Chapter 18: Ethics 2,  
Responsible Conduct of Research

In 1998 a medical journal called The Lancet published an article of interest to many psychologists. The researchers claimed to have shown a statistical relationship between receiving the combined measles, mumps, and rubella (MMR) vaccine and the development of autism—suggesting furthermore that the vaccine might even cause autism. One result of this report was that many parents decided not to have their children vaccinated, which of course put them at higher risk for measles, mumps, and rubella. However, follow-up studies by other researchers consistently failed to find a statistical relationship between the MMR vaccine and autism—and it is widely accepted now in the scientific community that there is no relationship. In addition, several more serious problems with the original research were uncovered. Among them were that the lead researcher stood to gain financially from his conclusions because he had patented a competing measles vaccine. He had also used biased methods to select and test his research participants and had used unapproved and medically unnecessary procedures on them. In 2010 The Lancet retracted the article, and the lead researcher’s right to practice medicine was revoked (Burns, 2010).

In the first chapter on research ethics, the focus was how to carry out scientific research in a manner that follows current expectations for best practice. We might summarize the core idea as “be nice to research participants.” Treat them with respect and design research to have value for the world (beneficence) and to make these benefits broadly available (justice). These ideas are then reflected in research processes related to obtaining informed consent and working with the Institutional Review Board as an external monitor of regulatory compliance with best practices.

The area described as Responsible Conduct of Research (RCR) reflects carrying out the scientific process in a fair and ethical manner related to the integrity of research and fairness in assigning credit to the researchers involved in the research process. The capitalization of the term is due to fairly recent changes in training of scientists requiring explicit engagement with these issues, especially at the student level (graduate students and undergraduate researchers). The focus on these specific topics has been driven by funding agencies (NIH, NSF) who seek to improve the reliability and quality of the research process.

[Adding this to Chapter 8]

## CITI Training

While the core issues in ethics are reviewed here, if you have the opportunity to work within a research lab in the future, you will very likely be asked to complete “CITI training” to obtain a certificate that verifies that you completed that process. At the time of writing, the Collaborative Institutional Training Initiative (CITI) is the main source of ethics training for researchers in psychology, health and medical and other research that depends on data collection from human participants. Completing your CITI training means signing up with their web service, completing training and test modules and then connecting that account to your university system. The university IRB then verifies your certification and allowability to be part of formal research processes.

The training provided by CITI is typically presented in modules. The ones you will most likely encounter for psychological research are Good Clinical Practice (GCP) which covers content similar to Chapter 8, Ethics 1. There are also specialized modules in Biomedical Research and Social, Behavioral, Educational (SBE) researchers. Psychology research is most typically requires knowledge of the SBE content, but if the research methodology includes biomedical methods (patients, health or biological assays, imaging, etc.) training may also be necessary in specific ethical issues related to biomedical research. In addition, all CITI training requires knowledge of Responsible Conduct of Research.

## Responsible Conduct of Research

As in Chapter 8, the core ideas can be expressed simply: Don’t Lie, Cheat or Steal. The main topic of RCR is to detail how these “kindergarten ethics” ideas apply to research processes in psychological science. Compliance with RCR principles is aimed to maintain the highest level of integrity in research processes so that the scientific community can rely on and trust the results of our scientific work. We have previously discussed the problem of the Type 1 error, a false claim of an effect among variables in research that turns out to be inaccurate. This can happen due to poor design or unexpected problems with extraneous variables. It can also happen due to an integrity violation where researchers do not follow best RCR practices.

Part of the motivation to increase awareness and training in RCR was due to the acknowledgement of external pressures on scientists engaged with the research processes. Successful science can produce substantial rewards for researchers including employment, promotion and access to research funding. Carrying out an unsuccessful research project is then costly in both time and opportunity loss that could reduce access to these rewards. The main content of this text is to illustrate the methods for carrying out research with the most rigor and care possible, but under external pressures, some researchers have failed to adhere to rigorous methods leading to results that are incorrect and/or retracted.

## RCR Failure: Fraud

An obvious way that integrity can be violated in research practice is the wholesale fabrication of data, which is then presented as if it were properly collected. This is fairly rare, although not unheard of, in scientific work as the majority of modern scientific work is done in teams of researchers. This type of violation is so blatant and obvious that it is unlikely not to be known to other members of a collaborative research team. Purely fraudulent findings are also unlikely to be useful as part of a research program to drive subsequent research, which is often an important part of the general operation of a research group. Being found responsible for research fraud also effectively ends a scientist’s research career, leading to immediate dismissal from the university or institution at which they work and a future bar on any external funding support.

## RCR Failure: Falsification

A more pernicious issue in maintaining rigor in research processes relates to “falsification” of research findings, which covers a range of inappropriate data handling methods that lead to presentation of a false conclusion. The simplest of these is to exclude data collected that does not support the researcher’s hypothesis. If we hypothesize that experimental condition A leads to higher scores on our DV than condition B, we can simply exclude all the low scorers from condition A (or high scores from B) and obtain an apparently statistically reliable result. This “data selection” is an obvious failure of research integrity and is treated in the same manner as wholesale fabrication.

There are subtler ways that aspects of falsification can creep into research. Scoring of subjective ratings of performance might not be done in a completely blind manner. Performance hints or clues could be given to participants in one condition. Bias could be covertly embedded in the task instructions or context. Participant recruiting could embed bias in assignment to conditions if not done properly randomly. These subtler issues are seen as problematic due to being difficult or nearly impossible to detect in the report of a completed research project. They could even be created accidentally by researchers who are simply so focused on research success that they implicitly deviate from best practice. A goal of RCR training for all lab personnel is to make everybody aware of these potential failures to provide checks on both their own work and the work of their collaborators.

The process of publishing research depends on peer-review of the methods, results and conclusions of a research project. Unfortunately, data handling problems are not visible to a peer reviewer, so this process does not effectively protect against RCR problems. In many modern journals, authors are encouraged or even required to publish their dataset in a publicly available location to support their results. However, even this may not protect against data selection if the publicly available data has already had the inconsistent data excluded.

A university or large research institution will typically have an office charged with evaluating scientific processes to assure compliance with best practices in research integrity, an Office of Research Integrity. If a concern is raised about a specific researcher or team, this office is charged with investigating and determining if an integrity violation has occurred. This investigation generally takes the form of an audit of research practices, review of raw data, preliminary analysis and evaluation of as many steps of the core research process as possible. These investigations can be complex and time-consuming as active research processes among a team of collaborators can often be fast-moving and sometimes important decisions are made quickly without immediate realization of how consequential they are.

Various recommendations have been put forth to improve the general process of research to maintain the highest levels of compliance with best practice. Some of these involve slowing the pace of research. For example, “pre-registration” of all research studies by reporting methodology, recruiting and planned analysis in advance of formal data collection. These obviously improve integrity but unfortunately can actually exacerbate the problem of external pressure to produce successful results by putting methodologically rigorous labs at a disadvantage in competition with labs that move faster. Another approach is to document and record as much of the research process as possible so that if a question about integrity is raised later, an audit can verify if a problem occurs. One way to accomplish is to work as if there were cameras recording every aspect of the research process in the laboratory so that everybody possibly biased decision about participant exclusion, assignment to conditions or scoring could be evaluated later.

## RCR Failures: Plagiarism and Research Privacy

One of the challenges with maintaining a robust record of all research practices is that some of the external pressure on scientists comes from competition to obtain an important discovery first. This problem is most evident in research fields like drug discovery where establishing the effectiveness of a new pharmacological agent can produce a patentable discover worth as much as a billion dollars. Large financial rewards for discoveries are exceedingly rare in psychological science but being the first to discover, name or characterize a novel aspect of psychology can be very rewarding in career advancement and scientific fame. As a result, cases do occur where multiple labs are considering very similar hypotheses and essentially racing to complete and publish their research project first. The lab that wins the race will accomplish a high-impact publication and lasting credit for the idea whereas the lab that finishes second will lament having gotten “scooped.”

For research being carried out in a context of this kind of competition, researchers will often work with a high degree of privacy about the research being carried out. Unfortunately, protecting research methods by operating in secrecy does not generally support the ability to provide oversight of those research practices. Thus, the areas with the most pressure to produce are also often the most difficult to verify. This is a difficult problem to solve, and the most common current approach depends entirely on training of all research staff.

This issue is where the concept of “plagiarism” plays a more visible role in ethical research practice. The most common form of the issue of plagiarism familiar to students has to do with copying another’s words or ideas and claiming them as your own. This is rare in scientific publication as the majority of the scientific record is easily available to all. In addition, best practices in science are to thorough review background research and cite the relevant research to support the latest findings. Once findings are published, the problem of claiming credit for another researchers’ idea is rare.

However, before publication of a novel finding, there is a risk of another research group finding out about a novel methodological approach to research and then appropriating this idea for their own without credit. This is an obvious integrity violation that can be difficult to deal with. A particularly famous example historically is the famous work of Francis Crick and James Watson to identify the double-helix structure of the DNA molecule. They had been working on this problem for some time, as had several other large research labs, all of which were working secretly in order to be first to solve the problem. Crick and Watson were also serving as reviews for research grants and in that role saw preliminary data obtained by Rosalind Franklin using x-ray crystallography of DNA that was consistent with the idea of a double helix. That directed their subsequent work to show that DNA was constructed that way, which led to substantial scientific fame and a Nobel prize. It was not until decades later that Franklin’s contribution to this discovery was fully appreciated and properly acknowledged. The ethical issues in this case are quite complex as there is no doubt that Crick and Watson developed completely novel methods and tools to come to their conclusion. When their Nobel prize was awarded, Franklin had passed away due to an unfortunately young case of ovarian cancer and since that award is not given posthumously, she would not likely have been included in any event. There are reports that she was even offered authorship on the original paper but declined (her own findings were published simultaneously). However, it is also clear that use of pre-publication data from another research lab without their knowledge or permission is clearly not following best practice for ethical research.

## RCR: Authorship

Credit for scientific findings is generally reflected in participating in authoring the scientific report. Surprisingly, the rules for who is officially an author on a published report are not completely clear and consistent across all domains of science, or even all subdomains within psychology. The APA provides guidance that researchers who provide “substantial intellectual contribution” to a project should receive authorial credit. However, “substantial intellectual contribution” is not defined. Cases where a key idea is appropriated from another research group and used without credit are clear violations of this policy.

More complicated are questions about a research result that comes out of a lab with a number of staff members, graduate students and/or undergraduate student researchers. Some levels of participation in the project are considered to merit co-authorship, but in other cases, there might merely be a mention in the Acknowledgements section of the published paper. There are no fixed rules for how this is decided, and guidelines vary across individual laboratories and also across different subdomains of research.

Even trickier is the question of who takes on the coveted role of lead or first author of a manuscript reporting the results of a project done by a collaborating group. Again, there are no fixed rules for this, but it is generally up to the PI of the group to set a clear policy that is known to all research personnel in advance and to follow that policy without bias or favoritism. Difficult situations can emerge where there is interpersonal conflict among research team members, especially if a graduate student and supervising faculty member have differing opinions about credit. One goal of improving RCR guidelines is to avoid conflict around this issue and minimize or eliminate conditions where a lead researcher appropriates credit for work done by laboratory members

.

## RCR: Analytic Flexibility

An issue that has gotten increased attention in modern scientific practice related to questions about reproducibility of research is the potential problems associated with a flexible approach to data analysis. In its more extreme forms, this produces the phenomenon known as **p-hacking** where a large number of analytic approaches are tried until one of them produces a result that meets the traditional criterion of p<.05. To understand the math behind this, remember that the p-value associated with an analysis is technically the probability of observing the data pattern under the null hypothesis. That is the same as saying that if there was no effect and all the DV measurements will random numbers drawn from the same distribution, what are the odds it would accidentally look like there was an effect? The standard criterion of .05 is effectively a 1:20 odds measure. So, if we had a true null effect but ran a study 20 times, we might accidentally observe what looks like a real result. If you are familiar with conditional probability math, you’ll know that it’s actually only about 65% like to get a false positive in this case, but the broader point that multiple rounds of data collection or multiple analysis techniques essentially mean that we are not effectively meeting the standard criterion.

There are statistical tools for correctly adjusting complex analysis of data where a large number of **multiple comparisons** are being considered. An example of this is where we might have a large number of measures on a large-scale survey and search for all possible relationships among all the variables. The conjoint probability of a false finding across tests is no longer .05 and tools such as **Bonferroni correction** should be applied to maintain rigor. The integrity concern arises from the fact that it cannot be easily determined from a published finding if the authors attempted a wide range of very different types of analysis but only published the one that produced an effect. A related phenomenon informally called the **file-drawer effect** is where a research group might have attempted a study multiple times and discarded all the data from null findings and selected only an unusual successful version of the experiment to present.

The real problem with issues related to analytic flexibility is that the normal process of experimental research often requires researchers to approach their project with a certain amount of creativity and flexibility. In particularly novel research, it can take several attempts to construct the best operational definition of a tricky construct, which will produce a series of unsuccessful experiments before a successful one. In other cases, the DV will not exhibit the exact effect predicted but another approach appears to show a robust finding (e.g., a recoding of the DV, or a related measure). As a result, researchers working with a high degree of rigor and integrity can still end up questioning their own processes as to whether they have accidentally biased their results through flexible procedures.

The idea of pre-registering experimental design and analysis plans is aimed at avoiding false positive reporting through analytic flexibility. This is highly effective, but has the unfortunate side effect of potentially eliminating report of interesting but unexpected findings that would be identified through more creative methods. Where possible, the best practice is to internally replicate an unexpected finding, that is, publish a report with multiple studies showing a consistent pattern. If flexibility was required for the first finding but it is a true effect, it should replicate in subsequent research. This approach is not easily applied in all areas of psychological science, however. Studies with expensive imaging methods, based on limited patient samples, or unique large-scale population surveys cannot be quickly or easily repeated. In those cases, considered analysis plans are important for the researchers and in addition, interested readers of the reports should likely look for parallel similar findings or subsequent research replications.

## RCR: Conflict of Interest

In standard reviews of RCR-related issues, the potential for problems related to **conflict of interest** is also covered. This is relatively rare in experimental psychology research but is a general concern in areas such as health, education, marketing, and other applied research domains. The core of the idea is simple. If a physician is contracted to carry out part of an efficacy study on a new pharmacological agent that is potentially worth a billion dollars if the drug reliable works, that research is being done under extreme external pressure. If the lead researcher has a financial interest in the company making the drug, great care needs to be taken to avoid any possible implicit bias in research procedures that might undercut the validity of the findings. In these cases, extremely rigorous double-blind procedures are used with active external oversight to guarantee integrity. And unfortunately, there are still cases found later where these processes failed.

Most psychological science research does not assess hypotheses with large financial implications. These studies do sometimes have implications for political policies, educational practices or some health interventions. In particular, this can occur when research funding is provided by a private foundation that seeks to advance a specific idea or agenda. To maintain integrity, researchers are required to disclose all financial support to university oversight and to the journal editors when research is published. This allows identification of possible conflicts of interest where, for example, research funding was contingent on a successful finding, which is not a good circumstance to foster best practices for integrity.

# Tension between Ethics and Science

Now that the basic structure of best practice for ethical research (Chapter 8) and ethical conduct of science (above) has been reviewed, it is useful to note a basic tension between best practices in research ethics and maximizing the internal validity of experimental research. It may already be clear that ethical conflict in psychological research is unavoidable. Because there is little, if any, psychological research that is completely risk-free, there will almost always be a conflict between risks and benefits. Research that is beneficial to one group (e.g., the scientific community) can be harmful to another (e.g., the research participants), creating especially difficult tradeoffs. We have also seen that being completely truthful with research participants can make it difficult or impossible to conduct scientifically valid studies on important questions.

Of course, many ethical conflicts are fairly easy to resolve. Nearly everyone would agree that deceiving research participants and then subjecting them to physical harm would not be justified by filling a small gap in the research literature. But many ethical conflicts are not easy to resolve, and competent and well-meaning researchers can disagree about how to resolve them. Consider, for example, an actual study on “personal space” conducted in a public men’s room (Middlemist, Knowles, & Matter, 1976). The researchers secretly observed their participants to see whether it took them longer to begin urinating when there was another man (a confederate of the researchers) at a nearby urinal. While some critics found this to be an unjustified assault on human dignity (Koocher, 1977), the researchers had carefully considered the ethical conflicts, resolved them as best they could, and concluded that the benefits of the research outweighed the risks (Middlemist, Knowles, & Matter, 1977). For example, they had interviewed some preliminary participants and found that none of them was bothered by the fact that they had been observed.

The point here is that although it may not be possible to eliminate ethical conflict completely, it is possible to deal with it in responsible and constructive ways. In general, this means thoroughly and carefully thinking through the ethical issues that are raised, minimizing the risks, and weighing the risks against the benefits. It also means being able to explain one’s ethical decisions to others, seeking feedback on them, and ultimately taking responsibility for them.

Several of the sources of risks to participants are both very common and very mild. Keeping participants blind to the experimental conditions has an element of deception built into the design. Maintaining a signed consent form and records of compensation (payment) to participants creates a privacy risk that somebody may discover they participated in the research study. The time spent completing the procedure is an inconvenience and even survey completion or simple cognitive tasks can be either cognitively challenging or boring. For these common design elements, providing advance information through the informed consent process and fairly compensating participants for their time addresses these mild issues.

However, it is worth acknowledging areas where the best experimental research practices become impossible due to ethical concerns. A number of these arise from research based on interventions aimed to provide a benefit to the participant as well as advancing science (Chapter 19). A clear example is research aimed to establish the effectiveness of life-saving medical interventions. When research on drugs to treat and cure AIDS were being developed, it quickly became clear that while the best research design is a double-blind randomized clinical trial using a placebo, that meant condemning the control group participants to poor health outcomes or even death. The same is true for treatments of life-threatening cancer.

In these cases, it was decided that control groups would not be included in research designs and the treatment condition, which would be all participants, would be compared to population based outcomes of the disease. This is a weaker research practice because the sampling may well be biased depending on availability of the clinical trial and who enrolls is being compared with a broader population sample. However, it was decided that the ethical issues created for researchers and participants if there was true random assignment to control conditions far exceeded the added scientific value. And it might also be noted that maximally effective research practices are primarily needed to detect subtle effect sizes. If we optimistically hope that these medical treatments are having a large effect on health improvements or preventing death, a technically weaker research approach is still enough to establish efficacy.

The problem of assignment to the control group was also a persistent issue with scientific research throughout the COVID-19 pandemic. The ability to truly establish the quantitative effectiveness of mask weaking on slowing the spread of the virus would have required random assignment of participants to non-mask-wearing conditions, which might have been life threatening. Instead, research in this area depended on non-experimental observations across mask wearing conditions that occurred based on personal choice or local culture. This led to widely varying estimates of mask effectiveness and criticism of the scientists attempting to study this. The problem was compounded by the fact that well-done science is not necessarily a fast process and there were demands for a more rapid answer. In addition, the estimates of mask effectiveness indicated it was not a huge effect at the individual level, although small effects over a large number of people showed robust positive epidemiological effects once that data could be collected.

## HIPAA and PHI

Research related to COVID epidemiology also exposed an aspect of health-based research familiar to scientists who work in medical contexts (e.g., Neuropsychology, Chapter 20) known as HIPAA. HIPAA stands for the Health Insurance Portability and Accountability Act, a law passed to protect research participants from negative consequences associated with potential loss of privacy in health research. The insurance related issue was the fact that health insurance was once provided by companies who could deny coverage based on health status, known as pre-existing conditions. As a consequence, if you were in a clinical trial testing an AIDS treatment, you could potentially lose your health insurance for either having or being at risk for AIDS. HIPAA added oversight to handling of Protected Health Information (PHI) to improve the ethical practice of health-based science which affects a great deal of medical research and some psychological studies.

In the normal course of research, scientists are expected to protect the privacy of participants as much as possible but in standard practice this can not be done perfectly, especially with respect to the fact of participation. Participants might be seen coming to the research laboratory by others. While approved research staff are charged with carrying out research procedures, other people around the lab or research collaborators might become aware of details. Oversight of research practices with respect to informed consent or financial compensation methods may leave information about participation available. If research processes have need for keeping information about PHI, HIPAA applies to the research protocol and requires a series of improved information protection procedures aimed to reduce risk for participants. A full discussion of these methods is beyond the scope of this text but awareness of these special cases highlights more challenging elements of ethical research practice.

A common discussion/misconception related to the COVID pandemic might be useful as illustration. Asking people about their COVID vaccination status was in some cases thought to be a violation of HIPAA due to the need to reveal what would normally be PHI. However, self-disclosing PHI is not covered by HIPAA. It would only be relevant if someone were to access your medical records directly without your consent. Asking somebody about PHI might be rude, but it not technically an ethical violation. It should also be noted that ethical issues like this are never thought to have clear black-and-white answers in all cases. It would not be unreasonable for a decision to be made that prioritizes public health and the spread of disease over individual privacy. This kind of issue is also made complex by technological advance. At the time of this writing, it is not clear how ethical practice related to PHI/HIPAA is applied to online systems that give external proof of vaccine status that appears to derive directly from medical records. While these would seem to be in a gray area, they have been widely accepted and seen as valuable, likely meaning that the broader understanding of how PHI is handled will continue to evolve and regulatory guidelines will continue to be modified and improved.

## Data Sharing

For sensitive data like PHI, researchers need to work carefully within guidelines in order to share data with collaborators. In general, data sharing is done by first **de-identifying data**, which is to remove all information in the data records that would link performance data back to the specific participant in the study. In many cases, this is as easy as coding data by participant id and avoiding the use of name (or email) in data records. If data cannot be effectively de-identified, then a **research sharing agreement** is written and reviewed by the IRB to evaluate any risks associated with possible privacy exposure.

Identification of participants is another area where technological advances have led to changes in oversight procedures. Many years ago, researchers might freely share biological specimens from human participants research with other labs. However, the advent of DNA sequencing means that blood or tissue samples can be analyzed in a way that reveals the original participant and now must be evaluated for privacy risk. It has been suggested that the same aspect may be true of some neuroimaging data, i.e., that brain images might be uniquely identifying eventually even though the tools for this do not yet exist. Machine learning techniques may also somebody be able to recover identity from extensive survey data if enough relatively individual data has been collected. In most cases, the privacy risk is minimal but data sharing should always be handled carefully to maintain compliance with best practice.

## Waiver of consent

In some field research studies, research practice requires data collection without the ability to first provide informed consent. This can be approved by the IRB through a request for a “waiver of consent.” This process is also used in cases of severe privacy risk, such as data collection about criminal or high-risk behavior. If participants are potentially asked about prior behavior that could have immediate legal consequences, carrying out the research requires absolute privacy protection. In this case, the written consent process can be waived to ensure that no method of tracing the data back to the participant exists. This is obviously a highly specialized case and one that it very carefully and extensively reviewed. One aspect of the complexity of this process is that although ethical practice might require protection of all research records, it cannot be guaranteed that a legal process such as a court order might override institutional preferences. Ethical practice guidelines from the IRB, the university or even federal funding agencies do not have legal standing to stop a court order or warrant.

## Mandatory Reporting

A fairly recent example of tension between legal understanding of ethical practice and science is recent decisions requiring some research personnel to be “mandatory reporters” for some kinds of observations. The most common situation for this is in developmental research with young children. If research personnel suspect that child abuse is occurring, they are required to report this to institutional authority for investigation. This appears to be a privacy violation for the child and their family, yet the decision was made that the need to protect children overrides the privacy concern. This policy has additional consequences such as the need for training of research staff in these laboratories to appropriately decide when reporting of suspicion is required.

A related issue arises in studies of mental health among adolescents, even university students. Measures related to clinical depression have to be used carefully in practice as there may be institutional or legal policies in place requiring intervention for adolescents at risk. If answers to a mental health survey indicate a potential for self-harm, it may be necessary to have trained and qualified mental health professionals available for immediate participant support. A counter-intuitive consequence of this policy is that questions related to high levels of risk (e.g., self-harm) are often practically removed from mental health surveys if the research team does not have access to adequate support services. That is, knowing that a participant is at risk and not acting is deemed to be a worse situation than not asking questions that would identify the risk. This is clearly an example where scientific practice and ethics have fallen into a complex gray area without an obvious solution.

## Summary

In most psychological research, effective compliance with best practices for ethical research are straightforward and easy to carry out. Showing respect for persons, ensuring voluntary cooperation, properly protecting privacy and data are usually easy to implement. Specialized procedures and training become critical for working in some select, more challenging subdomains. Awareness of the more complex issues in these domains is sometimes useful for understanding the long lists of questions asked in the process of obtaining approval for research from the IRB. Their processes have to be effective for all research and allow for determination of what the risks levels are without any assumptions about the usual case of minimal risk associated with much psychological science.