6.6.2 Understanding and managing informational risk

increased dramatically; and it is the hardest risk to understand. The second ethical challenge for digital-age research is informational risk, the potential

Informational risk is the most common risk in social research; it has

for harm from the disclosure of information (National Research Council 2014). Informational harms from the disclosure of personal information could be economic (e.g., losing a job), social (e.g., embarrassment), psychological (e.g., depression), or even criminal (e.g., arrest for illegal behavior). Unfortunately, the digital age increases informational risk dramatically—there is just so much more information about our behavior. And informational risk has proven very difficult to understand and manage compared with risks that were concerns in analog-age social research, such as physical risk. One way that social researchers decrease informational risk is "anonymization" of data. "Anonymization" is the process of removing obvious personal identifiers such as name,

address, and telephone number from the data. However, this approach is much less

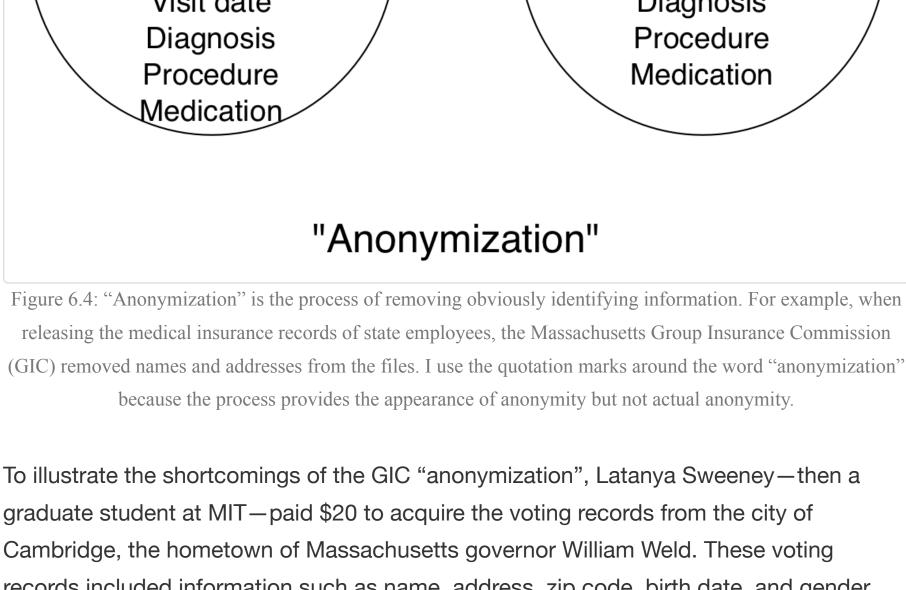
effective than many people realize, and it is, in fact, deeply and fundamentally limited. For

that reason, whenever I describe "anonymization," I'll use quotation marks to remind you that this process creates the appearance of anonymity but not true anonymity. A vivid example of the failure of "anonymization" comes from the late 1990s in Massachusetts (Sweeney 2002). The Group Insurance Commission (GIC) was a government agency responsible for purchasing health insurance for all state employees. Through this work, the GIC collected detailed health records about thousands of state employees. In an effort to spur research, the GIC decided to release these records to

researchers. However, they did not share all of their data; rather, they "anonymized" these

data by removing information such as names and addresses. However, they left other

information that they thought could be useful for researchers such as demographic information (zip code, birth date, ethnicity, and sex) and medical information (visit data, diagnosis, procedure) (figure 6.4) (Ohm 2010). Unfortunately, this "anonymization" was not sufficient to protect the data. Name Home Address Zip code Birth date Zip code Sex Birth date Sex **Ethnicity Ethnicity** Visit date Visit date Diagnosis



data. Using this fact, Sweeney was able to locate Weld's medical records, and, to inform him of her feat, she mailed him a copy of his records (Ohm 2010). Ethnicity Name Zip code Visit date Home Address Birth date Diagnosis **Party**

Sex

Date registered

Procedure

Medication

one shared Weld's zip code. Thus, the voting data showed that anyone in the medical

data with Weld's combination of birth date, gender, and zip code was William Weld. In

essence, these three pieces of information provided a unique fingerprint to him in the

"Anonymized" Voting records medical records Figure 6.5: Re-idenification of "anonymized" data. Latanya Sweeney combined the "anonymized" health records with voting records in order to find the medical records of Governor William Weld Adapted from Sweeney (2002), figure 1. Sweeney's work illustrates the basic structure of *re-identification attacks*—to adopt a term from the computer security community. In these attacks, two data sets, neither of which by itself reveals sensitive information, are linked, and through this linkage, sensitive information is exposed. In response to Sweeney's work, and other related work, researchers now generally remove much more information—all so-called "personally identifying information" (PII)

even after "anonymization." However, the examples that I'm about to give suggest that

all data are potentially identifiable and all data are potentially sensitive. In other words,

rather than thinking that informational risk applies to a small subset of projects, we

Both aspects of this reorientation are illustrated by the Netflix Prize. As described in

chapter 5, Netflix released 100 million movie ratings provided by almost 500,000

members, and had an open call where people from all over the world submitted

should assume that it applies—to some degree—to all projects.

social researchers need to change their thinking. As a first step, it is wise to assume that

Just two weeks after the data were released, Arvind Narayanan and Vitaly Shmatikov (2008) showed that it was possible to learn about specific people's movie preferences. The trick to their re-identification attack was similar to Sweeney's: merge together two information sources, one with potentially sensitive information and no obviously identifying information and one that contains people's identities. Each of these data sources may be individually safe, but when they are combined, the merged dataset can create informational risk. In the case of the Netflix data, here's how it could happen. Imagine that I choose to share my thoughts about action and comedy movies with my

Even though the Netflix data can be re-identified in either a targeted or broad attack, it still might appear to be low risk. After all, movie ratings don't seem very sensitive. While that might be true in general, for some of the 500,000 people in the dataset, movie ratings might be quite sensitive. In fact, in response to the re-identification, a closeted lesbian woman joined a class-action suit against Netflix. Here's how the problem was expressed in their lawsuit (Singel 2009): "[M]ovie and rating data contains information of a ... highly personal and sensitive nature. The member's movie data exposes a Netflix member's personal interest and/or struggles with various highly personal issues, including sexuality, mental illness, recovery from alcoholism, and victimization from incest, physical abuse, domestic violence, adultery, and rape." The re-identification of the Netflix Prize data illustrates both that all data are potentially identifiable and that all data are potentially sensitive. At this point, you might think that this only applies to data that purport to be about people. Surprisingly, that is not the case. In response to a Freedom of Information Law request, the New York City Government released records of every taxi ride in New York in 2013, including the pickup

sensitive information about people. To illustrate, he looked at all trips starting at the

city government had this in mind when it released the data. In fact, this same technique

could be used to find the home addresses of people who visit any place in the city—a

medical clinic, a government building, or a religious institution.

identifiable and all data are potentially sensitive.

Safe projects

Safe people

Safe data

Safe settings

Safe output

harm if a leak does somehow occur. The specifics of data protection plans, such as which form of encryption to use, will change over time, but the UK Data Services helpfully organizes the elements of a data protection plan into five categories that they call the *five* safes: safe projects, safe people, safe settings, safe data, and safe outputs (table 6.2) (Desai, Ritchie, and Welpton 2016). None of the five safes individually provide perfect protection. But together they form a powerful set of factors that can decrease informational risk. Table 6.2: The "Five Safes" are Principles for Designing and Executing a Data Protection Plan (Desai, Ritchie, and Welpton 2016) **Action** Safe

Limits projects with data to those that are ethical

(e.g., people who have undergone ethical training)

locked room) and software (e.g., password protection,

Research output is reviewed to prevent accidental privacy

encrypted) protection

process where informational risk is particularly salient is data sharing with other

breaches

Access is restricted to people who can be trusted with data

Unfortunately, there is no simple solution to the facts that all data are potentially

researchers. Data sharing among scientists is a core value of the scientific endeavor, and it greatly facilitates the advancement of knowledge. Here's how the UK House of Commons described the importance of data sharing (Molloy 2011): "Access to data is fundamental if researchers are to reproduce, verify and build on results that are reported in the literature. The presumption must be that, unless there is a strong reason otherwise, data should be fully disclosed and made publicly available." Yet, by sharing your data with another researcher, you may be increasing informational risk to your participants. Thus, it may seem that data sharing creates a fundamental tension between the obligation to share data with other scientists and the obligation to minimize informational risk to participants. Fortunately, this dilemma is not as severe as it appears. Rather, it is better to think about data sharing as falling along a continuum, with each point on that continuum providing a different mix of benefits to society and risk to participants (figure 6.6).

At one extreme, you can share your data with no one, which minimizes risk to

participants but also minimizes gains to society. At the other extreme, you can release

releasing data, release and forget offers both higher benefits to society and higher risk to

participants. In between these two extreme cases are a range of hybrids, including what

I'll call a walled garden approach. Under this approach, data are shared with people who

meet certain criteria and who agree to be bound by certain rules (e.g., oversight from an

benefits of release and forget with less risk. Of course, such an approach creates many

IRB and a data protection plan). The walled garden approach provides many of the

for Political and Social Research at the University of Michigan.

and forget, where data are "anonymized" and posted for everyone. Relative to not

Release and forget Walled garden Benefit to society No data sharing

So, where should the data from your study be on the continuum of no sharing, walled garden, and release and forget? This depend on the details of your data: researchers must balance Respect for Persons, Beneficence, Justice, and Respect for Law and Public Interest. Viewed from this perspective, data sharing is not a distinctive ethical conundrum; it is just one of the many aspects of research in which researchers have to find an appropriate ethical balance.

Some critics are generally opposed to data sharing because, in my opinion, they are

focused on its risks—which are undoubtedly real—and are ignoring its benefits. So, in

Risk to participants

Figure 6.6: Data release strategies can fall along a continuum. Where you should be on this continuum depends

on the specific details of your data, and third-party review may help you decide the appropriate balance of risk

and benefit in your case. The exact shape of this curve depends on the specifics of the data and research goals

(Goroff 2015).

order to encourage focus on both risks and benefits, I'd like to offer an analogy. Every year, cars are responsible for thousands of deaths, but we do not attempt to ban driving. In fact, a call to ban driving would be absurd because driving enables many wonderful things. Rather, society places restrictions on who can drive (e.g., the need to be a certain age and to have passed certain tests) and how they can drive (e.g., under the speed limit). Society also has people tasked with enforcing these rules (e.g., police), and we punish people who are caught violating them. This same kind of balanced thinking that society applies to regulating driving can also be applied to data sharing. That is, rather than making absolutist arguments for or against data sharing, I think we will make the most progress by focusing on how we can decrease the risks and increase the benefits from data sharing.

To conclude, informational risk has increased dramatically, and it is very hard to predict and quantify. Therefore, it is best to assume that all data are potentially identifiable and potentially sensitive. To decrease informational risk while doing research, researchers can create and follow a data protection plan. Further, informational risk does not prevent researchers from sharing data with other scientists.

records included information such as name, address, zip code, birth date, and gender. The fact that the medical data file and the voter file shared fields—zip code, birth date, and sex-meant that Sweeney could link them. Sweeney knew that Weld's birthday was July 31, 1945, and the voting records included only six people in Cambridge with that birthday. Further, of those six people, only three were male. And, of those three men, only

(Narayanan and Shmatikov 2010)—during the process of "anonymization." Further, many researchers now realize that certain data—such as medical records, financial records, answers to survey questions about illegal behavior—are probably too sensitive to release

algorithms that could improve Netflix's ability to recommend movies. Before releasing the data, Netflix removed any obvious personally identifying information, such as names. They also went an extra step and introduced slight perturbations in some of the records (e.g., changing some ratings from 4 stars to 3 stars). They soon discovered, however, that despite their efforts, the data were still by no means anonymous.

co-workers, but that I prefer not to share my opinion about religious and political movies. My co-workers could use the information that I've shared with them to find my records in the Netflix data; the information that I share could be a unique fingerprint just like William Weld's birth date, zip code, and sex. Then, if they found my unique fingerprint in the data, they could learn my ratings about all movies, including movies that I choose not to share. In addition to this kind of targeted attack focused on a single person, Narayanan and Shmatikov also showed that it was possible to do a broad attack—one involving many people—by merging the Netflix data with personal and movie rating data that some people have chosen to post on the Internet Movie Database (IMDb). Quite simply, any information that is a unique fingerprint to a specific person—even their set of movie ratings—can be used to identify them.

and drop off times, locations, and fare amounts (recall from chapter 2 that Farber (2015) used similar data to test important theories in labor economics). These data about taxi trips might seem benign because they do not seem to provide information about people, but Anthony Tockar realized that this taxi dataset actually contained lots of potentially Hustler Club—a large strip club in New York—between midnight and 6 a.m. and then found their drop-off locations. This search revealed—in essence—a list of addresses of some people who frequented the Hustler Club (Tockar 2014). It is hard to imagine that the

These two cases of the Netflix Prize and the New York City taxi data show that relatively skilled people can fail to correctly estimate the informational risk in the data that they release—and these cases are by no means unique (Barbaro and Zeller 2006; Zimmer 2010; Narayanan, Huey, and Felten 2016). Further, in many such cases, the problematic data are still freely available online, indicating the difficulty of ever undoing a data release. Collectively, these examples—as well as research in computer science about privacy lead to an important conclusion. Researchers should assume that all data are potentially identifiable and that all data are potentially sensitive. However, one way to reduce informational risk while you are working with data is to create and follow a data protection plan. This plan will decrease the chance that your data will leak and will decrease the

Data are de-identified and aggregated to the extent possible Data are stored in computers with appropriate physical (e.g., In addition to protecting your data while you are using them, one step in the research

questions—who should have access, under what conditions, and for how long, who should pay to maintain and police the walled garden, etc.—but these are not insurmountable. In fact, there are already working walled gardens in place that researchers can use right now, such as the data archive of the Inter-university Consortium

6.6.3 Privacy

Privacy is a right to the appropriate flow of information.

A third area where researchers may struggle is *privacy*. As Lowrance (2012) put it quite succinctly: "privacy should be respected because people should be respected." Privacy, however, is a notoriously messy concept (Nissenbaum 2010, chap. 4), and, as such, it is a difficult one to use when trying to make specific decisions about research.

A common way to think about privacy is with a public/private dichotomy. By this way of thinking, if information is publicly accessible, then it can be used by researchers without concerns about violating people's privacy. But this approach can run into problems. For example, in November 2007, Costas Panagopoulos sent letters about an upcoming election to everyone in three towns. In two towns—Monticello, Iowa and Holland, Michigan—Panagopoulos promised/threatened to publish a list of people who had voted in the newspaper. In the other town—Ely, Iowa—Panagopoulos promised/threatened to publish a list of people who had not voted in the newspaper. These treatments were designed to induce pride and shame (Panagopoulos 2010) because these emotions had been found to impact turnout in earlier studies (Gerber, Green, and Larimer 2008). Information about who votes and who doesn't is public in the United States; anyone can access it. So, one could argue that because this voting information is already public, there is no problem with a researcher publishing it in the newspaper. On the other hand, something about that argument feels wrong to some people.

As this example illustrates, the public/private dichotomy is too blunt (boyd and Crawford 2012; Markham and Buchanan 2012). A better way to think about privacy—one especially designed to handle issues raised by the digital age—is the idea of *contextual integrity* (Nissenbaum 2010). Rather than considering information as public or private, contextual integrity focuses on the flow of information. According to Nissenbaum (2010), "a right to privacy is neither a right to secrecy or a right to control but a right to appropriate flow of personal information."

The key concept underlying contextual integrity is *context-relative informational norms* (Nissenbaum 2010). These are norms that govern the flow of information in specific settings, and they are determined by three parameters:

- actors (subject, sender, recipient)
- attributes (types of information)
- transmission principles (constraints under which information flows)

Thus, when you as a researcher are deciding whether to use data without permission it is helpful to ask, "Does this use violate context-relative informational norms?" Returning to the case of Panagopoulos (2010), in this case, having an outside researcher publish lists of voters or nonvoters in the newspaper seems likely to violate informational norms. This is probably not how people expect information to flow. In fact, Panagopoulos did not follow through on his promise/threat because local election officials traced the letters to him and persuaded him that it was not a good idea (Issenberg 2012, 307).

The idea of context-relative informational norms can also help evaluate the case I discussed at the beginning of the chapter regarding the use of mobile phone call logs to track mobility during the Ebola outbreak in West Africa in 2014 (Wesolowski et al. 2014). In this setting, one could imagine two different situations:

- Situation 1: sending complete call log data [attributes]; to governments of incomplete legitimacy [actors]; for any possible future use [transmission principles]
- Situation 2: sending partially anonymized records [attributes]; to respected university researchers [actors]; for use in response to the Ebola outbreak and subject to the oversight of university ethical boards [transmission principles]

Even though in both of these situations call data are flowing out of the company, the informational norms concerning these two situations are not the same because of differences between the actors, attributes, and transmission principles. Focusing on only one of these parameters can lead to overly simplistic decision-making. In fact, Nissenbaum (2015) emphasizes that none of these three parameters can be reduced to the others, nor can any one of them individually define informational norms. This three-dimensional nature of informational norms explains why past efforts—which have focused on either attributes or transmission principles—have been ineffective at capturing common-sense notions of privacy.

One challenge with using the idea of context-relative informational norms to guide decisions is that researchers might not know them ahead of time and they are very hard to measure (Acquisti, Brandimarte, and Loewenstein 2015). Further, even if some research would violate contextual-relative informational norms that does not automatically mean that the research should not happen. In fact, chapter 8 of Nissenbaum (2010) is entirely about "Breaking Rules for Good." Despite these complications, context-relative informational norms are still a useful way to reason about questions related to privacy.

Finally, privacy is an area where I've seen misunderstandings between researchers who prioritize Respect for Persons and those who prioritize Beneficence. Imagine the case of a public health researcher who, in an effort to prevent the spread of a novel infectious disease, secretly watched people taking showers. Researchers focusing on Beneficence would focus on the benefits to society from this research and might argue that there was no harm to participants if the researcher did her spying without detection. On the other hand, researchers who prioritize Respect for Persons would focus on the fact that the researcher was not treating people with respect and might argue that harm was created by violating the privacy of participants, even if the participants were not aware of the spying. In other words, to some, violating people's privacy is a harm in and of itself.

In conclusion, when reasoning about privacy, it is helpful to move beyond the overly simplistic public/private dichotomy and to reason instead about context-relative informational norms, which are made up of three elements: actors (subject, sender, recipient), attributes (types of information), and transmission principles (constraints under which information flows) (Nissenbaum 2010). Some researchers evaluate privacy in terms of the harm that could result from its violation, whereas other researchers view the violation of privacy as a harm in and of itself. Because notions of privacy in many digital systems are changing over time, vary from person to person, and vary from situation to situation (Acquisti, Brandimarte, and Loewenstein 2015), privacy is likely to be a source of difficult ethical decisions for researchers for some time to come.

6.6.4 Making decisions in the face of uncertainty

Uncertainty need not lead to inaction.

The fourth and final area where I expect researchers to struggle is making decisions in the face of uncertainty. That is, after all the philosophizing and balancing, research ethics involves making decisions about what to do and what not to do. Unfortunately, these decisions often must be made based on incomplete information. For example, when designing Encore, researchers might have wished to know the probability that it would cause someone to be visited by the police. Or, when designing Emotional Contagion, researchers might have wished to know the probability that it could trigger depression in some participants. These probabilities were probably extremely low, but they were unknown before the research takes place. And, because neither project publicly tracked information about adverse events, these probabilities are still not generally known.

Uncertainties are not unique to social research in the digital age. When the Belmont Report described the systematic assessment of risks and benefits, it explicitly acknowledged these would be difficult to quantify exactly. These uncertainties, however, are more severe in the digital age, in part because we have less experience with this type of research and in part because of the characteristics of the research itself.

Given these uncertainties, some people seem to advocate for something like "better safe" than sorry," which is a colloquial version of the Precautionary Principle. While this approach appears reasonable—perhaps even wise—it can actually cause harm; it is chilling to research; and it causes people to take an excessively narrow view of the situation (Sunstein 2005). In order to understand the problems with the Precautionary Principle, let's consider Emotional Contagion. The experiment was planned to involve about 700,000 people, and there was certainly some chance that people in the experiment would suffer harm. But there was also some chance that the experiment could yield knowledge that would be beneficial to Facebook users and to society. Thus, while allowing the experiment was a risk (as has been amply discussed), preventing the experiment would also have been a risk, because it could have produced valuable knowledge. Of course, the choice was not between doing the experiment as it occurred and not doing the experiment; there were many possible modifications to the design that might have brought it into a different ethical balance. However, at some point, researchers will have the choice between doing a study and not doing it, and there are risks in both action and inaction. It is inappropriate to focus only on the risks of action. Quite simply, there is no risk-free approach.

Moving beyond the Precautionary Principle, one important way to think about making decisions given uncertainty is the minimal risk standard. This standard attempts to benchmark the risk of a particular study against the risks that participants undertake in their daily lives, such as playing sports and driving cars (Wendler et al. 2005). This approach is valuable because assessing whether something meets the minimal risk standard is easier than assessing the actual level of risk. For example, in Emotional Contagion, before the study began, the researchers could have compared the emotional content of News Feeds in the experiment with that of other News Feeds on Facebook. If they had been similar, then the researchers could have concluded that the experiment met the minimal risk standard (M. N. Meyer 2015). And they could make this decision even if they didn't know the absolute level of risk. The same approach could have been applied to Encore. Initially, Encore triggered requests to websites that were known to be sensitive, such as those of banned political groups in countries with repressive governments. As such, it was not minimal risk for participants in certain countries. However, the revised version of Encore—which only triggered requests to Twitter, Facebook, and YouTube—was minimal risk because requests to those sites are triggered during normal web browsing (Narayanan and Zevenbergen 2015).

A second important idea when making decisions about studies with unknown risk is power analysis, which allows researchers to calculate the sample size they will need to reliably detect an effect of a given size (Cohen 1988). If your study might expose participants to risk—even minimal risk—then the principle of Beneficence suggests that you should impose the smallest amount of risk needed to achieve your research goals. (Think back to the Reduce principle in chapter 4.) Even though some researchers have an obsession with making their studies as big as possible, research ethics suggests that researchers should make their studies as small as possible. Power analysis is not new, of course, but there is an important difference between the way that it was used in the analog age and how it should be used today. In the analog age, researchers generally did power analysis to make sure that their study was not too small (i.e., under-powered). Now, however, researchers should do power analysis to make sure that their study is not too big (i.e., over-powered).

The minimal risk standard and power analysis help you reason about and design studies, but they don't provide you with any new information about how participants might feel about your study and what risks they might experience from participating in it. Another way to deal with uncertainty is to collect additional information, which leads to ethicalresponse surveys and staged trials.

In ethical-response surveys, researchers present a brief description of a proposed research project and then ask two questions:

- (Q1) "If someone you cared about were a candidate participant for this experiment, would you want that person to be included as a participant?": [Yes], [I have no preferences], [No]
- (Q2) "Do you believe that the researchers should be allowed to proceed with this experiment?": [Yes], [Yes, but with caution], [I'm not sure], [No]

Following each question, respondents are provided a space in which they can explain their answer. Finally, respondents—who could be potential participants or people recruited from a microtask labor markets (e.g., Amazon Mechanical Turk)—answer some basic demographic questions (Schechter and Bravo-Lillo 2014).

Ethical-response surveys have three features that I find particularly attractive. First, they happen before a study has been conducted, and therefore they can prevent problems before the research starts (as opposed to approaches that monitor for adverse reactions). Second, the respondents in ethical-response surveys are typically not researchers, and so this helps researchers see their study from the perspective of the public. Finally, ethical-response surveys enable researchers to pose multiple versions of a research project in order to assess the perceived ethical balance of different versions of the same project. One limitation, however, of ethical-response surveys is that it is not clear how to decide between different research designs given the survey results. But, despite these limitations, ethical-response surveys appear to be helpful; in fact, Schechter and Bravo-Lillo (2014) report abandoning a planned study in response to concerns raised by participants in an ethical-response survey.

While ethical-response surveys can be helpful for assessing reactions to proposed research, they cannot measure the probability or severity of adverse events. One way that medical researchers deal with uncertainty in high-risk settings is to perform staged trials —an approach that might be helpful in some social research. When testing the effectiveness of a new drug, researchers do not immediately jump to a large randomized clinical trial. Rather, they run two types of studies first. Initially, in a phase I trial, researchers are particularly focused on finding a safe dose, and these studies involve a small number of people. Once a safe dose has been determined, phase II trials assess the efficacy of the drug; that is, its ability to work in a best-case situation (Singal, Higgins, and Waljee 2014). Only after phase I and II studies have been completed is a new drug allowed to be assessed in a large randomized controlled trial. While the exact structure of staged trials used in the development of new drugs may not be a good fit for social research, when faced with uncertainty, researchers could run smaller studies explicitly

focused on safety and efficacy. For example, with Encore, you could imagine the researchers starting with participants in countries with strong rule of law.

Together, these four approaches—the minimal risk standard, power analysis, ethical-

face of uncertainty. Uncertainty need not lead to inaction.

response surveys, and staged trials—can help you proceed in a sensible way, even in the

Powered by Open Review Toolkit





6.7 Practical tips

In addition to high-minded ethical principles, there are practical issues in research ethics.

In addition to the ethical principles and frameworks described in this chapter, I'd also like to offer three practical tips based on my personal experience conducting, reviewing, and discussing social research in the digital age: the IRB is a floor, not a ceiling; put yourself in everyone else's shoes; and think of research ethics as continuous, not discrete.



6.7.1 The IRB is a floor, not a ceiling

Many researchers seem to hold contradictory views of the IRB. On the one hand, they consider it to be a bumbling bureaucracy. Yet, at the same time, they also consider it to be the final arbitrator of ethical decisions. That is, many researchers seem to believe that if the IRB approves it, then it must be OK. If we acknowledge the very real limitations of IRBs as they currently exist—and there are many of them (Schrag 2010, 2011; Hoonaard 2011; Klitzman 2015; King and Sands 2015; Schneider 2015)—then we as researchers must take on additional responsibility for the ethics of our research. The IRB is a floor not a ceiling, and this idea has two main implications.

First, *the IRB is a floor* means that if you are working at an institution that requires IRB review, then you should follow those rules. This may seem obvious, but I've noticed that some people seem to want to avoid the IRB. In fact, if you are working in ethically unsettled areas, the IRB can be a powerful ally. If you follow their rules, they should stand behind you should something go wrong with your research (King and Sands 2015). And if you don't follow their rules, you could end up out on your own in a very difficult situation.

Second, the IRB is not a ceiling means that just filling out your forms and following the rules is not enough. In many situations you as the researcher are the one who knows the most about how to act ethically. Ultimately, you are the researcher, and the ethical responsibility lies with you; it is your name on the paper.

One way to ensure that you treat the IRB as a floor and not a ceiling is to include an ethical appendix in your papers. In fact, you could draft your ethical appendix before your study even begins, in order to force yourself to think about how you will explain your work to your peers and the public. If you find yourself uncomfortable while writing your ethical appendix, then your study might not strike the appropriate ethical balance. In addition to helping you diagnose your own work, publishing your ethical appendices will help the research community discuss ethical issues and establish appropriate norms based on examples from real empirical research. table 6.3 presents empirical research papers that I think have good discussions of research ethics. I don't agree with every claim by the authors in these discussions, but they are all examples of researchers acting with *integrity* in the sense defined by Carter (1996): in each case, (1) the researchers decide what they think is right and what is wrong; (2) they act based on what they have decided, even at personal cost; and (3) they publicly show that they are acting based on their ethical analysis of the situation.

Table 6.3: Papers with Interesting Discussions of the Ethics of their Research

Field experiments without consent Avoiding contextual harm Field experiments in developing country Research on sensitive topic Complex consent issues
Avoiding contextual harm Field experiments in developing country Research on sensitive topic
Field experiments in developing country Research on sensitive topic
Research on sensitive topic
·
Complex consent issues
Remediation of possible harms
Research without consent
Balancing risks and benefits when risks are hard to quantify
Social implications of research
Jsing leaked data files
Field experiments without consent
/iolated terms of service





6.7.2 Put yourself in everyone else's shoes

Often researchers are so focused on the scientific aims of their work that they see the world only through that lens. This myopia can lead to bad ethical judgment. Therefore, when you are thinking about your study, try to imagine how your participants, other relevant stakeholders, and even a journalist might react to your study. This perspective taking is different than imaging how *you* would feel in each of these positions. Rather, it is trying to imagine how these *other people* will feel, a process that is likely to induce empathy (Batson, Early, and Salvarani 1997). Thinking through your work from these different perspectives can help you foresee problems and move your work into better ethical balance.

Further, when imagining your work from the perspective of others, you should expect that they are likely to fixate on vivid worst-case scenarios. For example, in response to Emotional Contagion, some critics focused on the possibility that it might have triggered suicide, a low-probability but extremely vivid worst-case scenario. Once people's emotions have been activated and they focus on worst-case scenarios, they may completely lose track of the probability of this worst-case event occurring (Sunstein 2002). The fact that people might respond emotionally, however, does not mean that you should dismiss them as uninformed, irrational, or stupid. We should all be humble enough to realize that none of us have a perfect view of ethics.





6.7.3 Think of research ethics as continuous, not discrete

Debate about the ethics of social research in the digital age frequently happens in binary terms; for example, Emotional Contagion was either ethical or it was not ethical. This binary thinking polarizes discussion, hinders efforts to develop shared norms, promotes intellectual laziness, and absolves researchers whose research is labeled "ethical" from their responsibility to act more ethically. The most productive conversations that I've seen involving research ethics move beyond this binary thinking to a continuous notion about research ethics.

A major practical problem with a binary conception of research ethics is that it polarizes discussion. Calling Emotional Contagion "unethical" lumps it together with true atrocities in a way that is not helpful. Rather, it is more helpful and appropriate to talk specifically about the aspects of the study that you find problematic. Moving away from binary thinking and polarizing language is not a call for us to use muddled language to hide unethical behavior. Rather, a continuous notion of ethics will, I think, lead to more careful and precise language. Further, a continuous notion of research ethics clarifies that everyone—even researchers who are doing work that is already considered "ethical"—should strive to create an even better ethical balance in their work.

A final benefit of a move toward continuous thinking is that it encourages intellectual humility, which is appropriate in the face of difficult ethical challenges. The questions of research ethics in the digital age are difficult, and no single person should be overly confident in her own ability to diagnose the correct course of action.





6.8 Conclusion

Social research in the digital age raises new ethical issues. But these issues are not insurmountable. If we, as a community, can develop shared ethical norms and standards that are supported both by researchers and the public, then we can harness the capabilities of the digital age in ways that are responsible and beneficial to society. This chapter represents my attempt to move us in that direction, and I think the key will be for researchers to adopt principles-based thinking, while continuing to follow appropriate rules.

In section 6.2, I described three digital-age research projects that have generated ethical debate. Then, in section 6.3 I described what I think is the fundamental reason for ethical uncertainty in digital-age social research: rapidly increasing power for researchers to observe and experiment on people without their consent or even awareness. These capabilities are changing faster than our norms, rules, and laws. Next, in section 6.4, I described four existing principles that can guide your thinking: Respect for Persons, Beneficence, Justice, and Respect for Law and Public Interest. Then, in section 6.5, I summarized two broad ethical frameworks—consequentialism and deontology—that can help you with one of the deepest challenges that you might face: when is it appropriate for you to take ethically questionable means in order to achieve an ethically appropriate end. These principles and ethical frameworks will enable you to move beyond focusing on what is permitted by existing regulations and increase your ability to communicate your reasoning with other researchers and the public.

With that background, in section 6.6, I discussed four areas that are particularly challenging for digital-age social researchers: informed consent (section 6.6.1), understanding and managing informational risk (section 6.6.2), privacy (section 6.6.3), and making ethical decisions in the face of uncertainty (section 6.6.4). Finally, in section 6.7, I concluded with three practical tips for working in an area with unsettled ethics.

In terms of scope, this chapter has focused on the perspective of an individual researcher seeking generalizable knowledge. As such, it leaves out important questions about improvements to the system of ethical oversight of research; questions about regulation of the collection and use of data by companies; and questions about mass surveillance by governments. These other questions are obviously complex and difficult, but it is my hope that some of the ideas from research ethics will be helpful in these other contexts.