4

Data Management Practices

░ INTRODUCTION

Data is the plural form of *datum,* the Latin word for something given or admitted, known, or assumed. The sense is that data represent or describe some aspect of the real world accurately or truthfully. They are facts representing what is in the world. Researchers collect data so they can test their theories and hypotheses against the reality of the world and try to determine whether the theories and hypotheses are correct. The goals of a theory are to describe and explain what the world is like and how it works. Theories form the bases for predictions. Without reliable and valid data, theories cannot be tested and thus have little value.

Consequently, data are the raw material of research. Data can be qualitative or quantitative. Regardless of the format of the data, whether they consist of numbers, written or spoken words, photographs, video recordings, or other materials, the researcher has the responsibility to manage these resources in an effective and responsible way in order to permit the most objective and rigorous test of theory while at the same time protecting the interests of research subjects.

According to the Office of Research Integrity, located in the U.S. Department of Health and Human Services, data management consists of several elements:[1]

- Data ownership
- Data collection
- Data storage
- Data protection
- Data retention
- Data analysis
- Data sharing
- Data reporting

Issues in data management are related to the responsible conduct of research in a variety of ways. Ethical issues can arise when there is conflict over who owns the data and thus who gets to use them. There can be problems with how data will be collected and stored. Especially important are concerns about the protection of data that might contain personal or other sensitive information. Such data must be stored in such a way as to keep them confidential. They cannot be freely shared or left unsecured, because the subjects of the data could be harmed. How long should researchers keep data in case they may need to reexamine them? Data analysis is the basis of all theory testing, and incorrect analysis is not always obvious. Because research is a search for truth and data are the foundation of much of that search, data must be treated with careful attention.

Outright falsification of data is clearly unethical behavior. However, more subtle ways to stray from the ethical path can confront researchers, especially novices, because so many different stakeholders may be involved in the enterprise and data analysis comprises so many elements. Bruce L. Brown and Dawson Hedges present some good general guidelines to follow:

> Methodological rigor is closely related to ethical vigilance: When research, statistical calculation, and data presentation can be done better and more accurately, they should be. That is, there is an ethical imperative to demand and use the highest standards of research and data presentation.[2]

This suggestion should be a basic guideline for data collection, analysis, and presentation. Beyond that, other ethical issues arise in the publication of results. Brown and Hedges suggest that it is unethical to withhold negative results because their absence from the public record could bias the results of meta-analyses. In addition, they state: "It is a breach of one's ethical responsibility to coworkers to let a paper sit in one's drawer unpublished. Continuing on until the paper reaches publication is an ethical responsibility for those committed to a career in scholarly work" (p. 382).

Brown and Hedges emphasize that when reporting the results of statistical analyses, researchers should provide complete and detailed information regarding such key points as the sample size, response rate, level of statistical significance, effect sizes, and information bearing on the assumptions of the data. They recommend as a sound practice the regular and rigorous use of power analysis to plan studies with sufficient sample sizes to provide reliable tests of hypotheses. Small-scale studies with negative results might be misleading and are a waste of time and effort, and thus borderline unethical. Reporting of effect size is important to help readers gauge the likely value of a study. Researchers should be willing to share data in ways that preserve the interests of those who collected them and at the same time permit other interpretations that might be of great value.

Our cases in this chapter focus on ethical issues in data analysis. By themselves, data are only raw materials, but they contain information about the world. This information is used by the researcher to test theory and make predictions. Thus, how the data are managed has an important bearing on the conclusions drawn from the study and the uses to which these conclusions are put. The following cases present situations in which the researchers must make decisions that have substantive and ethical consequences.

Notes

1. Office of Research Integrity, U.S. Department of Health and Human Services, "Guidelines for Responsible Data Management in Scientific Research," http://ori.hhs .gov/education/products/clinicaltools/data.pdf.

2. Bruce L. Brown and Dawson Hedges, "Use and Misuse of Quantitative Methods," in *The Handbook of Social Research Ethics,* ed. Donna M. Mertens and Pauline E. Ginsberg (Thousand Oaks, CA: Sage, 2009), 375. Further citations of this work appear in text.

〽 CASE 1. PRUNING DATA

Michelle and her team of research assistants are working with a family clothing store in the local mall trying to answer questions about how merchandise displays affect customer shopping patterns. Michelle has done this kind of research before, but she is working with a new team of graduate students on this project. The team both collects video footage and conducts paper-and-pencil surveys.

Michelle hired a technical expert to install video cameras around the store that could view customers unobtrusively. Signs were placed at the store entrance to let patrons know that cameras were in use. The cameras were turned on whenever the store was open, and they visually covered the entire floor space. During store hours, customers were videotaped walking around the store, handling the merchandise, taking clothes into the dressing rooms, and speaking with the sales staff. (The dressing rooms were not under video surveillance.) Upon reviewing the videotapes, the research team first classified each customer's age, gender, and ethnic origin and then followed that person's activities on the tape to record his or her movement and other activities in the store. These data were then analyzed using a variety of statistical methods.

During certain times, part of the research team stood outside the store to administer paper-and-pencil surveys to customers exiting the store. Team members

approached customers and asked them to participate. If a customer consented, a team member gave the customer the survey form on a clipboard to fill out him- or herself. Michelle personally trained the graduate students in how to collect the surveys, which were an integral part of the research plan.

Over the course of three weekends, the researchers collected video of 730 customers shopping in the store. The survey team did less well, garnering only 263 completed surveys. Team members missed many customers while approaching others, and some customers did not agree to fill out the survey. Despite the discrepancy in numbers, the researchers agreed to include both data sets in their analysis.

As data analysis proceeded, the team noticed something. Results from the video data and the survey data did not match. In fact, some hypotheses that were supported by the video footage were not supported by the survey responses and vice versa. Needless to say, the group was disheartened by these findings, and a number of potential solutions were floated.

Michelle has been working with the survey data and feels that a fair number of the customers who filled out the forms were not diligent and marked responses randomly to speed the process. She feels that she can identify these approximately 70 respondents and delete them from the analysis. Will the results then change?

The results of the video footage support the group's original hypotheses, and the number of subjects is impressive. The results of the survey contradict the hypotheses, and the subject pool is small. What to do?

Learning Objectives

1. Students should become acquainted with some of the real problems that arise in the collection and handling of data.

2. Students are faced with an ethical decision-making task: How should Michelle and the team proceed?

Questions

1. How should Michelle handle the discrepancy between the findings of the survey and the findings of the observational data?

2. How would you go about identifying and discarding the "suspicious" survey respondents?

3. How much detail should Michelle provide in the written report? Full disclosure or less than that?

4. What are Michelle's options? Do they have ethical dimensions, or are these just "technical, methodological" issues?

5. What are Michelle's responsibilities to her client and to the graduate students?

⑈ CASE 2. SECONDARY DATA

Mary Smith is a 44-year-old doctoral student in the insurance program at a major university in the eastern United States. She left an executive job in Manhattan to retreat from the stress of the corporate world into the calm of the ivory tower. She is wondering about that decision today as she looks at her dissertation data.

As part of her dissertation research, Mary is collecting records from business casualty insurance companies. The records are of insurance policies and associated claims from a variety of business categories. Her dissertation adviser helped her get a list of contacts from five large insurers who might be able to provide more records, but they have not been very forthcoming with data. It has taken Mary more than a year to come up with the 51 cases she has in her possession now. It seems that each company has different reasons for not providing the client data. Initially, Mary's proposal called for 200 cases in order to test her hypotheses adequately. Mary is woefully short of data. She is also tired of being a student and is missing her former Manhattan lifestyle.

At a regularly scheduled progress meeting with her adviser, Mary asks her for help. The professor agrees to call in a few favors from industry friends and see what she can do for her student. In the meantime, she advises, Mary should try to find some new contacts in some different insurance companies. This last piece of advice frustrates Mary, as developing the current contacts has taken her 14 months and yielded only 51 data points.

One other option is on the table for Mary. When she left her insurance industry job 5 years ago, she took with her some data tapes. They contained data she used at work, with the names of the companies represented masked. The data are also old enough that using them would not reveal anything about her old company's clients. Some records in the tapes go back as far as 12 years now. Mary has also used parts of the data tapes in other research studies. On the other hand, the old data would blend in with the new. It would not be obvious to anyone that some of the data are "historical" rather than current. The only problem she can see with using these data is that her adviser and doctoral committee do not want her to.

Although the committee members have objected to the use of the older data, citing the data's datedness and the possibility that they are biased because they

were not collected "independently," Mary decides to take a look at the older data. When she merges the old and new data sets, she finds that she gets good results on her hypotheses tests.

Six months pass, the contacts from her adviser do not come through, and Mary analyzes and writes up the results from her merged data set and turns in her final chapters.

Learning Objectives

1. Students should come to see that honest and fair use of data is not always a simple matter.

2. Students should put themselves in Mary's position and propose alternative courses of action for her.

3. Students should gain an understanding of the difficulties of data collection and the ethical dilemmas it can pose.

Questions

1. Are there any ethical problems in terms of Mary's former employer and her use of the historical data?

2. Is Mary's committee being fair in forbidding her to use the "old" data?

3. Could Mary's committee have been more helpful?

4. What could Mary do to make her use of the old data more appropriate?

5. How much should Mary reveal to her committee about her data?

6. What are the potential downsides of doing what Mary has done? How might her actions come back to haunt her?

⟍ CASE 3. DATA OWNERSHIP

As part of his graduate program in the Communications Department, Dave is working with his favorite professor, Dr. Jaynes, to collect data on the outcomes by age of children treated for sigmatism, or lisp. The professor is the principal investigator (PI) on the study, which seeks to determine the optimal age for treating this speech impediment so that treatment can be both most effective and least expensive. Dr. Jaynes had the idea for the project and applied for and received

the grant. He also wrote the surveys and the general protocol for the study. Dr. Jaynes has enlisted Dave to put all of that into action. He felt it would be good research experience for Dave, who has been calling, mailing, and e-mailing speech therapy clinics all over the United States in order to get historical records and enlist new children in the study. Although this is not a paid job, Dave has been routinely working more than 30 hours per week on the project. In addition, while this presses into his time for his other studies, he feels that there is a good chance the resulting data will be just what he needs for his dissertation, which he plans on starting in one more year.

Over the 10 months Dave has been at work on this project, he has collected data from more than 1,300 cases and enlisted 425 current children in the study. This will be the largest study to date of outcomes of sigmatism by age of onset of therapy, and because treatment of this functional speech disorder is common, the research is valuable to patients and to therapists. Dave is pretty excited that he is getting all of this work done before he even starts on his dissertation. None of his classmates are this far along.

One day, Dr. Jaynes is in his office as Dave is passing by on the way to his cubical. Dave stops in to talk about the study and happens to mention that he is planning to use some part of the data in his dissertation research. Immediately Dr. Jaynes's eyebrows shoot up, and Dave gets a funny feeling in the pit of his stomach.

"Those data are not available to you for use in your dissertation," says Dr. Jaynes. "They belong to the study and therefore to me as the P.I., and I am going to use them to write a book. It's too bad we didn't think of this at the beginning of the study, so you could have included some of your own questions in the survey forms. You could still do that, you know." With only two months left in the data collection period, Dave knows that he would not get much useful information out of that. He is only really cleaning up stragglers at this point.

In addition to being crestfallen, Dave is also really angry. After all, he has collected all of the data in the enormous survey. He has coded it all, made all the phone calls, and cajoled the therapists into cooperating and giving their time and energy. As he sits down for a cup of coffee with one of his fellow doctoral students, it is all he can do not to threaten to kill Dr. Jaynes for being greedy and for misleading him. Plus, what is he going to do his dissertation on now?

Learning Objectives

1. Students should think about what their "ownership" is of work they undertake as research or graduate assistants.

2. Students should consider the proper role of a faculty mentor regarding the work of student research assistants.

Questions

1. What could Dave have done differently to have saved himself this misery?

2. What should Dr. Jaynes have done to forestall this current unhappy set of events?

3. Who really does own the data? Why do you believe ownership lies with that person?

4. In general, how should researchers working together on a project handle ownership and use of data that are collected by the team?

5

Mentor and Trainee Responsibilities

⚒ INTRODUCTION

Graduate students take part in a number of activities in the course of their education and training, including attending classes, writing papers, conducting research, making presentations, and participating in informal discussions with faculty and other students. In addition, graduate students often pair up with faculty members in mentor–trainee relationships that can be either informal or formal, depending on the specifics of their discipline or department policies. Mentors are more experienced (usually older) faculty members who take younger scholars under their wing, advising and training them. As Margaret King puts it:

> Mentors ease the assimilation of students to a new culture, facilitate the integration that may improve a student's chances of successfully completing degree requirements, and help students navigate the rapids of a sometimes uncertain job market. Mentors often inculcate students in the norms of professional socialization that may not be an explicit part of a program's formal curriculum.[1]

Mentors should not take on that role for personal gain, other than the satisfaction of shaping a young person's career and the intellectual stimulation the relationship provides.

Many academics and other professionals can refer to trusted mentors who influenced their education for the better. Mentoring is an important component of graduate education because it is one of the primary means by which one

generation of scholars and scientists transmits knowledge and experience to the next. Mentoring can thus be a crucial aspect of a graduate student's education that often does not end with graduation, but continues throughout the student's professional career.

The knowledge the mentor passes to the trainee comes largely from his or her own experience. Much of it consists of informal rules, practices, and skills that are part of each organization's culture and thus depend on this informal process for dissemination. In addition, mentoring is an important means by which the ethical standards of conduct in a discipline are communicated. The trainee learns what should and should not be done. Thus, mentoring is a type of socialization through which the beginning scientist or scholar benefits from the experience and wisdom of the senior faculty.

Several important aspects of the mentor–trainee relationship touch on ethical issues. One of these is the importance of establishing the responsibilities that each party owes to the other. This helps avoid misunderstandings down the road. Another stems from the potential problems that can arise from the imbalance of power between the two parties—the student mentee is at a disadvantage in the unequal relationship. A third type of ethical issue is related to the rewards of the relationship. Who gets how much credit from the outcomes of the collaborative research and possibly the financial rewards as well must be considered. In other words, are the rewards distributed fairly or not, and how do the parties determine what is fair?

Celia B. Fisher, Frederick J. Wertz, and Sabrina J. Goodman discuss several ethical aspects of the mentor–trainee relationship.[2] They make the point that faculty who did not experience positive relationships with mentors in their own formative years might not have the skills to perform well as mentors themselves. Moreover, some mentors may lack formal education in ethics, or may even pass along unethical or improper behaviors to their trainees. Such barriers to effective mentoring could be removed by formal training in RCR and mentoring. Additional possible contributors to unsuccessful mentoring include the multisite and multi-investigator collaboration characteristics of much scientific research, the large size and complexity of institutions, the nature of the reward system, demands on faculty time, and the lack of rewards for the mentor. Institutions can improve the mentoring situation by encouraging respect for the mentor–trainee relationship, adopting codes of conduct, creating procedures to investigate allegations of misconduct (with sanctions), offering education opportunities, and monitoring the entire RCR environment for self-improvement.

Our focus in this chapter is on mentoring relationships in which ethical breaches might have occurred. You are encouraged to think about these stories and put yourself in the shoes of the different parties to assess the ethical issues from a variety of perspectives.

Notes

This chapter is based on "Mentor and Trainee Responsibilities," Chapter 7 in Nicholas H. Steneck, *ORI Introduction to the Responsible Conduct of Research,* rev. ed. (Washington, DC: U.S. Government Printing Office, 2007), http://ori.hhs.gov/documents/rcrintro.pdf.

1. Margaret King, *On the Right Track: A Manual for Research Mentors* (Washington, DC: Council of Graduate Schools, 2003), p. v.

2. Celia B. Fisher, Frederick J. Wertz, and Sabrina J. Goodman, "Graduate Training in Responsible Conduct of Social Science Research: The Role of Mentors and Departmental Climate," in *The Handbook of Social Research Ethics,* ed. Donna M. Mertens and Pauline E. Ginsberg (Thousand Oaks, CA: Sage, 2009), 550–564.

⧊ CASE 1. APPROPRIATE RESEARCH ASSISTANT DUTIES

Tom sat nervously outside the Anthropology Department chair's office, waiting for a meeting with the chair and with his mentor, Dr. Smith. He was not especially looking forward to this confrontation even though it had been scheduled because of his complaints to the chair about Dr. Smith. Now he worried that he had not only poisoned his relationship with his mentor but had also endangered his career. How did things get to this point?

Tom started the graduate program in anthropology with great enthusiasm and high hopes of becoming a successful academic researcher. He chose this school with that end in mind and learned a great deal in his classes about conducting successful research. He felt his preparation, however, would not be complete unless he acquired a mentor to show him how to become preeminent in the field. He selected Dr. Smith because she had a national reputation, so Tom thought she would be the ticket to his becoming the success he sought to be.

Tom had not talked to anyone about his plans because he did not want another graduate student to affiliate with Dr. Smith first, and possibly he did not want to hurt the feelings of any of the other faculty. He simply went to Dr. Smith's office one day and asked if she would mentor him. "Show me the tricks to becoming a top-flight researcher," he said. Dr. Smith had just signed a contract for a large and lucrative consulting project at a building site and knew she needed help to complete it on time. She saw Tom as the solution to that problem and said to herself that he would benefit from seeing the consulting activity firsthand. So she agreed.

The meeting with the chair did not go well. Tom repeated his complaints to the chair. "Dr. Smith has me answering her phone when she is in class because her clients on this project often call unexpectedly. She has me going through an endless pile of files from the client company, looking for key pieces of information relevant to the project and coding them into a database for later analysis. I'm not learning what I thought I would be learning: how to be a great researcher." Dr. Smith replied, "I thought Tom understood what I expected from him when we first talked. I feel that he is ungrateful for this wonderful opportunity to see a client problem solved. In addition, I am angry because Tom complained to the department chair without talking to me first. I feel that this is a personal attack on my integrity."

The chair listened to both parties and thought about what to do. He was vexed with this squabble and wanted to make it go away as quickly and with as little trouble as possible. He suggested that Tom and Dr. Smith just work it out for themselves. "I don't want to have to call a department meeting and have the faculty discuss this," he said.

Learning Objectives

1. Students should see that the different parties in the mentor relationship (student, mentor, department) can have different expectations of the procedures and outcomes of it.
2. Students should understand that their reasons for seeking a mentor should be overt and that they should clarify their goals to any potential mentor prior to entering the relationship.

Questions

1. Did Tom make the best choice of mentor at the onset? Why or why not?

2. Is Tom justified in his complaints about Dr. Smith? Why or why not?

3. Is there any merit in Dr. Smith's complains about Tom? Why or why not?

4. What could Tom and Dr. Smith have done differently to avoid the confrontation?

5. How effectively do you think the department chair handled this matter? What would you recommend he do?

6. Should the other members of this department play a role in this matter, or should it be strictly between Tom and Dr. Smith?

\\\\\\ CASE 2. CHOOSING A DISSERTATION CHAIR

Marcie is a second-year Ph.D. student in the educational research department at a public university. She is an excellent student with a promising career in research and publication ahead of her. Currently, she is sitting in the hall, waiting for a meeting with one her professors, Bob Johnson, an older member of the department. While she waits, Marcie is rehearsing what she wants to talk to Bob about. Her problem is this:

Another member of her department, Mike Smith, has been talking with her. Praising her ability and potential, Mike has discussed whom she should choose as her dissertation supervisor and partner for a program of research. Mike told her that she should choose him because his national reputation practically guarantees that she will get a first-class job when she graduates and that the top journals will publish her articles if his name is also on them. Being an ambitious student, Marcie has been considering this carefully. However, her topic of interest does not exactly correspond to Mike's. They are close enough that it would not be unrealistic for him to direct her dissertation, but his work is not what she really wants to do. Other members of the department—Bob Johnson, for example—would be more suitable in this regard, and Marcie wonders how she should make the choice. Specifically, she worries whether she would be able to work well with Mike.

She took Mike's seminar during her first year of course work, and she was struck by several of his offhand comments that betrayed a negative attitude toward women, especially women students. She also thought that he sometimes dismissed the ideas of other scholars without considering them thoroughly, something he also seemed to do when students disagreed with him. In addition, she often had problems scheduling meetings with Mike to discuss her term paper. She has heard older graduate students mention that they worked hard on projects that never came to fruition because Professor Smith did not contribute his share of the work. Now she wants to talk to Bob about how she should proceed. Should she select Mike as her dissertation adviser or choose another member of the department?

Learning Objectives

1. Students should gain an understanding of what is important in choosing a dissertation chair and research mentor.

2. Students should see that forming working relationships with faculty could have more than a purely "intellectual" dimension based on shared interests.

3. Students should become aware that not all faculty members put their students' best interests first.

Questions

1. What should Marcie do? What questions should she ask Bob?

2. Has Mike behaved ethically toward Marcie and the other graduate students, as well as toward his colleagues?

3. What should Bob Johnson tell Marcie?