECTS: LCSH: Consumer protection. | Consumer protection–Law and legislation. | Consumer profiling. | Privacy, Right of.
CLASSIFICATION: LCC HC79.C63 C36 2018 | DDC 381.3/4–dc23 LC record available at https://lccn.loc.gov/2017054702

ISBN 978-1-107-18110-6 Hardback

12

Ethical Considerations When Companies Study – and Fail to Study – Their Customers

Michelle N. Meyer

Experimentation has been a hallmark of science for 400 years, but only recently – since the advents of the computer and the Internet – have relatively quick, inexpensive experiments at scale become feasible for businesses. Today, the practice of companies studying their customers is ubiquitous. In 2011, Google ran 100–200 experiments per day on its products, services, algorithms, and designs.[1] By 2016, it was running over 10,000 experiments per year.[2] If you are a Bing user, on any given day, you are participating in some fifteen experiments out of the company's 200 or so concurrent daily experiments, each of which exposes several million users to experimental conditions.[3] And if you are a Facebook user, you are a subject in some ten experiments at any given time.[4] Yet frameworks for thinking about the ethical conduct of research – long a mainstay of research throughout the academy and in industries whose research is subject to regulation by the Food and Drug Administration (FDA) – have lagged behind in most business contexts. This chapter considers the ethical implications of company attempts to learn from their customers or users. It also considers the overlooked ethical implications of companies that *fail* to do so.

A FEW WORDS ABOUT TERMINOLOGY

First, this chapter concerns, and refers to, companies' "learning activities." "Research" might be an adequate term, except that in US federal law – and perhaps colloquially – it has a particular meaning, namely, "a systematic investigation . . . designed to develop or contribute to generalizable knowledge."[5] As discussed below, many companies study their customers for reasons in addition to, or in lieu of, a desire to contribute to generalizable knowledge. For instance, companies may wish to assure or improve the quality of their products or services. Such quality assurance (QA) and quality improvement (QI) activities fall outside this definition of "research"

[1] Erik Brynjolfsson & Andrew McAfee, *The Big Data Boom Is the Innovation Story of Our Time*, The Atlantic (Nov. 21, 2011), https://www.theatlantic.com/business/archive/2011/11/the-big-data-boom-is-the-innovation-story-of-our-time/248215/ (citing Google Chief Economist Hal Varian).

[2] Hal Varian, *Intelligent Technology*, 53 Finance & Development 6, 7 (Sept. 2016).

[3] Ron Kohavi, Alex Deng, Brian Frasca, Toby Walker, Ya Xu & Nils Pohlmann, *Online Controlled Experiments at Large Scale*, Proc. of the 19th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining 1168, 1168 (2013).

[4] *Radiolab Podcast*, "The Trust Engineers," Feb. 9, 2015, http://www.radiolab.org/story/trust-engineers.

[5] Department of Health and Human Services, Policy for Protection of Human Research Subjects, 45 C.F.R. § 46.102 (d) (1991).

(and, hence, outside of federal regulations governing human subjects research), but comprise an important reason why companies study their customers. The term "learning activities" refers broadly to the range of methods and motivations for studying people, some but not all of which meet the traditional definition of "human subjects research."

Second, although this chapter generally refers to "customers" as the focus of companies' learning activities, when companies such as Google, Bing, and Facebook conduct experiments, they are generally studying their end users rather than their consumers (which tend to be advertisers). Even when this chapter refers solely to "customers," what is said will generally be applicable to users, as well. In addition, some companies study the general public – usually, but not always, because they are potential customers or users, former customers or users, employees or independent contractors, and even suppliers, distributors, and competitors. Because companies have different relationships with these stakeholders than they do with their customers or users, somewhat different ethical issues may be raised. Those nuances, however, are outside the scope of this chapter.

Third, although the traditional term for a person being studied is "subject," a term that federal regulations continue to use even as they undergo the first substantive changes in decades,[6] the trend in the research ethics literature and in research ethics practice (e.g., Institutional Review Board [IRB] deliberations) is towards referring to these individuals instead as "participants." The latter term is widely thought to be more respectful of the goal – if not quite yet the reality – of including individuals who are the focus of study as participants or even partners in the research enterprise. Nevertheless, this chapter generally uses the term "subjects." Because there are many kinds of "participants" in learning activities involving human beings, it often makes for ambiguous prose to refer to those who are the focus of study as participants. More important, the focus of this chapter is learning activities conducted by companies, and most of these proceed without anything like the study-specific, voluntary, informed consent of individuals that is typically (but, importantly, not always) involved in academic research. I defend the ethics of some such activities below, but it strains the normal meaning of the word to refer to unwitting individuals as "participants."

HOW COMPANIES STUDY THEIR CONSUMERS

In order to see when and how studying consumers implicates their privacy and other interests, it will be helpful to briefly review some of the most common types of study design that companies use, and the kinds of ethical issues these study designs raise. The broadest methodological distinction is between *observational* and *experimental* studies.

Observational Studies and the Ethics of Data Collection

A purely observational study involves passively observing customer or user behavior and other phenomena of interest as they occur in a "natural" environment, unchanged by the study. In such a study, where investigators refrain from altering or otherwise intervening in the company's normal business practices, ethical concerns will generally center on potential loss of informational privacy as a result of data collection.

[6] Department of Health and Human Services et al., Final Rule: Federal Policy for the Protection of Human Subjects, 82 Fed. Reg. 7,149 (Jan. 19, 2017) (to be codified at 45 C.F.R. 46).

One factor that may help determine the level of ethical concern involved in data collected is the *kind of data* that are collected. Some data are more sensitive than others. Moreover, some data used in a study may already be collected by the company. For instance, customer billing, sales, or inventory information may be analyzed for research purposes. In other cases, however, a company may newly collect additional kinds of data for the purpose of learning. Without appropriate transparency, the collection of additional data may violate a company's terms of service (ToS) or be unexpected by users, whose trust in the company may thereby be undermined. Even when study data have already been collected for routine business purposes, the repurposing of that data for other purposes may be perceived by users to be (or may actually be) problematic.

Different *methods of data collection* also raise different privacy and other ethical concerns. Perhaps the most benign data collection methods are surveys, interviews, and focus groups. Unless the purpose of this data collection is actively hidden from participants, the fact of data collection is necessarily revealed to participants, and consumers may usually choose to participate or not, or to stop at any point. Ethical issues can, of course, be raised by the way this data is stored, how the company uses it (both during and after the study), and with whom the company shares the data (both within and beyond the company).

On the other end of the spectrum, data collection tends to raise the most acute privacy concerns when it is conducted directly by human investigators who watch consumers in real time or who record the relevant phenomena for later viewing. Data collection through audio and visual recordings can involve two particularly troublesome features. First, they are somewhat indiscriminate, and may capture data beyond that which is sought, including data from third party "bystanders." Second, recordings create potentially permanent archives to which investigators (or others) can return repeatedly; this is a substantive difference from mere one-time observations of behavior that are imperfectly preserved in the researcher's memory.

One notable example of proposed human-driven data collection for research purposes occurred in late 2016, when Evernote – the popular platform for creating and sharing notes, photos, and other documents – announced a new policy under which designated employees could read user notes in order to "spot check" the company's efforts to use machine learning to help users automate functions ("like creating to-do lists or putting together travel itineraries").[7] According to the announcement, Evernote users would have been able to opt out of the policy (if they became aware of it – hardly a sure thing). And Evernote did seek to limit the privacy implications of its planned policy: only a few, carefully selected employees would have access to user notes; the company would randomly select notes for inspection (rather than targeting certain users); and the notes would be stripped of names and "any personal information."[8] Still, there was considerable pushback from users, and Evernote quickly reversed course, announcing that users will have to expressly opt *in* before Evernote employees may examine their notes to further the company's efforts to improve the user experience.[9]

[7] Chris O'Neill, *A Note From Chris O'Neill about Evernote's Privacy Policy*, Evernote Blog, Dec. 15, 2016, https://blog.evernote.com/blog/2016/12/15/note-chris-oneill-evernotes-privacy-policy/. Designated employees had always been (and remain) able to access user notes in other, limited circumstances, including "responding to a warrant, investigating violations of [Endnote's] Terms of Service such as reports of harmful or illegal content, and troubleshooting at the request of users." *Id.*

[8] *Id.*

[9] Greg Chiemingo, *Evernote Revisits Privacy Policy Change in Response to Feedback*, Evernote Blog, Dec. 15, 2016, https://blog.evernote.com/blog/2016/12/15/evernote-revisits-privacy-policy/.

Evernote's announced policy was notable, in large part, for how unusual it was. Much more typically in contemporary business, data are collected by computers and analyzed by humans only in aggregate or anonymized form. Companies typically do not, for instance, directly observe individual customers at their computers as they click an online ad (or not) or in stores as they select one brand of cereal from the grocery aisle over another. Instead, most companies analyze data that *represent* these behaviors. For instance, a company may conduct an A/B test, in which different users are randomly assigned to see one of two versions of an ad, and user click rates for each ad are tracked. Or, during a given time span, a company may compare cereal sales in stores randomized to display boxes relatively higher on store shelves to cereal sales in stores randomized to display boxes relatively lower (say, where they might attract the attention of children).

Experiments and the Ethics of Interventions

Unlike observational studies, in an *experiment*, the investigator actively manipulates for investigational purposes the exposure received by different groups of subjects. Company experiments, for instance, frequently involve altering the existing product or service environment in order to test the effect of that change on customer behavior. In a true experiment, individuals are randomly assigned by the investigator to two or more different exposure conditions. At least one condition is a "treatment" designed to have an effect, and at least one serves as a control. Randomized, controlled trials (RCTs) are generally considered the "gold standard" for establishing a causal relationship between a treatment and a measured outcome. When subjects are sorted into different study conditions by their own preferences or other nonrandom factors, as is the case with observational studies and quasi-experiments, it is likely that the groups themselves will differ, and it may be those group differences, rather than the treatment, that cause any observed differences in outcome.

Between purely observational studies and true experiments lie a variety of other study designs. For instance, a company may impose a treatment on an entire population (or a sample thereof) and collect data both before and after administering the treatment. Or a company may study the outcomes of a "natural experiment" in which a factor external to the investigators (e.g., varying state laws) "assigns" customers to different conditions, approximating true random assignment.

Because all studies depend on some form of data collection, the informational privacy concerns that arise in observational studies (discussed earlier) are also present in experimental and quasi-experimental studies. Experiments pose additional ethical concerns, however, arising from the interventions they impose on some subjects. For instance, treating some consumers differently from others – as is necessarily done in randomized studies – may be said to raise equality concerns. Similarly, assigning customers to one condition or another necessarily deprives them of the ability to sort themselves into conditions. When an experiment proceeds without customers' consent (as is often the case), it may be thought that the experimenter's assignment deprives customers of an important autonomy interest (sometimes characterized as "decisional privacy").

Normally, however, randomizing people to different "treatments" will be ethically problematic on equality grounds only when there is good reason (i.e., evidence) *prior to conducting the experiment* to believe that the risk-benefit profile of one arm is significantly different from that of the other(s) or from what subjects would experience outside of the trial. In clinical research, it is traditionally (if controversially[10]) regarded as an ethical prerequisite of a randomized, controlled

[10] *See infra* note 65.

trial for the relevant expert community to be in collective "equipoise" about which condition, on net, most advances patients' interests. Similarly, depriving subjects of a choice between experimental conditions normally will be ethically problematic on autonomy grounds only when individuals have an ethical right to make that choice outside of the experiment.

WHY COMPANIES STUDY THEIR CONSUMERS

So far, we have discussed some of the major methods by which companies learn about their customers and the typical ethically relevant effects – or consequences – for customers of each method. We have yet to consider companies' *reasons* for studying their customers. Few, if any, ethicists believe that intentions are the only thing that counts when considering whether an act (or omission) is morally justified. But count they often do: accidents are generally not as blameworthy as intentional acts, for instance. So, too, with company research: the different purposes behind such research raise different ethical considerations. Of course, companies study their customers for many reasons – and any particular study is often conducted for multiple reasons. With that caveat aside, we now turn to the range of considerations that motivate company research and the ethical relevance of different motivations.

Some company research is conducted, at least in part, in order to contribute to *generalizable knowledge*. Although developing or contributing to generalizable knowledge is typically thought of as the province of academic research, companies such as Facebook and Microsoft compete for the best data scientists, many of whom have academic pasts (and perhaps academic futures or part-time presents) and want to pursue at least some research that is publishable in academic journals. Competitive hiring aside, it is also not implausible to argue that such companies as these have an obligation to conduct some basic research. Some effects are small enough that they require very large study samples to detect. Outside of governments, only companies such as Facebook have large enough user bases for some studies. Some company research may be conducted with this sense of obligation in mind.[11]

Nevertheless, although contributing to generalizable knowledge is an important social good, it may be orthogonal to the consumer's relationship to the company. In the quintessential basic research study, any benefits that result are likely to accrue to others (such as society at large) in the future, while a study's risks are concentrated on subjects. That mismatch in the distribution of research-related risks and expected benefits does not render basic research unethical, but it makes a moral difference. In the ethical and legal frameworks that govern most academic research, this mismatch is addressed in two ways: by a requirement that a study's risks are "reasonable" in relation to its expected benefits,[12] and by a strong default rule (subject to important exceptions) that subjects provide voluntary, informed consent to participating in research,[13] usually as a form of altruism. Study-specific informed consent of this kind is not usually feasible in the kind of continuous, robotically conducted research in which most large companies engage. Yet if a company fails to make it clear to its customers or users that basic research will be conducted, their expectations for what data the company collects, how it uses that data, and how and why it may manipulate the user environment may be upended.

[11] Data scientists at large companies are also making advances in statistics and research methodologies, itself a public good.

[12] 45 C.F.R. § 46.111(a)(2). [13] *Id.* § 46.116.

It is relatively rare, however, that a company's research has no bearing at all on its products and services. Instead, the vast majority of company research aims at either new product development or the quality assurance (QA) or quality improvement (QI) of products and services. While the noble pursuit of generalizable knowledge will generally be orthogonal to the company-customer relationship, research that seeks to develop new products and services may not be, and QA/QI activities are by definition related to the business that brings the company and its users together and are the most likely to directly benefit participants.[14]

Before concluding that the same population that bears any risks of a particular QA/QI activity also stands to enjoy its benefits, however, it is important to distinguish QA/QI activities that aim to assure or improve the quality of a company's products or services *from the customer's or user's point of view* from those that aim to increase business efficiency or maximize profits in ways that are unlikely to benefit customers (and may even set back their interests). For instance, Facebook makes money by selling ad space. Presumably, the longer its users stay on the platform, and perhaps the more they "engage" with the platform (through "likes," shares, and comments), the more likely they are to see and click on these ads – something the company can tout to advertisers when it sets ad prices. But the time that users spend on Facebook and their level of engagement are also reasonably good (if imperfect) proxies for user enjoyment. In this case, then, research designed to determine how best Facebook can maximize user engagement plausibly benefits both Facebook and its users.

In other cases, however, company and user interests diverge, even when it may appear at first blush that they dovetail. Consider, for instance, the differences between people's first order preferences (what we normally mean by preferences) and their second order preferences (the preferences they have *about* their own preferences). Whenever we want to lose weight or remain healthy but also want a second helping of dessert, we have a conflict between our first order preference for the immediate gratification of cake and our second order preference for ranking health above gastronomic delights. To a large extent, it is reasonable to hold individuals responsible for their indulgences, and not the companies who create the products and services in which some customers (over)indulge. If someone were, say, to procrastinate writing a book chapter by indulging in Facebook browsing, we probably ought not to blame Mark Zuckerberg, and still less his employees who conduct A/B tests.

But companies can learn when their customers are most vulnerable to succumbing to such first order preferences, and deliberately use this information to frustrate their second order preferences. For example, a company may try to identify the times of day when users are most susceptible to indulgences (whether sweets or in-app purchases that help boost gamers' play) and offer those indulgences (or ads for them) then.

Some companies take an even more individualized approach. Gary Loveman, who holds a PhD in economics from MIT and is a former Harvard Business School professor, instilled a culture of experimentation[15] at Caesars Entertainment when he took over as COO (he would later become President and CEO before moving on to Aetna). Loveman famously made use of

[14] This is more likely to be true the more that a business's customers or users are returning rather than transient. Returning, but not transient, customers may benefit from QA/QI activities that result in improvements to a product or service that they will use in the future.

[15] According to Loveman, "there were three ways to get fired at Harrah's: steal, harass women, or institute a program or policy without first running an experiment." Jeffrey Pfeffer & Robert I. Sutton, Hard Facts, Dangerous Half-Truths, and Total Nonsense: Profiting from Evidence-Based Management 15 (2006). *See also* Brynjolfsson & McAfee, *supra* note 1 (noting that Loveman is known for saying, "There are two things that will get you fired [at Caesar's]: stealing from the company, or running an experiment without a properly designed control group.").

the large quantities of customer data Caesars collects in its Total Rewards card program to predict which marketing strategies would be effective for different individuals. The results were industry-changing: as of 2014, Caesars boasted 45 million Total Rewards members worldwide.

Soon, however, came charges that Caesars' marketing is predatory and exploits those who are addicted to gambling. Caesars responded to the charge as follows:

> We look for ways to attract customers, and we make efforts to maintain them as loyal customers. When our customers change their established patterns, we try to understand why and encourage them to return. That's no different than a hotel chain, an airline, or a dry cleaner.[16]

The difference harkens back to the distinction between first and second order preferences. Few people have deep second order preferences that conflict with their first order preferences for, say, inexpensive or convenient dry cleaning (unless the cheapest, most convenient dry cleaner engages in practices to which one is ideologically opposed). But too many struggle to reconcile the rush of the quick win with their need to hold down a job, maintain their marriages, or pay for their kids' education.

Similarly, companies may conduct research to determine whether the business can extract more money from its customers without offering anything more of value. Again, Loveman provides a particularly stark example. He wanted to know whether patrons of his casino could tell the difference between a slot machine that pays out only 5 percent of what players put in, as machines in some competitor casinos in Atlantic City did, and one that pays out a slightly more generous 7 percent, as his own casino's slot machines then did. His hypothesis was that they could not, since he calculated that someone would have to play the slots for approximately forty hours before being confident of such a difference, and he suspected that vanishingly few, if any, customers played so long in one sitting. To rigorously test that hypothesis, however, he put two groups of otherwise identical machines on the casino floor: one with the usual 5 percent hold and the other with a 7 percent hold. Sure enough, customers used the 7 percent hold slot machines just as often as the 5 percent machines, but they lost more money when they played those machines, which the house pocketed as a 40 percent increase in gross profit.[17] Caesars used the results of the experiment to set a new policy of 7 percent holds in all slot machines throughout its casinos, leading to over $300 million in profit a decade later.

Of course, most companies extract as much profit from their customers as the market will bear, and the point of this example is not necessarily to suggest that a 7 percent hold rate is unethical (though the addictive nature of gambling for some people makes this business different from many others). Moreover, assuming for the sake of argument that a 7 percent hold rate is ethical (a questions of business ethics well outside the scope of this chapter), discovering through an experiment, as opposed to through luck or trial and error, that the market will bear a 7 percent hold rate does not magically convert that ethical practice to an unethical one. Rather, the point for present purposes is merely that not all company learning activities serve customer interests equally well.

[16] Transcript of Episode 466: *Blackjack*, This American Life, June 8, 2012, http://www.thisamericanlife.org/radio-archives/episode/466/transcript.

[17] Ryan Jacobs, *Turning Data Into Profit at Caesars Palace*, Inside Ops, March 14, 2014, http://insideops.com/rljacobs/turning_data_into_profit_at_caesars_palace/.

US FEDERAL REGULATION OF HUMAN SUBJECTS RESEARCH: THE COMMON RULE

With the increase in corporate experimentation and the considerable public attention that some of these experiments have recently received[18] has come increased interest in how company experiments and other forms of learning can be conducted ethically. One obvious source of potential wisdom in this regard is the governance of academic research with human participants. In the United States, much human subjects research is governed directly or indirectly by a set of federal regulations best known as the Common Rule.[19]

The Common Rule reflects the work of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. That ad hoc commission was created by the National Research Act of 1974[20] to distinguish research from practice, identify the ethical principles that should govern the conduct of research with human participants, and recommend specific rules, including the appropriate role of risk-benefit assessment, that local Institutional Review Boards (IRBs) should apply in reviewing and approving research before it may proceed. The Commission's most influential product is known as the Belmont Report, which identified the aforementioned principles as respect for persons, beneficence, and justice.[21]

In prospectively reviewing, and determining whether or not to approve, research, IRBs have three sets of tasks. The first involves *voluntary, informed consent*, which is the main means by which the principle of *respect for persons* is applied to research. The Common Rule provides a list of informational elements (such as a statement that participation is voluntary and a description of the risks and expected benefits of, and individuals' alternatives to, participation) that ordinarily must be disclosed to prospective subjects to help them decide whether or not to participate.[22] Before research may proceed, an IRB must either review and approve the written informed consent that the investigator has drafted to ensure that it accurately and accessibly conveys all appropriate information,[23] or it must determine that Common Rule criteria for waiving some or all of the required elements of informed consent have been met (on which more later).[24]

[18] For instance, Facebook's mood contagion experiment and OkCupid's matching algorithm experiment, discussed later in the chapter, received very strong responses from many privacy and technology critics, regulators, ethicists, and others.

[19] Department of Health and Human Services, Policy for Protection of Human Research Subjects, 45 C.F.R. 46 (1991).

[20] National Research Act, Pub. L. 93–348 (July 12, 1974).

[21] National Commission, Ethical Principles and Guidelines for the Protection of Human Subjects of Research (1978) (hereinafter "Belmont Report," so named because the commission meeting at which its rough outlines were hashed out took place at the Smithsonian Institute's Belmont Conference Center). The commission's staff philosopher, Tom Beauchamp, drafted the Belmont Report. At the time, Beauchamp was simultaneously in intense collaboration with Jim Childress in writing the first edition of what would become perhaps the most influential treatise on method in bioethics, Tom L. Beauchamp & James F. Childress, Principles of Biomedical Ethics (7th ed. 2012) (1979). Although *Principles* and the Belmont Report clearly have much in common, Beauchamp was not able to convince the Commission that the Belmont Report should be written in every aspect exactly as he advised, and as a result there are important differences between the two frameworks. For instance, *Principles* adds to the Belmont principles a fourth principle of non-maleficence, which more completely fills out beneficence by recognizing the importance not only of securing benefits and protecting from harm but also of avoiding harm. In addition, *Principles* refers to respect for autonomy, rather than to the broader principle of respect for persons. Respect for autonomy refers to the prima facie obligation to respect the decision-making capacities of autonomous individuals. The Belmont Report's "respect for persons" awkwardly combines in one principle both that tenet and the idea that *nonautonomous* individuals should be protected. In *Principles*, protection of individuals who are incapable of self-determination is captured by the principles of beneficence, nonmaleficence, and justice, rather than fused with a principle that aims primarily at respect for autonomy.

[22] *Id.* § 46.116(a)–(b). [23] *Id.* § 46.111(a)(4)–(5). [24] *Id.* § 46.116(c)–(d).

The second set of tasks that IRBs perform, which is an application of the principle of *beneficence*, is *risk-benefit analysis*. Before it may approve a study, an IRB must find that any research-related physical, psychological, social, legal, or economic risks to subjects are "reasonable in relation to" the study's expected benefits to subjects (if any) and to society (in the form of knowledge production), and it must ensure that these risks are minimized.[25] An IRB must also ensure "adequate protections" to protect subjects' privacy, "when appropriate."[26] And it must – again "when appropriate" – ensure that data collected during the research will be monitored on an ongoing basis to ensure that research risks don't turn out to be higher than expected, possibly necessitating termination of the study.

The third set of tasks broadly concerns *justice* in the *distribution of research risks and expected benefits*. An IRB must ensure that the selection of subjects is "equitable."[27] Investigators should not, for instance, choose vulnerable subjects simply because they are vulnerable or otherwise constitute a convenient sample. IRBs must also ensure that "additional safeguards" are in place if some or all study subjects are likely to be "vulnerable to coercion or undue influence."[28]

By its terms, the Common Rule applies to all activities that (1) meet the regulatory definitions of "research" involving "human subjects," (2) are not exempt from the policy,[29] and (3) are either conducted or funded by one of fifteen federal departments and agencies[30] that have adopted the Common Rule. (Anyone conducting human subjects research involving an item subject to regulation by the Food and Drug Administration – e.g., investigational drugs, medical devices, or biologics – is subject to a similar set of regulations that also includes a requirement of IRB review.[31]) When an institution receives funding to conduct human subjects research, it enters into an agreement with the federal government, called a Federalwide Assurance (FWA), in which it agrees to subject the funded study to IRB review.

However, the current version of the FWA invites institutions to voluntarily subject *all* human subjects research conducted at the institution to IRB review regardless of the source (or even existence) of funding. Institutions that accept this invitation are colloquially referred to as having "checked the box," and historically, somewhere between 74 percent and 90 percent of institutions with an FWA have done so (though with a declining trend).[32] Because the vast majority of colleges, universities, academic medical centers and healthcare systems conduct at least some federally funded human subjects research, and because the majority of those institutions elect to check the box, the vast majority of human subjects research conducted at these institutions is subject to the Common Rule through the FWA.

Federal regulators recently announced their intention to revise the FWA to omit the possibility of checking the box, but it is unclear that this will have a dramatic impact on the scope of IRB review. Even today, virtually every academic and nonprofit research institute requires IRB review of all of its human subjects research as a matter of institutional policy. That is unlikely

[25] *Id.* § 46.111(a)(1)–(2). [26] *Id.* § 46.111(a)(7). [27] *Id.* § 46.111(a)(3). [28] *Id.* § 46.111(b). [29] Id. §46.101(b).

[30] The fifteen federal departments and agencies that have codified the Common Rule into their own regulations are the Departments of Agriculture, Commerce, Defense, Education, Energy, Health and Human Services, Housing and Urban Development, Justice, Veterans Affairs, and Transportation and the Agency for International Development, Consumer Product Safety Commission, Environmental Protection Agency, National Aeronautics and Space Administration, and National Science Foundation. In addition, the Central Intelligence Agency complies with the Common Rule pursuant to an Executive Order and the Food and Drug Administration (FDA) has promulgated its own regulations, which are as similar to the Common Rule as possible, commensurate with the limits of the FDA's enabling statute.

[31] *See* Food and Drug Administration, 21 C.F.R. parts 50, 56, 312, and 812.

[32] *See* Carol Weil et al., *OHRP Compliance Oversight Letters: An Update*, 32 IRB 1 (2010) ("informal review of a sample of institutions" showed that in 2000, more than 90% of domestic institutions had agreed to extend the regulations to all human subjects research, while in 2010 only 74% of these institutions did so).

to change if and when the FWA no longer offers institutions the opportunity to opt into oversight of all its human subjects research by the federal government; the difference is likely to be only that violations would be reportable solely to the relevant authorities at individual institutions, and not also to the federal government.

In sum, the Common Rule is a well-established ethical-regulatory framework for the governance of research involving human subjects. Given the significant overlap between the academy and company data scientists and other researchers – after the US government, Microsoft reportedly employs the largest number of anthropologists in the world – it is also a framework with which many people conducting company research are already familiar.[33]

CHALLENGES IN APPLYING THE COMMON RULE TO COMPANY RESEARCH

There are several problems, however, with simply importing the Common Rule into the company research context.

Limited Substantive Scope of the Common Rule

First, even if the Common Rule applied to research conducted and funded by private companies, much of the research that companies conduct either doesn't meet the Common Rule's definition of "human subjects research" or is human subjects research that is nevertheless exempt from the Common Rule.

The Common Rule defines "research" as "a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge."[34] Activities that are instead designed solely to allow companies to measure and either assure or improve the quality of its products or services fall outside the scope of the Common Rule, even if these activities are otherwise identical to research designed to contribute to generalizable knowledge.

Determining when an activity is research, QI/QA, or both is notoriously difficult. Here is a taste of how federal regulatory guidance attempts to thread the needle: an institution engages solely in QI not subject to the Common Rule if it implements a practice that is either untested or "known to be" effective (in some context), with the aim of improving the quality of that institution's work (say, patient care or user experience) and collects (patient or user) data about this implementation for "clinical, practical, or administrative purposes," including in order to determine whether the practice had the expected effects. If, however, these data are *also* collected "for the purpose of establishing scientific evidence to determine how well the intervention achieves its intended results," then the QI project "may also constitute nonexempt human subjects research."[35]

To see how the research/QI distinction might play out in the (nonmedical) business context, consider the infamous Facebook mood contagion experiment.[36] The purpose of that experiment

[33] Graeme Wood, *Anthropology Inc.*, The Atlantic (Feb. 20, 2013), http://www.theatlantic.com/magazine/archive/2013/03/anthropology-inc/309218.

[34] 45 C.F.R. 46.102(d).

[35] Off. for Hum. Research Protections, *Quality Improvement Activities FAQs*, U.S. Dep't Health & Hum. Serv., https://www.hhs.gov/ohrp/regulations-and-policy/guidance/faq/quality-improvement-activities (no date; last accessed May 6, 2017).

[36] For background on and a detailed ethical analysis of this experiment, see Michelle N. Meyer, *Two Cheers for Corporate Experimentation: The A/B Illusion and the Virtues of Data-Driven Innovation*, 13 Colo. Tech. L.J. 273 (2015).

was to attempt to better understand the effects of exposing users to both positive and negative content in their News Feeds. The mood contagion experiment gave Facebook potentially useful information about the negative effects of its News Feed on users and how to avoid them – a classic QI aim. It also contributed to the academic literature in social psychology pertaining to the general phenomenon of mood contagion. Although mere publication of results, without more, does not render an activity research,[37] the fact that the article reporting the results of the experiment situated them within this broader academic conversation is suggestive of a research, rather than solely a QI, purpose.[38]

Yet the results could have been written up a bit differently – or not published at all – and easily have constituted a pure QI activity not subject to the Common Rule's requirements of IRB review, informed consent, and risk-benefit analysis. The same will be true of virtually all company research.[39] Insistence that company research be subject to the Common Rule will likely only push companies to simply conduct those learning activities as QI, depriving the public of both generalized knowledge and knowledge of the effects of company products and services, and generally reducing transparency about how companies study their users.[40]

Even if an activity involves "research," the Common Rule only applies if that research involves human subjects. Under the Common Rule, a "human subject" is "a living individual about whom an investigator . . . conducting research obtains (1) Data through intervention or interaction with the individual, or (2) Identifiable private information."[41] Most company research does not involve interaction with consumers or users but is instead conducted at a distance. Much company research *does* involve intervention, which includes not only physical interventions common in clinical trials such as drawing blood or administering drugs but also "manipulations of the subject or the subject's environment that are performed for research purposes."[42]

What about research that involves no interaction or intervention, only obtaining data about subjects? For such research to involve "human subjects," the data must be "private," which "includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record)."[43] How this

[37] *See* Off. for Hum. Research Protections, *Quality Improvement Activities FAQs*, U.S. Dep't Health & Hum. Serv., https://www.hhs.gov/ohrp/regulations-and-policy/guidance/faq/quality-improvement-activities (no date).

[38] *See* Adam D. I. Kramer, Jamie E. Guillory & Jeffrey T. Hancock, *Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks*, 111 Proceedings of the Nat'l Acad. of Sciences 8788 (2014).

[39] Perhaps the chief reason why some company research does in fact appear to be designed to contribute to generalizable knowledge (in addition to improving business quality and otherwise guiding business decisions) is that the best data scientists tend to have academic ambitions and therefore are more likely to accept positions in companies that allow them to publish some research in general academic journals such as *PNAS*.

[40] *See* Michelle N. Meyer, John Lantos, Alex John London, Amy L. McGuire, Udo Schuklenk & Lance Stell, *Misjudgements Will Drive Social Trials Underground*, 511 Nature 265 (July 16, 2014).

[41] 45 C.F.R. 46.102(f).

[42] The 2018 Common Rule, however, newly exempts "benign behavioral interventions." Basic HHS Policy for the Protection of Human Research Subjects, 45 C.F.R. § 46.104(d)(3) (2017) (as of this writing, this provision of the revised Common Rule will go into effect on January 19, 2018; it appears in the Code of Federal Regulations following the version of the subpart currently in effect as a convenience to users). The exemption requires, among other things, that the subject "prospectively agrees to the intervention and information collection." *Id.* Until such time as regulators provide guidance interpreting this new exemption, it is unclear whether broad "prospective agreement" to future, unspecified interventions included in something like a ToS would suffice, or whether the exemption requires study-specific agreement. Other exemptions are discussed later in the chapter.

[43] 45 C.F.R. 46.102(f).

rather odd definition of private data applies to the kind of data companies most often use in research is unclear. For example, information knowingly given to a company for business purposes would not seem to qualify as private (since we expect a company to record information given to them), even if its use for company research purposes is contrary to customer expectations.[44]

Even if research data are private, they must also be individually "identifiable" if an activity that only involves obtaining data (without intervention or interaction) is to constitute research with "human subjects." Under the Common Rule, data are "identifiable" only if the subject's identity is either directly associated with the data or is "readily ascertainable" by the researcher.[45] Regulatory guidance further explains that coded data are *not* readily ascertainable if the researcher is denied access to the code key.[46] The vast majority of consumer data used in company research that is not aggregated or anonymous will be "coded" by customer ID, device ID, cookies, and the like. Under federal guidance interpreting the Common Rule, so long as there is an agreement in place under which those conducting research with such data will under no circumstances have access to the key linking these pseudonymous identifiers to actual consumer identities, no "human subjects" are involved and the Common Rule does not apply to the research.

Most data privacy scholars would find this result inadequate for at least two reasons. First, nonidentifiable data can sometimes to be reidentified. Some scholars have argued that HIPAA de-identification is inadequate to protect data,[47] and the Common Rule's definition of "non-identifiable" does not set even so high a bar as that. Although federal regulators are aware of the possibility for reidentification, they were unable to achieve consensus among the many federal Common Rule departments and agencies about how to address the problem. As a result, the 2018 Common Rule kicks the ball down the road, merely requiring agencies to regularly

[44] *Cf.* Fed. Trade Comm'n, *Privacy Online: A Report to Congress* 7 (June 1998) (describing the Fair Information Practice Principles (FIPPs), including Notice/Awareness, which calls on businesses to disclose their data practices to consumers, usually including "the uses to which the data will be put"); Org. for Econ. Co-operation & Dev. (OECD), *OECD's Privacy Framework* 14 (2013) (orig. 1980) (calling on businesses to respect principles of Purpose Specification ("The purposes for which personal data are collected should be specified no later than at the time of data collection and the subsequent use limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.") and Use Limitation ("Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with [the purpose specification principle] except: a) with the consent of the data subject; or b) by the authority of law.")); White House Report, *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy*, 4 J. Priv. & Confidentiality 95, 97, 109–113, 136–7 (2012) (proposing a Consumer Privacy Bill of Rights, including a principle of Respect for Context that calls on companies to "collect, use, and disclose personal data in ways that are consistent with the context in which consumers provide the data"); White House, Administration Discussion Draft: Consumer Privacy Bill of Rights Act of 2015 (Feb. 27, 2015) (requiring that out-of-context data collection, e.g., for research, require either affirmative consumer consent in response to a clear and prominent prompt or approval by an FTC-approved "Privacy Review Board" that finds that such opt-in consent is impractical and that privacy risks have been minimized and are outweighed by "substantial benefits" to those beyond merely the business, e.g., to consumers or society).

[45] 45 C.F.R. 46.102(f).

[46] Off. for Hum. Research Protections, *Coded Private Information or Specimens Use in Research, Guidance*, U.S. Dep't Health & Hum. Serv., Oct. 16, 2008, https://www.hhs.gov/ohrp/regulations-and-policy/guidance/research-involving-coded-private-information/index.html.

[47] *See, e.g.*, Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. Rev. 1701 (2010). *But see* Daniel Barth-Jones, *The Debate Over 'Re-Identification' of Health Information: What Do We Risk?*, Health Affairs Blog, Aug. 10, 2012, http://healthaffairs.org/blog/2012/08/10/the-debate-over-re-identification-of-health-information-what-do-we-risk.

reconsider the Common Rule's definition of "identifiable" and whether certain research techniques or technologies (such as genomic sequencing) inherently yield identifiable data.[48]

Finally, several categories of human subjects research are exempt from the Common Rule despite meeting these definitions. For instance, research involving the collection or study of existing data is exempt if the data are publicly available or recorded by the researcher in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.[49] Research involving surveys, interviews, and observations of public behavior[50] is exempt if either data are recorded such that subjects cannot be identified directly or through identifiers linked to subjects or, when data are identifiable, if their disclosure would not place subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.[51] Some have worried about "the 'extraordinary' lengths to which food manufacturers go to scientifically engineer craving,"[52] but taste and food quality evaluation and consumer acceptance studies, too, are exempt from the Common Rule.[53] As for the broader social implications of manufacturing ever more addictive food during an obesity epidemic, the Common Rule explicitly instructs, "The IRB should not consider possible long-range effects of applying knowledge gained in the research (for example, the possible effects of the research on public policy) as among those research risks that fall within the purview of its responsibility."[54]

Consent

The Common Rule typically requires researchers to obtain subjects' informed consent to participate in a study. The traditional form such consent takes is often a lengthy written document, signed by the subject, that describes a specific study's purpose, risks, and potential benefits. This kind of formal, study-specific consent is infeasible for the continual A/B testing in which companies such as Facebook, Google, and Bing engage, and for much other business research. A pop-up window that required Facebook users to read and agree to a consent document every time they use the service, for instance, would quickly spell the demise of the platform. Similarly, much company research is behavioral – i.e, designed to investigate which aspects of the retail environment affect customer behavior – and disclosing the details of these studies in advance would badly (often irreparably) bias their results.[55]

Precisely for these reasons, much company research would meet the Common Rule's criteria for a waiver or alteration of informed consent. That is arguably as it should be (there are sound normative reasons why the criteria for waiver and alteration of informed consent are as they are) – and so this is not an argument for the inapplicability of the Common Rule to company research, per se. But it will not appease those who believe that participating in an experiment or other study should always involve the kind of consent one expects in, say, a clinical trial of an experimental drug.

[48] *See* Holly Fernandez Lynch & Michelle N. Meyer, *Regulating Research with Biospecimens Under the Revised Common Rule*, 47 Hastings Ctr. Report 3 (May–June 2017).

[49] 45 C.F.R. § 46.101(b)(4). [50] What constitutes "public behavior" is not well defined.

[51] *Id.* § 46.101(b)(2). *Cf.* Consumer Privacy Bill of Rights Act of 2015 at §4(g) (defining "privacy risk" as the potential of data disclosure to "cause emotional distress, or physical, financial, professional or other harm to an individual").

[52] Calo at 99–100 (citing Michael Moss, *The Extraordinary Science of Addictive Junk Food*, N.Y. Times Mag. (Feb. 20, 2013)).

[53] *Id.* § 46.101(b)(6). [54] *Id.* § 46.111(a)(2).

[55] *See, e.g.*, Meyer, *supra* note 36, at p. 298 (explaining why study-specific consent would have biased the results of Facebook's mood contagion and OkCupid's matching algorithm experiments).

An increasingly common alternative to study-specific consent is one-time "broad consent," in which participants agree at the time of collection to have their data (or specimens) stored for future, unspecified research. Research institutions have long used broad consent, and the 2018 Common Rule explicitly blesses this alternative in certain circumstances.[56] Broad consent is not a blank check; it requires, among other things, a description of the *kinds* of studies that might be conducted using stored data or tissue.[57] For instance, participants who give broad consent to contribute to a biobank may be told that their tissue and resulting data will be used in health research. Facebook or Google users who will be subject to a variety of A/B testing might be told, in general terms, that they may see one version of a page while others see a different version, and that the length of time they remain on that page, or the extent to which they click through to another page, may be (pseudonymously) recorded and compared with the data of users who see an alternative page in order to help the company decide which page is preferable (either for users or for the business).

Broad consent has been criticized by some: consent that is – by definition – not informed about the specific nature of research, they argue, is insufficiently protective of autonomy.[58] But we knowingly make considered decisions based on incomplete information all the time. As discussed above, however, the prima facie obligation to obtain research subjects' informed consent is primarily rooted in the principle of respect for autonomy. And broad consent need be no less autonomous a decision than study-specific consent: they are simply different kinds of autonomous decisions, one involving an understanding and acceptance of what will happen in a particular study and one involving an understanding and acceptance that the participant's knowledge will be limited to the kinds of studies that may occur.[59]

The closest existing business research analog to broad consent is perhaps the ubiquitous and much-maligned Terms of Service (ToS), or the similarly dense Data Use Policy and Privacy Policy. Whatever the demerits of incompletely informed broad consent, the main difference between it and these company agreements is that virtually everyone "accepts" the ToS without even attempting to read its contents.

There are, however, ways of ensuring that individuals read and even meaningfully comprehend documents such as ToS and consent forms. Research studies increasingly give prospective participants short (online, automatically "graded") comprehension quizzes after they've read the consent (whether broad or study-specific). A company that wants to inform its customers that it learns from customer data in various ways (perhaps including through A/B tests and other experiments) can require one-time passage of such a broad research ToS. Asking customers to agree to a separate Research ToS, labeled as such, is another way of highlighting that particular term of service and prompting those who are concerned to peruse its contents. Assuming that the disclosures and questions are sufficiently clear and accurately reflect the company's learning practices and their implications for customers, a customer who successfully completes the quiz and signs up for the product or service can be said to have given ethically meaningful informed consent.

Users might be required to accept the Research ToS as a condition of opening an account, or they may be given a choice about whether to opt into (or out of) the Research ToS.[60] More

[56] *See* Fernandez Lynch & Meyer, *supra* note 47.

[57] *See, e.g.*, 45 C.F.R. §46.116(d) (2017). For instance, a biobank might collect tissue under a consent that limits secondary research use of the tissue to health-related research aims.

[58] *See, e.g.*, Björn Hofmann, *Broadening Consent – and Diluting Ethics?*, 35 J. OF MED. ETHICS 125 (2009).

[59] Mark Sheehan, *Can Broad Consent Be Informed Consent?*, 4 PUB. HEALTH ETHICS 226 (2011).

[60] For instance, the direct-to-consumer genetic testing company 23andMe requires all of its consumers to accept both its Privacy Statement and its Terms of Service. These documents disclose (among many other things) that the company and its partners use consumer data not only to provide genetic testing services and to bill consumers for the same, but

customers may be willing to give broad consent to participate in a business's various learning activities than might be expected, especially if some of those learning activities give customers opportunities to be alpha or beta users or the latest innovation. Studying only those customers who opt in to (or fail to opt out of) participation in research, and not the remaining customers, could bias the results, however, because there may be important differences between these two groups of people. Even so, a company might conduct pilot studies on such "empaneled" customers and, if the results are promising, roll out an intervention to successively broader groups of customers.

Whether informed consent-by-ToS is also sufficiently *voluntary* is a separate issue. As with virtually every academic or industry research study, there is no negotiating over the ToS: both research consent forms and commercial ToS are almost always take-it-or-leave-it offers. If a company has a monopoly on a critical product or service, then even if customers understand what they're getting into, thanks to clear ToS and a forced comprehension quiz, they may still lack any meaningful choice over whether to accept those terms or not.[61] The result will be similar if all major companies in a sector require their customers to agree to participate in research and competition has not led to the emergence of an alternative that offers more preferable ToS.

On the other hand, the Common Rule's requirement that subjects are "volunteers"[62] for whom research participation should ordinarily be voluntary – and, indeed, that subjects must ordinarily be told that they may revoke their consent to participate at any time and for any reason, regardless of how that might affect the study[63] – may be less appropriate for an arm's length transaction between company and customer. In many cases, it is ethically and legally permissible for a company to offer a product or service premised on certain nonnegotiable conditions that, all else being equal, consumers would not choose. Both law and ethics require that those conditions not be unconscionable, but otherwise a consumer's right to choose may be limited to her initial decision whether to avail herself of a company's product or service or not.

A very great deal of company research easily meets that test. It is hardly unconscionable, for instance, if a company requires its customers or users to agree to have their data used for QI/QA purposes. That is quite routine[64] and the fact that a QI/QA activity may take the form of a

also to "perform research & development activities, which may include, for example, conducting data analysis and research in order to develop new or improve existing products and services, and performing quality control activities." 23andMe, *Privacy Statement*, https://www.23andme.com/legal/privacy/, §4(a)(viii). The company may also share aggregate or (undefined) "anonymous" consumer data with third parties for various purposes. *Id.* at §4(d)(iii). By contrast, before individual-level consumer data may be used by 23andMe in research designed to contribute to generalizable knowledge, consumers must opt in via a separate "research consent." See 23andMe, *Research Consent*, https://www.23andme.com/about/consent/ (undated; last accessed May 6, 2017); 23andMe, *Privacy Statement*, *supra* at §4(b).

[61] For instance, several people from across the political spectrum have argued that Facebook use is so ubiquitous and central to daily modern life that it should be regulated like a utility. *See*, *e.g.*, danah boyd, *Facebook Is a Utility: Utilities Get Regulated*, Apophenia, May 15, 2010, http://www.zephoria.org/thoughts/archives/2010/05/15/facebook-is-a-utility-utilities-get-regulated.html; Scott Adams, *Should Twitter and Facebook Be Regulated as Utilities?*, Scott Adams' Blog, Jan. 25, 2017, http://blog.dilbert.com/post/156377416856/should-twitter-and-facebook-be-regulated-as. *See also* Philip N. Howard, *Let's Nationalize Facebook*, Slate, Aug. 16, 2012, http://www.slate.com/articles/technology/future_tense/2012/08/facebook_should_be_nationalized_to_protect_user_rights_.html; Timothy B. Lee, *Facebook Has More Influence Over Americans Than Any Media Company in History*, Vox, May 10, 2016, https://www.vox.com/2016/5/10/11640140/facebook-media-influence.

[62] Belmont Report, *supra* note 20, at Part C.1. [63] 45 C.F.R. § 46.116(a)(8).

[64] In the furor over the Facebook mood contagion experiment, much was made of the fact that Facebook had only after the fact included a disclosure in its ToS that it conducted "research." Aside from the important point that the content of a standard ToS is close to morally irrelevant, given the accepted fact that virtually no one reads them, Facebook's Data Use Policy in effect at the time did disclose that the company might use users' data "as part of our efforts to keep

randomized, controlled evaluation rather than an observational approach does not, per se, change the moral calculus.[65] Indeed, in some cases – as discussed in the final section of this chapter – it would be unconscionable (or at least ethically inferior) for a company *not* to use nonidentifiable data to ensure that its product or service works the way the company advertises or is safe for users. Nor is it unconscionable if a company decides to exercise its social responsibility (or recruit top data scientists) by conducting research that contributes to the public good of generalizable knowledge.

Many companies conduct pilot studies with their employees. Perhaps ironically, voluntariness is more likely to pose a significant issue in these cases than in cases of customers, who can (barring monopolies) always opt to use a competitor's product or service. Employees, of course, generally have a great deal invested in their employment (financially and professionally, but often also socially and psychologically). This investment can make declining an employer's request to serve as a pilot subject difficult or practically impossible. Employers should therefore generally make such service completely voluntary, especially where studies involve potentially sensitive data or other significant risks.

Similarly, data privacy will generally be more important in this context. Most of us are more concerned about our employers or colleagues learning sensitive things about us than we are about a stranger we are unlikely ever to meet (say, a company data scientist) learning the same thing about us. When employees do opt in to participate in pilot studies, their data should generally be at least coded (if not anonymous), with access to the key linking employee data to their employee identities limited to as few employees as possible (ideally, only one, and strict firewalls should separate these data from company departments such as human resources and benefits), and the key deleted as soon as possible after data collection and analysis.[66]

It is important to note that a different situation is posed by many employer learning activities aimed at assuring or improving workplace policies and conditions. The purpose of such activities – unlike pilot testing a company's new products on its employees – is less orthogonal to the employer–employee relationship. Imagine, for instance, a CEO who worries that some of her employees aren't taking optimal advantage of the company's generous 401(k) matching program.[67] She has a hunch that adding to the usual matching letter information telling each

Facebook products, services and integrations safe and secure," Kashmir Hill, *Facebook Added "Research" To User Agreement 4 Months After Emotion Manipulation Study*, Forbes (June 30, 2014), http://www.forbes.com/sites/kashmirhill/2014/06/30/facebook-only-got-permission-to-do-research-on-users-after-emotion-manipulation-study/ (referring to users as "guinea pigs made to have a crappy day for science"), a QA/QI purpose that, I have argued, the mood contagion experiment served (even if it also contributed to generalizable knowledge). *See* Meyer, *supra* note 36.

[65] All else being equal, randomized, controlled QA/QI activities are ethically preferable to QA/QI activities using observational methods, because the former permit causal inferences to be made about the effects of a company's goods, services, policies, or practices. The primary ethical issue posed by randomized evaluations is whether different arms have different risk-benefit profiles. The traditional answer of research ethics to this issue is that randomizing subjects to different arms is ethical if the relevant expert community as a whole is in "equipoise" as to which arm, if any, is superior to the other(s). For the seminal statement of this position, see Benjamin Freedman, *Equipoise and the Ethics of Clinical Research*, 317 New Eng. J. of Med. 141 (1987). Clinical equipoise is, however, the subject of lively and continuing debate in the field of research ethics. *See, e.g.*, Frank G. Miller & Howard Brody, *A Critique of Clinical Equipoise: Therapeutic Misconception in the Ethics of Clinical Trials*, 33 Hastings Ctr. Report 19 (2003).

[66] Selecting employees as pilot subjects may well also bias the data, if employees tend to be more homogeneous than end users.

[67] This scenario, loosely based on actual studies and the challenges that economists tell us they have faced in conducting them, is adapted from Michelle N. Meyer & Christopher F. Chabris, *Please, Corporations, Experiment on Us*, N.Y. Times (June 19, 2015), https://www.nytimes.com/2015/06/21/opinion/sunday/please-corporations-experiment-on-us.html.

employee what his or her similarly aged peers are saving would nudge the low-savers to save more. The CEO could implement this peer information letter policy in the next enrollment cycle and observe the result, but regardless of whether savings went up or down, she'd never really know what effect, if any, the policy had without running a randomized, controlled experiment in which some employees receive letters with no peer information and others receive letters with peer information. Telling employees in advance that the company would be sending out different letters and why would badly bias the results by altering employees' behavior; no social scientist worth her salt would take the results of such a study seriously.

Although some insist that "human experimentation" always requires informed consent, the CEO would be "experimenting" on her employees no matter what; the only question is whether her experiment will be properly controlled, enabling her to learn from the experience and make sound choices going forward. The CEO will either send the usual matching letter or the alternative letter to all employees, using her intuition to choose between them, without ever knowing which one is better for her employees – and surely without anyone insisting that employees provide informed consent to receive one version of the letter instead of the other. Or she will have the epistemic humility to realize that she does not know which letter will better serve her employees, and will conduct an A/B test to answer that question. To have more qualms about conducting that A/B test than about enacting an untested practice across the board is what I have elsewhere called the "A/B illusion."[68]

Notably, the above challenges in applying the Common Rule's substantive criteria to company research are in addition to the many critiques of it as applied to the academic and biomedical industry context. For instance, many critical terms in the Common Rule are ambiguous or difficult to apply, allowing the processing of IRB review, in which those terms are applied to a study protocol, to be captured by interests other than participant welfare or to reflect various forms of bias. Although the Common Rule requires IRBs to assign each study a single risk-benefit profile, heterogeneity among subjects means that a study will in fact have very different risk-benefit profiles for many subjects.[69] The Common Rule is based on a medical model of research and has crowded out other disciplines' traditions of ethical reflection of research.[70]

Different business sectors, if not individual businesses, will benefit from thinking about the right set of ethical principles to guide the particular learning endeavors in which they are engaged. Still, the Common Rule does have some general wisdom to offer: risks and expected benefits are ethically relevant inputs into decisions about the appropriateness of research. Risks to subjects are important to consider, and although the Common Rule directs IRBs not to consider them, businesses that wish to be socially responsible may well wish to consider the risks and costs to third parties of their learning activities. Risks should be minimized to those necessary to achieve the learning aim. The distribution of risks and expected benefits matters: imposing risks on customers is more easily justified if those customers are also among those who stand to benefit from the results of an activity. In the absence of study-specific consent, meaningful notice at the point of service of a company's general research practices and policies can help people make better informed decisions about whether or not to use a product or service. Amending one's products, services, or practices in light of evidence about their effects,

[68] *See* Meyer, *supra* note 36. See also the discussion, below, of HiPPOs.

[69] *See* Michelle N. Meyer, *Regulating the Production of Knowledge: Research Risk-Benefit Analysis and the Heterogeneity Problem*, 65 Admin. L. Rev. 237 (2013).

[70] *See* Zachary M. Schrag, Ethical Imperialism: Institutional Review Boards and the Social Sciences, 1965–2009 (2010).

or disseminating the results of internal research so that they can guide decision-making by end users, is a good idea.

Indeed, the "basic formula" of Facebook's internal research review process (discussed further below) "is the same as an IRBs [sic]: We consider the benefits of the research against the potential downsides."[71] In particular, the company considers "how the research will improve our society, our community, and Faceboook"; "whether there are potentially adverse consequences that could result from the study," especially due to vulnerable populations or sensitive topics, and "whether every effort has been taken to minimize them"; "whether the research is consistent with people's expectations"; and whether it has "taken appropriate precautions designed to protect people's information."[72]

OVERSIGHT OF COMPANY RESEARCH

Although it is unclear how useful or appropriate it would be to apply the details of the Common Rule's *substantive* criteria to company research, businesses should experiment with adopting something *like* the *process* of IRB review.[73] This somewhat cautious recommendation for a mere approximation of the IRB review process reflects the many legitimate criticisms of that process. For instance, most IRB members are inexpert in both ethical reasoning and the scientific methods of many of the studies they review, and they frequently fail to base their decisions on evidence of risks and expected benefits.[74] Although both researchers and prospective participants are biased in how they approach decisions to participate in research, IRB members are also burdened by their own biases.[75] Finally, compared to other regulators, IRBs are lawless: they are largely unaccountable, have nearly boundless jurisdiction, and typically make ad hoc decisions generally subject to no review or other mechanisms for due process.[76]

Nevertheless, boiled down to its essence, IRB review is simply a mechanism for gathering feedback about a planned learning activity from someone other than the person likely to be most invested in it, who may therefore have a biased view of its risks and potential benefits.[77] In some cases, both the informed consent process and the fact that most prospective participants are free

[71] Molly Jackman & Lauri Kanerva, *Evolving the IRB: Building Robust Review for Industry Research*, 72 Wash. & Lee L. Rev. Online 442, 454 (2016). See also Michelle De Mooy & Shelten Yuen, *Toward Privacy Aware Research and Development in Wearable Health: A Report from the Center for Democracy & Technology and Fitbit, Inc.* (May 2016), available at https://healthblawg.com/images/2016/06/CDT-Fitbit-report.pdf (suggesting that both the Belmont Report and the Common Rule are ethically relevant to wearables companies' research and development processes – but somewhat mischaracterizing both those sources of ethical guidance).

[72] *Id.* at 454–56.

[73] *See, e.g.*, Ryan Calo, *Consumer Subject Review Board: A Thought Experiment*, 66 Stan. L. Rev. Online 97 (Sept. 3, 2013), https://www.stanfordlawreview.org/online/privacy-and-big-data-consumer-subject-review-boards/ (proposing that corporations rely on "consumer subject review boards"); De Mooy & Yuen, *supra*, at 20–21 (developing ethical guidance for wearables companies' R&D and proposing that "a company might create a 'Privacy Board' comprised of selected company staff to review the research practices and policies for in-house volunteers and device users, and offer ongoing feedback. Companies may find it useful to bring together privacy experts, advocates, academics, and their R&D teams regularly to discuss existing and emerging privacy issues, market concerns, and technical challenges facing wearable companies.").

[74] *See* Robert Klitzman, The Ethics Police? The Struggle to Make Human Research Safe (2015).

[75] *See* Michelle N. Meyer, *Three Challenges for Risk-Based (Research) Regulation: Heterogeneity among Regulated Activities, Regulator Bias, and Stakeholder Heterogeneity*, *in* Human Subjects Research Regulation: Perspectives on the Future 313 (I. Glenn Cohen & Holly Fernandez Lynch eds., 2014).

[76] *See* Carl E. Schneider, The Censor's Hand: The Misregulation of Human-Subject Research (2015).

[77] When companies use external IRBs or internal IRB-like boards, the decisions of those boards are almost always advisory only. Ordinarily, as with most business decisions, management will have final say over whether a project goes forward. Often, representatives from various relevant management sectors (e.g., the Chief Privacy Officer, Chief

either to leave or take a research offer might plausibly operate as checks on the researcher's self-interest and biased perspective. But because so much of company research neither is nor practically can be conducted with study-specific consent, and because consumers frequently have little choice in choosing to patronize a business that doesn't engage in learning activities, some other check on the researcher's perspective is in order.

Indeed, when the Obama administration was confronted with the dilemma that there will often be important but non-contextual uses of consumer data for which consent is infeasible, it proposed, in its draft Consumer Privacy Protection Act of 2015, that "privacy review boards" be empowered to authorize such non-contextual uses of consumer data if a study met criteria very much like the Common Rule's criteria for waiving consent.[78]

Although the Consumer Privacy Protection Act of 2015 and its privacy review boards did not become law, similar ideas already exist in the business world. Since about 2007, for example, Google has employed an "experiment council" – a group of internal engineers who try to ensure proper experimental design by reviewing proposed experiments using a "lightweight," continually updated checklist of best practices that the experimenter completes. Following the completion of an experiment, results are then presented at a separate "discussion forum" open to anyone at the company but facilitated by experts. The discussion forum aims to ensure that results are valid and properly interpreted, and that there is agreement about the consequences of the interpreted results for the user experience.[79] Although it may appear otherwise, ensuring that experiments are properly designed and that their results are appropriately understood and disseminated are matters not only of scientific and business but also of ethical import. A poorly designed study cannot answer the question it seeks to answer (no matter how important that question may be), rendering any risks or burdens imposed on subjects gratuitous.[80]

Since about 2009, direct-to-consumer genetics company 23andMe, which also conducts research with genetic data from customers (who provide broad consent to such research), has voluntarily employed an external IRB to review most of its research.[81] That may not be surprising, since the research that 23andMe conducts is similar to research conducted by academic geneticists and is familiar to IRBs. Both Microsoft and Fitbit, Inc. also submit some portion of their research projects to external IRB review and review the remaining projects internally.[82]

Very little is known about such internal review processes, including how common they are among businesses. Following the public fallout over the June 2014 publication of its mood contagion experiment, Facebook announced, in October of the same year, that it was

Communications Officer, or in-house counsel) will sit on an internal review board, and if they object to a project for legal or business reasons, that will be the end of the conversation.

[78] White House, *Administration Discussion Draft: Consumer Privacy Bill of Rights Act of 2015* (Feb. 27, 2015), https://www.democraticmedia.org/sites/default/files/field/public/2015/draft_consumer_privacy_bill_of_rights_act.pdf.

[79] Diane Tang, Ashish Agarwal, Deirdre O'Brien & Mike Meyer, *Overlapping Experiment Infrastructure: More, Better, Faster Experimentation*, *in* Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 17–26 (2010), https://research.google.com/pubs/archive/36500.pdf.

[80] *See, e.g.*, 45 C.F.R. § 46.111 (criteria for IRB approval of research include a determination that "risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result").

[81] *Protecting People in People Powered Research*, 23andMeBlog, July 30, 2014, https://blog.23andme.com/23andme-research/protecting-people-in-people-powered-research. *See also 23andMe Improves Research Consent Process*, 23andMeBlog, June 24, 2010, https://blog.23andme.com/23andme-research/23andme-improves-research-consent-process.

[82] Daniela Hernandez & Deepa Seetharaman, *Facebook Offers Details on How It Handles Research*, Wall St. J., June 14, 2016, https://www.wsj.com/articles/facebook-offers-details-how-it-handles-research-1465930152.

significantly revising (and strengthening) its internal research review process.[83] Employees spearheading that iterative process later published a description of it, which involves training across the company, various research review "checkpoints," and a standing, five-person "research review group" comprised of employees with expertise in policy, ethics, law, and communications (in addition to substantive expertise in Facebook's areas of research) that reviews projects triggered at earlier checkpoints as raising especially complex ethical or other issues.[84] In addition to this internal process, Facebook periodically consults outside subject matter experts.[85]

One important consideration in developing an internal review board is whether and how perspectives external to the company will be brought to bear on decision-making. Two kinds of outside perspective may be especially important: that of professional (usually academic) ethicists and that of laypersons who can represent the end user's experience and expectations. Outsiders may not only be more likely to recognize ethical issues; they may also feel freer than employees to voice concerns about them.[86]

One way to incorporate one or both kinds of perspectives is to include them as external members of a company's otherwise internal review board. The Common Rule, for instance, requires IRBs to include at least one member who is a non-scientist – someone who isn't steeped in the perspective of scientific research and may not see proposed studies the same way — and at least one member who is a non-affiliate of the institution – so that they do not have any inherent financial or other professional investment in the institution being permitted to conduct the research.[87]

Lay perspectives are critical, among other reasons, as a way of correcting insiders' "curse of knowledge." Each of us in an insider in some domain, and in that domain, it is often difficult to know what knowledge we can expect outsiders to share and what knowledge is inside baseball. For instance, lay criticism of the Facebook mood contagion experiment revealed that many users did not realize that their News Feeds have always been curated by an algorithm that prioritizes some posts and deprioritizes others; because they were unaware of this fact, the experimental manipulation was a more dramatic departure from what they believed to be the status quo than it was from Facebook's actual practice, likely increasing the level of outrage. The News Feed algorithm is fundamental to the Facebook platform, and it may never have occurred to some within the company that many users viewed their Facebook News Feed as a simple reverse chronological list of all their friends' posts, unmanipulated by Facebook in any way. Sharing project plans with actual or potential end users can reveal these misunderstandings and suggest ways that research projects should be conducted, telegraphed, or disclosed that will help close the gap between company behavior and end user expectations.

There are advantages and disadvantages to including lay members on internal review boards. Some empirical studies of academic IRBs have found that such members do not always feel empowered to speak up, at least unless there is something like a critical mass of them on the

[83] Reed Albergotti, *Facebook Tightens Oversight of Research*, Wall St. J., Oct. 2, 2014, https://www.wsj.com/articles/facebook-tightens-oversight-of-research-1412292491.

[84] Molly Jackman & Lauri Kanerva, *Evolving the IRB: Building Robust Review for Industry Research*, 72 Wash. & Lee L. Rev. Online 442 (2016).

[85] For example, before conducting "research on trends in the LGBT community on Facebook, [the company] sought feedback from prominent groups representing LGBT people on the value that this research would provide and on what data to collect and report." *Id.* at 453.

[86] The extent to which this relative freedom exists depends, however, on whether and how external members of review boards are compensated. The greater the financial or professional compensation she receives for serving, the more she has to lose by speaking out, and the less she is distinguishable in this regard from an employee.

[87] 45 C.F.R. § 46.107(c)–(d).

board. Laypersons' lack of expertise in both ethics and the substantive matter under consideration (e.g., data science or wearable technology) can also make it difficult for them to meaningfully participate on the same level as other members.[88] An alternative way of learning about user expectations is through surveys, focus groups, or interviews, which can be tailored to specific planned learning activities.

Ethics experts, too, are critical to ethics review. They have thought long and deeply about certain ethical issues and are therefore well positioned to spot those issues as they arise and bring practical tools to bear in thinking about how to address them. Some companies employ professional ethicists, and more retain professional ethicists as standing consultants. But if a company creates an internal review board, it will be more helpful if ethics experts are included in those discussions as frequently as possible rather than occasionally being consulted.

What roles can internal review boards play? Company Research Review Boards (RRBs) can ensure that research risks are reasonable in light of the expected benefits and in light of the distribution of risks and expected benefits. They can also consider whether the proposed activity would likely match customers' expectations and, if not, whether the project should be abandoned or reconceived to meet expectations, or whether user or consumer expectations should be reset before proceeding. Company RRBs can develop, disseminate, and enforce policies regarding which kinds of data must be anonymous, coded, or de-identified. They can review manuscripts before they are submitted for publication (whether in an academic or trade journal or on a company research blog) to ensure that subject privacy is maintained and that the study results are clearly and accessibly explained and do not needlessly alarm customers. They may also be able to help advise management about the business implications of results.

It must be kept in mind that prospective group review isn't a panacea. IRBs and similar boards don't always prevent unethical activities from occurring, both because their review isn't perfect and because, following approval, they typically exercise minimal oversight over the actual activity. IRBs can also delay, alter, or block valuable and ethical research in ways that can be harmful to people's welfare. In short, IRBs make Type I and Type II errors (or false positives and false negatives), and both are worth watching out for and trying to mitigate.

HUNGRY, HUNGRY HIPPOS: THE ETHICS OF *NOT* LEARNING

"If we have data, let's look at data. If all we have are opinions, let's go with mine."

– Jim Barksdale, former CEO of Netscape[89]

Most discussions of research ethics limit themselves to the ethical issues raised by *decisions to conduct research* or some other learning activity. This is a mistake (and perhaps the result of a common human bias toward noticing acts more than omissions). An appropriately comprehensive ethical analysis includes not only the risks and expected benefits of studying something, but also the risks and expected benefits of *not studying* that thing. Previous sections of this chapter have alluded to companies' responsibility for learning about the effects on others (and especially on their users or customers) of their products, services, policies and practices and to the critical

[88] Institutions sometimes appoint to various boards "lay" members who are actually quite sophisticated about the matter under discussion, whether in order to avoid this problem or because such lay members are more interested in participating. Unfortunately, such lay members begin to look and think more like the "insiders" than like the modal end user, which largely defeats the point of their presence.

[89] This quote has been widely attributed to Barksdale.

role that data collection and analysis, experiments, and other learning activities play in meeting that responsibility.

Most business decisions are made on the basis of what Avinash Kaushik and Ron Kohavi call HiPPOs: the Highest Paid Person's Opinions.[90] Some might assume that those who rise to the level of "highest paid person" must have done so by virtue of possessing unique knowledge about what works in business, but the reality is obviously more complex.[91] CEOs and other leaders do often have extensive experience observing the results of various business decisions. But it is rare for such observations to be systemically captured and analyzed, and their lessons extracted and implemented. Moreover, even the best observational study cannot determine whether what is observed is caused by a particular business decision. Instead of evidence, business leaders typically make decisions on the basis of tradition ("how it's always been done," or how their own mentors did things) or instinct. This is problematic because gut instincts – while often romanticized – all too often prove wrong, sometimes leading to harm incurred or to welfare gains foregone.

As an early Amazon engineer, for instance, Greg Linden had the idea of recommending additional purchases to customers based on what they already had in their virtual baskets. A marketing senior vice president was "dead set against it." While Linden's hypothesis was that the feature would tend to result in customers buying more products, the HiPPO's hypothesis was that it would distract customers from checking out, and they would tend not to buy anything at all. The HiPPO forbade Linden from pursuing the project. But Linden decided to test their competing hypotheses on Amazon.com, and his prevailed: "Not only did it win, but the feature won by such a wide margin that not having it live was costing Amazon a noticeable chunk of change." The shopping cart recommendations feature was then rolled out across the platform.[92]

Similarly, at JoAnn.com, marketers trying to sell more sewing machines online guessed that the *least* effective promotion of those they were considering would be one that advertised 10 percent off the sale of two machines. The CFO's instinct was that few people wanted two sewing machines, and therefore opting for this promotion would amount to "wast[ing] a week's worth of sales on this promotion." But when the company tested his intuition, they discovered that this promotion increased conversion rates (the rate at which online users who click on a promotion actually purchase the product the promotion advertises) by 70 percent – far more than any of the other promotions they tried. (It turns out that many sewers are members of sewing groups, who simply pursued the promotion in pairs.)[93]

In the Amazon and JoAnn examples, what were primarily at stake were corporate profits. Other business decisions more directly impact customer or user welfare. Consider, again, Facebook's mood contagion experiment. Observational studies by academics of Facebook use had reached contrasting conclusions about what psychological risks, if any, Facebook's News Feed poses to users. Some studies found that friends' positive posts make users feel worse about their own lives (a

[90] Ron Kohavi, Randal M. Henne & Dan Sommerfield, *Practical Guide to Controlled Experiments on the Web: Listen to Your Customers Not to the HiPPO*, in Proc. of the 13th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining 959, 966 (2007).

[91] *See, e.g.*, then-CEO of OkCupid Christian Rudder, *We Experiment on Human Beings!*, The OkCupid Blog, July 27, 2014, https://theblog.okcupid.com/we-experiment-on-human-beings-5dd9fe280cd5 ("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.").

[92] Greg Linden, *Early Amazon: Shopping Cart Recommendations*, Geeking With Greg blog, April 25, 2006, http://glinden.blogspot.com/2006/04/early-amazon-shopping-cart.html.

[93] Russ Banham, *Power to the Little People*, CFO Magazine, Dec. 14, 2005, http://ww2.cfo.com/technology/2005/12/power-to-the-little-people/.

social comparison effect). Others found that it is friends' negative posts that pose the risk, by depressing users (through a mood contagion effect). Others concluded that both of these hypotheses hold some truth, and that the effect of the News Feed on users depends on an individual user's characteristics (such as personality). Still others believed that the News Feed has no significant psychological effect on users and that studies finding otherwise were mere noise, based on low sample sizes, self-reported outcomes, and other less-than-gold-standard methods.

The News Feed has always been curated by a proprietary algorithm that results in some posts being more likely to be seen than others. Moreover, all users are exposed to undulating levels of both positive and negative content in their feeds, depending on local and national events and other factors. In order to rigorously evaluate the effects on users' mood of exposure to both positive and negative words, Facebook conducted a randomized, controlled experiment in which some user feeds had relatively more positive posts suppressed, others had relatively more negative posts suppressed, and still other feeds served as controls. Experimental exposures to positive and negative words were almost certainly within the range of exposure that users experience over the course of, say, several months of Facebook use. The effects were extremely small, detectable only because the sample size (about 700,000 users across all conditions) was so large. The investigators found that users who were exposed to more positive words themselves used slightly more positive words when writing their own posts, whereas those who were exposed to more negative words used slightly more negative words in their own posts – findings that they interpreted as evidence for the mood contagion hypothesis.

After the results of the experiment were published in an academic journal and reported by the media, widespread public condemnation of Facebook ensued, including contemplated lawsuits and calls for federal and state agency investigations. That condemnation, however, seems mostly misplaced. Facebook's algorithm and other practices almost certainly affect users' exposure to their friends' positive and negative words. The existing evidence about the effects of negative and positive words pointed in different directions: if the social comparison hypothesis was correct, then Facebook's platform poses risks through positive words. If the mood contagion hypothesis was correct, then it is negative words that pose the threat to Facebook's users; positive words actually should be beneficial. If one takes seriously these risks, then the decision that is more difficult to ethically defend would have been a decision *not* to try to determine the truth about these risks through a rigorous, yet low-risk, experiment.

A similarly laudable example of a company using research methods to improve the safety of its platform comes from the online gaming company, Riot Games. The company uses social psychology experiments to combat a range of troublesome behaviors, from incivility to harassment, that plague its popular online game, League of Legends.[94] We all wish harassers and those who are uncivil online would just stop or go away. But since that is unlikely to happen, we are better off if companies quickly gather rigorous evidence of what works through low risk, nonconsensual experiments than if they simply guess about how to address the problem.

OkCupid, the online dating company, conducted an experiment that nicely illustrates how research methods can be used to investigate the effectiveness of products and services. OkCupid markets itself as offering customers a science-based approach to dating. And indeed, those pairs of customers who the company's touted "matching algorithm" rates as compatible do tend to engage in longer conversations on the platform. But the company wasn't actually sure whether these pairs engage in more extensive conversation because the matching algorithm accurately

[94] Jeremy Hsu, *Inside the Largest Virtual Psychology Lab in the World*, Backchannel, Jan. 27, 2015, https://backchannel.com/inside-the-largest-virtual-psychology-lab-in-the-world-7c0d2c43cda5#.3d48yzqzi.

predicts compatibility or whether simply telling pairs of people that they are compatible makes them believe and behave as if they were. So it conducted an experiment in which it told some pairs of customers that its algorithm predicted would be compatible that they were not so compatible, it told other pairs that the algorithm predicted would not be compatible that they were, and it compared both of these groups to pairs who were told exactly what the algorithm predicted. As measured by how many messages these pairs of customers exchanged, OkCupid found that although its algorithm did, in fact, predict which pairs would exchange more messages, the act of merely telling pairs that they were compatible (regardless of how compatible the algorithm said they were) was a slightly better predictor.[95] The company reported the results of the experiment on its blog, where customers could use the results to gauge how much reliance to place on the algorithm's suggested pairings.[96]

The Facebook, Riot Games, and OkCupid examples show that nonconsensual "human experimentation" (whether characterized as research or as QI) can be ethically acceptable – indeed, laudable or even ethically obligatory – when the learning activity imposes *no more than minimal risks* on subjects (above and beyond the risks they are already exposed to as customers); when it seeks to *determine the effects of an existing or proposed company practice or product* on those who will be affected by it, so that subjects exposed to (minimal) risks are also those who stand to benefit from the results; and when *the validity of the data* or other practical constraints precludes obtaining fully informed consent.

In the wake of its mood contagion experiment, Facebook was accused of abusing its power, depriving its users of important information, and treating its users as mere means to corporate ends. Yet it is declining to investigate probative, but conflicting and indeterminate, evidence that its platform harms users that would have exhibited these failures, not conducting a QI experiment. Similarly, critics of OkCupid's matching algorithm experiment found it outrageous that the company briefly told some pairs of customers that they were compatible, when the algorithm said that they were not. But the point of the experiment was that, despite its marketing campaign, the company lacked evidence that its algorithm in fact accurately predicted compatibility. Subjects were not deprived of information known to be accurate or given information known to be false. And in fact, the experiment found that the algorithm is slightly *less* predictive of which pairs will get along well than is simply telling people that they *should* get along well.

It was better, ethically speaking, for Facebook and OkCupid to have conducted these experiments than for them not to have done so, given the probative but inconclusive evidence that emotionally laden Facebook content may post psychological risks to users and that the success of OkCupid's touted matching algorithm may have been partly or wholly due to the power of persuasion rather than to an algorithm that scientifically detects romantic compatibility.[97]

CONCLUSION

There has always been tension between the idea that people need to be protected from research and the idea that people need access to the sometimes considerable benefits of participating in research.[98] Companies increasingly engage in a wide variety of learning activities. These activities may be ethically dubious, praiseworthy, or even obligatory, depending on how much

[95] For a more detailed discussion of this experiment and its ethical implications, see Meyer, *supra* note 36, at 312–21.
[96] *Rudder*, *supra* note 91.
[97] *See generally* Meyer, *supra* note 36 (discussing both experiments at length and defending this viewpoint).
[98] Anna Mastroianni & Jeffrey Kahn, *Swinging on the Pendulum: Shifting Views of Justice in Human Subjects Research*, 31 Hastings Center Rep. 21 (2001).

unique privacy and other risks they impose on subjects, whether they are designed to yield important information, whether those who bear the burdens of the learning activity also stand to reap its potential rewards, and whether the learning activity is consistent with subjects' reasonable expectations of how the company will behave. Company learning activities almost always take place outside the jurisdiction of the regulations governing human subjects research. Moreover, even if company learning activities were subject to those regulations, the vast majority would be partly or fully exempt. Nevetheless, companies would benefit from judiciously borrowing elements from both the substance and the process of human subjects regulations, including attention to notice and broad consent, risk-benefit balance and risk minimization, and prospective review of proposed learning activities by a diverse group of individuals who can help anticipate actual and perceived ethical problems.