



istock/Devrimb

Evaluators must use their own moral compass, in conjunction with the guidance of the profession's principles and standards, to take the most ethical and socially just course of action possible.

Evaluation Ethics and Quality Standards

Making ethical decisions throughout the evaluation process is a relatively easy task when the facts are clear and the choices black-and-white. But it is a totally different story when the evaluation context is clouded by ambiguity, incomplete information, cultural incongruence, biases, multiple points of view and values, conflicting responsibilities, and political pressures. Different stakeholders often prioritize different values and weigh risks and benefits differently. In such situations, which evaluators experience frequently, engaging in ethical behaviors and producing a quality evaluation depends not only on evaluators' methodological skills and experience, but also on their critical insight, integrity, cultural competence, and willingness to take a socially just and defensible course of action in view of the standards and guidelines of the profession and the context in which the evaluation takes place.

Introduction

It is expected that an evaluator's work results in an evaluation that is ethical, culturally responsive, and high quality. In reality, achieving these important outcomes is more daunting than it appears on the surface. Due to the idiosyncratic, political, and applied nature of **social programming**, and subsequently the evaluation of those programs, ethical issues can arise that, if not properly handled, will negatively impact the integrity and quality of the evaluation. While the evaluation profession has written, and publically disclosed, ethical principles and quality standards, these represent only a reference point for guiding the attitudes and behaviors of evaluators. Professional evaluators also need to have a keen awareness and deep understanding of contextual factors that may give rise to the complexity and urgency of dealing with ethical concerns in diverse settings.

Ethics is a branch of philosophy focusing on values relating to human conduct with respect to the "rightness" and "wrongness" of actions. It involves standards for

After reading this chapter and participating in the activities, readers will be able to meet the following learning objectives:

- Explain the origins of research ethics codes and their relevance to evaluation
- Identify sources of ethical thinking in evaluation
- Demonstrate ethical sensitivities particularly in settings where there might be cultural incongruence and cultural conflicts of interest between the evaluator and the evaluation context
- Discuss and distinguish between the *Evaluators' Ethical Guiding Principles* and the *Program Evaluation Standards*
- Use the *Evaluators' Ethical Guiding Principles* and *Program Evaluation Standards* to support ethical decision making and evaluation quality in diverse contexts

responsible conduct prescribed by an external source (e.g., society, businesses, professional organizations) aimed at guiding individual decision making and behaviors. This is in contrast with morals, which stem from within individuals—that is, individuals' own personal boundaries and principles of right and wrong. The following activity provides an opportunity for readers to reflect on and discuss ethics from their own perspective.

Reflect and Discuss

What Ethics Means to You

Most adults have a vision or image in their mind of an ethical business, ethical organization, ethical government, or ethical society. On the individual level, ethics has a specific and oftentimes unique meaning and source. Using the prompts that follow, reflect on and discuss what ethics means to you.

- What does ethics mean to you as a student, parent, spouse/partner, employee, and/or other roles?
- What is the basis of your own ethical decision making and behaviors?

Questions adapted from Patton, M. (2002). *Qualitative research and evaluation methods*, (3rd ed.). Thousand Oaks, CA: SAGE

With today's 24/7 cable news cycle, we are frequently bombarded with reports of unethical behaviors engaged in by businesses, politicians, and other professionals. Reported ethical infractions range from cheating (e.g., testing scandals in schools), stealing, and misuse of funds to inappropriate sexual behaviors, abuse of power, misrepresentations, and lying. Reports of these and other unethical behaviors, ultimately, erode confidence and public trust. As social, economic, and political discord manifests in myriad ways throughout the United States, many politicians, educators, activists, practitioners, and researchers, as well as the general public, have become increasingly attuned to the importance of ethics in navigating various aspects of our daily existence.

If evaluations are to be useful to program administrators, staff, participants, sponsors, and the public, the work must be planned, implemented, and disseminated in an honest, objective, and fair manner. In the American Evaluation Association's (AEA, 2018a) *Evaluator Competencies*, the first competency listed under the professional practice domain is

- 1.1 Acts ethically through evaluation practice that demonstrates integrity and respects people from different cultural backgrounds and indigenous groups.

It is essential that evaluators conduct ethically grounded evaluations in the diverse settings in which they work. **Ethically grounded evaluations** are characterized by ongoing critical thinking, reflection, judgment, and decision making. This is squarely aimed at protecting the rights of stakeholders and building daily ethical routines into evaluation planning, implementation, and reporting. There is no separate stage in an evaluation during which ethical issues must be addressed; instead, they arise throughout the entire evaluation process and, thus, must be dealt with continuously. Ethical evaluations require evaluators with more than knowledge of ethics. They also need

evaluators with sensitivities, such as the ability to recognize the ethical dimensions of a situation, and a commitment to ongoing ethical self-examination throughout their work. Furthermore, ethical issues in evaluation are not limited to highly egregious acts such as falsification or fabrication of data and violations of confidentiality, but they can also involve seemingly small, everyday decisions and behaviors such as what information or which stakeholders to engage or ignore, what data to collect or dismiss, and how and when to report the evaluative information.

This chapter examines evaluation ethics and quality standards that are expected to govern the behavior of evaluators and the outcomes of an evaluation. While ethical issues in evaluation undoubtedly extend beyond the behavior of the evaluator (to include others such as clients, sponsors, and users), this chapter is primarily concerned with the evaluator's ethics in relation to various stakeholders within the evaluation context. The chapter begins with a brief historical perspective on research ethics, including the origins of our present-day approach to research ethics and ethical principles emerging from the Belmont Report, which serves as the foundation for protection of individuals involved in research and evaluation studies. The importance of ethics in evaluation and the discipline's major professional guidelines and principles are discussed. The chapter's content and activities are designed to help evaluators develop a keen awareness of how ethical issues can manifest themselves across all stages of the evaluation process, particularly in settings where there might be cultural incongruence and a cultural conflict of interest between the evaluator and the evaluation context.

A Brief Historical Perspective on Research Ethics

While research and evaluation serve different purposes, historical knowledge of research ethics can be valuable in helping evaluators understand the state of contemporary evaluation ethics and why ethics remains an important issue. **Research ethics** are core professional behaviors and institutional and federal standards by which every researcher is guided to protect the dignity, rights, and welfare of research participants. However, there have not always been explicit ethical codes and principles to guide the behavior of researchers. In the 1900s, no regulations existed regarding the ethical use of human participants in research. The field of research ethics has largely been built upon disastrous and egregious treatment of research participants, particularly when those individuals were poor or from minority or other vulnerable populations. The sections that follow provide a brief history of research ethics. More detailed historical perspectives can be found elsewhere (e.g., Kitchener & Kitchener, 2009; National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1978).

The Nuremberg Code of 1947

The origins of our present-day approach to research ethics can be traced back to the **Nuremberg Code** established in 1947. This code, including a set of 10 research ethics principles, followed the December 1946 criminal proceedings against 23 leading German physicians and administrators for their willing participation in war crimes and crimes

against humanity. Horrifying procedures, such as the breaking and rebreaking of bones to see how many times they could be broken before healing, were conducted for research purposes on thousands of concentration camp prisoners without their informed consent; most of these prisoners either died or were permanently disabled as a result. The first principle of the Nuremberg Code, which was a radical idea at the time, was that voluntary, **active consent** of the human participant is absolutely essential. Two other principles included the rights of people to withdraw from research and to protect themselves. The Nuremberg Code also included a principle articulating that it is the duty of researchers to act in the best interests of those who take part in research for the good of society.

The Tuskegee Syphilis Study of 1932–1972

Despite the Nuremberg Code being given status of an international code for the ethical conduct of research and its substantial influence on international documents such as the Universal Declaration of Human Rights (adopted by the United Nations General Assembly in 1948), for many years after the introduction of these documents, researchers continued with unethical practices. In the United States, one of the most infamous biomedical research abuse cases, the Tuskegee Syphilis Study, involved experiments, over a 40-year period, on low-income African American males in Tuskegee, Alabama. As part of a research project conducted by the U.S. Public Health Service, 600 low-income (mostly illiterate) African American males were recruited, 399 of whom were infected with syphilis and 201 of whom served as a control group not infected with the disease. The researchers advertised for “colored” participants with the slogan “Last Chance for Special Free Treatment” by the Macon County Health Department and government doctors. Researchers never obtained informed consent from the men and did not inform the men with syphilis that they were not being treated but were simply being monitored and left to suffer with syphilis long after a cure (penicillin) became available in 1947. By the end of the study in 1972, only 74 of the test subjects were still alive. It was not until May 16, 1997, that President Bill Clinton issued a formal apology for the Tuskegee Syphilis Study, denoting that it destroyed the trust many African Americans had for medical institutions. The following activity provides readers with an opportunity to reflect on the lasting impact of the Tuskegee Syphilis Study.

Reflect and Discuss

The Tuskegee Timeline

Review the timeline of the Tuskegee Syphilis Study and then reflect on and discuss the questions that follow.

The Study Begins

In 1932, the Public Health Service, working with the Tuskegee Institute, began a study to record the natural history of syphilis in hopes of justifying treatment programs for [B]lacks. It was called the “Tuskegee Study of Untreated Syphilis in the Negro Male.”

The study initially involved 600 Black men—399 with syphilis, 201 who did not have the disease. The study was conducted without the benefit of patients’ informed consent. Researchers told the men they were being treated for “bad blood,” a local term used to describe several ailments, including syphilis, anemia, and

fatigue. In truth, they did not receive the proper treatment needed to cure their illness. In exchange for taking part in the study, the men received free medical exams, free meals, and burial insurance. Although originally projected to last 6 months, the study actually went on for 40 years.

What Went Wrong?

In July 1972, an Associated Press story about the Tuskegee Study caused a public outcry that led the Assistant Secretary for Health and Scientific Affairs to appoint an Ad Hoc Advisory Panel to review the study. The panel had nine members from the fields of medicine, law, religion, labor, education, health administration, and public affairs.

The panel found that the men had agreed freely to be examined and treated. However, there was no evidence that researchers had informed them of the study or its real purpose. In fact, the men had been misled and had not been given all the facts required to provide informed consent.

The men were never given adequate treatment for their disease. Even when penicillin became the drug of choice for syphilis in 1947, researchers did not offer it to the subjects. The advisory panel found nothing to show that subjects were ever given the choice of quitting the study, even when this new, highly effective treatment became widely used.

The Study Ends and Reparation Begins

The advisory panel concluded that the Tuskegee Study was “ethically unjustified”—the knowledge gained was sparse when compared with the risks the study posed for its subjects. In October 1972, the panel advised stopping the study at once. A month later, the Assistant Secretary for Health and Scientific Affairs announced the end of the Tuskegee Study.

In the summer of 1973, a class-action lawsuit was filed on behalf of the study participants and their families. In 1974, a \$10 million out-of-court settlement was reached. As part of the settlement, the U.S. government promised to give lifetime medical benefits and burial services to all living participants. The Tuskegee Health Benefit Program (THBP) was established to provide these services. In 1975, wives, widows and offspring were added to the program. In 1995, the program was expanded to include health as well as medical benefits. The Centers for Disease Control and Prevention was given responsibility for the program, where it remains today in the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. The last study participant died in January 2004. The last widow receiving THBP benefits died in January 2009. There are 11 offspring currently receiving medical and health benefits.

Source: National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. (2020, March 20). *U.S. Public Health Service syphilis study at Tuskegee: The Tuskegee timeline*. Retrieved from <https://www.cdc.gov/tuskegee/timeline.htm>

Now, reflect on and discuss the interplay of ethics and social justice issues raised in the Tuskegee Syphilis Study.

Guiding Questions

- Do you think the Tuskegee Syphilis Study has current and/or lasting societal impact? If so, what?
- Might certain individuals or groups still be suspicious of evaluators and other researchers because of experiments like the Tuskegee Syphilis Study? Elaborate your response.

Courtesy of the Center for Disease Control

The Radiation Studies of 1940–1960

Another example of unethical research practice includes the abuse of human participants during World War II and the early Cold War when U.S. officials studied the effects of radiation through experiments on hospital patients, pregnant women, children with intellectual disabilities, and enlisted military personnel. While officials authorized the wartime experiments to establish health and safety standards for the thousands of workers in atomic bomb plants, few of the participants in the experiments gave informed consent. In fact, most had no knowledge that they were being exposed to radioactive materials. After the war ended, officials justified expanding the study of the effects of radiation on grounds of national security. In the 1990s, following congressional investigations, numerous official reports, scholarly studies, and lawsuits, the government offered apologies and financial compensation to some of the victims of human radiation testing.

The HeLa Story: 1950s and Beyond

Until the *New York Times* best seller *The Immortal Life of Henrietta Lacks* (Skloot, 2010) was published, and subsequently presented as a dramatic television film, Henrietta Lacks was virtually unknown to the general public. At the age of 30, Lacks, a working-class, African American tobacco farmer and mother of five, living near Baltimore, Maryland, was a cancer patient in the “colored ward” of Johns Hopkins Hospital. Lacks died of an unusually aggressive form of cervical cancer in 1951 at 31 years of age. Tissue samples from Lacks were taken during her diagnosis and treatment, and portions were passed along to a researcher without her knowledge or permission, as was common practice at the time (Beskow, 2016). Lacks’s cell line, referred to as HeLa (using the first two letters of her first and last names), is the first immortal human cell line in history. HeLa remains viable today and has been used in laboratories around the world for a vast array of biomedical research, contributing to some of the most important medical advances of all time, including the polio vaccine, chemotherapy, cloning, gene mapping, and in vitro fertilization. Further, HeLa cells were the first human biological materials ever bought and sold, which helped launch a multibillion-dollar industry. In 1971, *Obstetrics and Gynecology* (a scientific journal) named Henrietta Lacks as the HeLa source, and this disclosure was subsequently revealed by other scientific publications, including *Nature* and *Science*, as well as the mainstream press. Reportedly, it was not until 1973, two years after Lacks’s name was published in a scientific journal as the source of HeLa cells, that her family learned about the HeLa cells. This case, while consent was not required, raises serious ethical concerns about privacy and respect for family members.

Beyond Medical Studies and Physical Harm: The Milgram Study of 1963

Biomedical researchers were not alone in engaging in unethical practices. There, too, are historical examples of horrific ethical violations occurring in social and behavioral science research. Such violations often resulted in psychological or social harm to participants including feelings of shame, embarrassment, loss of self-confidence, and depression. One of the most infamous such studies was conducted in 1961 by Stanley Milgram, a psychologist at Yale University, on the conflict between obedience

to authority and personal conscience. Participants were led to believe they were administering real pain through electric shocks to another participant as part of a learning experiment, which was designed to see if ordinary Americans would obey immoral orders as many Germans had done during the Nazi period. Baumrind (1964) noted that participants became distressed and nervous when they thought they were administering severe shocks, but when participants asked for the experiment to be stopped, the researcher in charge insisted that they continue. In the Milgram study, participants sustained no physical harm; however, they suffered shame and embarrassment for having behaved inhumanely toward their fellow human beings. Please reflect upon and discuss the questions presented in the textbox.

Reflect and Discuss

Ethical Considerations and Authority Figures

In contrast to the 1940s and 1950s, in today's society do you think everyday citizens are more aware of ethical behavior related to social and behavioral sciences research—that is, right versus wrong—or do you believe that an authority figure will always be able to sway people's judgment toward unethical behavior? Under what conditions do you believe this is less or more likely to be the case? Provide examples.

The National Research Act of 1974

Due primarily to publicity from the Tuskegee Syphilis Study and after a series of congressional hearings on human subjects research, the National Research Act of 1974 was passed. It authorized federal agencies to develop human research regulations and established the first institutional review boards as a mechanism through which human subjects would be protected. Additionally, the National Research Act created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Decades after the Nuremberg Code of 1947 and the 1964 Declaration of Helsinki that provided guidance to medical doctors conducting research involving human subjects, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1978) released the *Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. This document, known colloquially as the **Belmont Report**, established three basic ethical principles that are the cornerstone for regulations involving human participants:

1. *Respect for persons*, or recognition of the personal dignity and autonomy of individuals, and special protection of those persons with diminished autonomy
2. *Beneficence*, or obligation to protect persons from harm by maximizing anticipated benefits and good outcomes for science, humanity, and the individual and by minimizing or avoiding unnecessary risk, harm, or wrong
3. *Justice*, or ensuring that the benefits and burdens of research are distributed fairly and that research procedures are reasonable, nonexploitative, carefully considered, and fairly administered (pp. 4–6)

The principles identified in the Belmont Report relate both to the participant as an individual and to the participant as a member of social, racial, sexual, or ethnic groups. This means that research participants should not be either favored or disfavored (e.g., more or less likely to be involved in risky research) simply because they are a member of a particular class of people.

The Continuing Importance of Research Ethics

Despite regulations and consciousness-raising regarding ethics in medical, social, and behavioral sciences research, it is an area that still needs attention and monitoring by internal and external bodies. In the article “What Is Ethics in Research and Why Is It Important?” Resnik (2015) cites a number of important reasons for a continued emphasis on ethics in research:

- To promote the goals of research, such as the search for knowledge and truth and the avoidance of error
- To promote essential values for collaborative work, including trust, accountability, mutual respect, and fairness
- To ensure that researchers are held accountable for their actions given policies on conflict of interest, misconduct, and research involving humans or animals
- To build public support for research to the extent that people feel they can trust its quality and integrity
- To support additional important social and moral values, including the principle of doing no harm to others (paras. 7–11)

By introducing evaluators to the history of research ethics and important principles such as “doing no harm,” “beneficence,” and “respect for persons,” they are better positioned to, first, identify potential ethical dilemmas when they emerge and, second, be deliberate about how to address the issue in a manner that is fair and just. Such an understanding can help the evaluator identify **ethical triggers** (Munteanu et al., 2015), or elements that indicate potential challenges during the evaluation, particularly when working with participants who belong to vulnerable or marginalized populations (e.g., persons with disabilities), when dealing with sensitive topics (e.g., abortion services), and when there is the possibility of blurred lines between the evaluator and the end user.

Ethics in Evaluation

Virtually every scientific field or discipline has guidelines that are codified in a set of statements put forth by its relevant professional association. Organizations such as the **American Evaluation Association (AEA)**, National Society of Professional Engineers (NSPE), American Psychological Association (APA), American Sociological Association (ASA), National Association of Social Workers (NASW), and American Nurses Association (ANA) have a publicly disclosed, aspirational set of ethical guidelines for professionals. While the guidelines may differ in content and detail across various organizations,

all represent an aspirational set of principles, values, and beliefs that help define the organization and address four overarching issues: (a) *respect others' rights* to act freely and make their own choices while protecting the rights of those who may be unable to fully protect themselves; (b) *do no harm* including both physical injury and psychological harm such as damage to an individual's reputation, self-esteem, or emotional well-being; (c) *act fairly* by treating individuals equitably and without regard to race, gender, socioeconomic status, or other characteristics; and (d) *help or benefit others* through promoting the common good and interests of individuals and society.

In contrast to an organization's **code of conduct**, which is a directional document focusing on compliance and rules that describe how its members should behave in specific situations (e.g., forbid sexual harassment or racial intimidation), ethical guidelines include broad aspirational values and principles intended to provide an organization's members with a general idea of the types of decisions and behaviors (e.g., treat others with respect) that are acceptable and encouraged by the profession. Individuals must interpret the organization's ethical principles and adapt them in practice. A discipline's ethical statements are used to guide practitioners of that discipline in determining the right course of action in a situation. On the one hand, evaluation shares ethical challenges similar to those of other types of social and behavioral science research. For example, public scrutiny of evaluation, like that occurring in other types of research, has resulted in heightened attention to examining whose views are included or excluded, determining design choices, considering how findings are checked for accuracy, and focusing on how results are reported (Wolf, Turner, & Toms, 2009). On the other hand, evaluation can be challenged by a variety of ethical issues beyond those confronted in other social and behavioral science research. In a presentation on ethical land mines in program evaluation delivered at the 1997 American Educational Research Association (AERA) Annual Meeting in Chicago, Linda Mabry stated that

evaluation is the most ethically challenging of the approaches to research inquiry because it is the most likely to involve hidden agendas, vendettas, and serious professional and personal consequences to individuals. Because of this feature, evaluators need to exercise extraordinary circumspection before engaging in an evaluation study. (quoted in McDavid, Huse, & Hawthorn, 2013, p. 468)

In 1999, Mabry added that of all the methods of social science inquiry, evaluation occurs in the most intensely political milieu where the heaviest assaults to ethics are threatened. Evaluation, she contended, involves risks that are rare in research such as generating final reports that can lead to expansion of a program or to its reconfiguration, shrinkage, or termination. Furthermore, evaluators are "often flattered, coddled, granted selective access, indoctrinated, misinformed, disregarded, challenged, or discredited according to the interests and opportunities of clients, program personnel, or other stakeholders" (Mabry, 1999, p. 200). Barnett and Camfield (2016) point out that the wholesale adoption of research ethics may not provide the most appropriate solution for evaluation. Instead, they propose a different approach to evaluation ethics that addresses stakeholder relationships, helps rebalance the primary focus on the respondent, and focuses on the duties and responsibilities of evaluation to society more broadly.

Research ethics has a predominant focus on the researcher–research participant relationship. In evaluation, however, other equally important relationships must be considered from an ethical perspective. These include relationships with a host of other stakeholders such as program administrators, service providers, program

participants, and community members. Evaluators may experience role conflict with certain stakeholders, such as program staff, funders, or program clients, that raises ethical dilemmas. Happy clients (e.g., program administrators) make for pleasant working conditions for the evaluator, and these stakeholders' favorable comments can generate additional contracts for the evaluator that ultimately enhance both income and professional standing (Mabry, 1999). Thus, client or stakeholder appeasement creates positive bias, which raises not only a validity issue but also an ethics issue. Evaluators who choose to comfort or enrich themselves by providing reports that cheer more than inform may hinder clients' access to the information promised, foreclose on opportunities for program improvement, warp managerial decisions, and yield preventable negative human consequences (Mabry, 1999). The following case study provides an example of a key stakeholder (i.e., the director) attempting to cajole an evaluator in a manner that could present an ethical dilemma depending on the evaluator's action.

Case Study

Identifying Hidden Agendas and Ethical Land Mines

The following description was adapted from one of the Ethical Challenges found in the *American Journal of Evaluation*. Review the case and then respond to the questions at the end of the scenario.

The Case of the Sensitive Survey

The Health Services Center (HSC) at North Central Southeastern State University, a small public institution (popularly known as Where Are U?), has recently engaged you, an external consultant, to conduct an evaluation of its programs and services. Last year, the state legislature mandated that the various administrative units within public colleges and universities be systematically evaluated on a periodic basis, and actually went so far as to allocate funds for this purpose.

The HSC has been included in the "first wave" of offices to be reviewed at Where Are U?, and the HSC director is less than thrilled. He has occupied the director's position for the past 10 years, and he is convinced that (1) several highly placed administrators at the university would like to replace him and (2) these administrators view the evaluation as an opportunity to build a case to support such a move. This afternoon you are meeting with the director to discuss a draft version of a survey that you have prepared, focusing on students' experiences with and opinions of the HSC. One section of the survey contains an extensive list of items that describe both

positive and negative encounters that the diverse student body might have had with the HSC, and asks respondents to indicate which ones they have personally experienced.

In previous evaluations, you have found the survey to be an extremely useful, efficient, and comprehensive method for examining stakeholders' perceptions. The survey will also give students an opportunity to rate how important these various issues and experiences are to them in terms of influencing their overall opinion of the HSC.

It is noteworthy to mention that the survey represents just one of the data-gathering strategies that you intend to use in this project. The director is, to put it mildly, unhappy with certain aspects of the survey. He believes that the items focusing on negative experiences will turn the HSC into a "punching bag" (the director's words) for disgruntled students with an "axe to grind," and that it would only be fair to use such a list if comparable ones were employed in the evaluation of other departments on campus. You, of course, are not involved in the evaluation of these other departments.

The director also maintains that many of the negative survey items pertain to matters that are not fully under the control of the HSC (e.g., the hours when certain medical specialists are available at the HSC). As the discussion continues, you become increasingly convinced that the key issue here is not the methodology

you have proposed for gathering data. Rather, it seems to be the director's intense desire to declare certain domains of the HSC's functioning as "off-limits" in this evaluation. Although you can understand why he might be motivated to take such a position, you also strongly believe that an adequate, professionally respectable evaluation of the HSC cannot omit consideration of the areas that the director wants you to avoid. Interestingly, during the entry/contracting stage of this project, the director had not voiced the "domain concerns" that appear to worry him so greatly now.

Source: American Journal of Evaluation (AJE) Morris, M. (2000). The Case of the Sensitive Survey. American Journal of Evaluation, 21(2), 263 -263. <https://doi.org/10.1177/109821400002100216>

Your conversation with the director is cut short when he is called out of the office due to an emergency.

- What do you see as the major ethical issues that the evaluator is facing in this case?
- What hidden agendas could become ethically problematic for the evaluator?
- What might be some professional consequences for this evaluator if the ethical issues remain unresolved?

A number of evaluation-related circumstances could pose ethical dilemmas for the evaluator. Some examples include (a) being contracted to conduct an evaluation on a program not yet ready to be evaluated; (b) being given insufficient time to complete the evaluation, (c) treatment of people associated with the program around issues such as confidentiality, informed consent, and assignment to program groups; and (d) role conflicts facing evaluators (Posavac & Carey, 2016). Ethical issues can also be the result of inappropriate behaviors such as exerting inappropriate influence on program participants, usurping the role of others (e.g., service providers), or having conflicts of interests (e.g., promoting personal interests over those of the client).

Evaluation ethics concerns the responsibility of evaluators to be competent, honest, and respectful to all individuals and groups that are affected by the evaluation. This includes not only evaluators' responsibilities to participants, program staff, and other beneficiaries, whom they must treat with respect, but also their responsibilities to evaluation sponsors to yield a quality and credible evaluation. Because of the different roles (e.g., consultant/administrator, data collector/researcher, reporter, member of profession, or member of society) evaluators fulfill at any given time while in the field, an evaluator may face a variety of circumstances in which ethical dilemmas arise (Newman & Brown, 1996). As a result, evaluators must strike the right balance between what is in the best interest of their client and society (Wolf et al., 2009). Ethics in evaluation does not exist in a vacuum. No set of guidelines or standards can cover every possible situation that evaluators will encounter in the field; thus, evaluators' ethical decision making and behavior must be nuanced to the context with consideration of things such as the values held by the stakeholders, cultural issues, and potential conflicts of interest within the particular setting.

How evaluators make decisions and exercise their ethical and professional judgment in practice will determine the evaluation ethics in context. While professional ethics are designed to protect against the reoccurrence of major atrocities such as the Tuskegee experiments, they also are needed to guard against less obvious, yet still potentially harmful, effects of evaluation when there might be issues such as inaccurate representation of stakeholders' perceptions and viewpoints, invasion of privacy, deception, and insensitivity to those being studied, especially when they represent minority and other vulnerable populations. Evaluators, like other individuals, are likely to

overestimate how ethical they are, which, in turn, can be a barrier against behaving ethically, especially in ambiguous situations. Ethical behavior is not just making the easy ethical choices between right and wrong, but also working through the more complex decisions that involve hierarchies of values, prioritized according to circumstances, and deciding what is, and is not, valid and credible information (Mabry, 2004).

Ethics in evaluation can be considered at both the individual and professional levels—both affecting evaluation quality. At the individual level, ethics is concerned with the behavior of an individual evaluator. The evaluator is expected to act with integrity throughout the entire evaluation process from conceptualization and design through dissemination and use of findings. Ethical issues can exist in the way the evaluator presents (or misrepresents) him- or herself during planning and recruitment, data collection, and dissemination activities; ethical issues can arise due to how an individual evaluator interacts and shares responsibility with stakeholders.

Evaluators are presented with an array of messy, complicated issues in the field requiring them to think quickly and ethically in a given situation. They may face what Guillemín and Gillam (2004) refer to as a number of “ethically important moments.” These are the difficult, often subtle, and usually unpredictable situations that arise in the practice of doing all kinds of research, including evaluations. During the course of the evaluation, important moments arise without warning, and the evaluator must be prepared to respond in an ethically defensible manner at the time the situation occurs. Sometimes, ethically important moments arise and must be addressed immediately. In such cases, the evaluator must decide what action to take in real time. For example, in a situation where a project administrator offers a new evaluator a gift, the evaluator has to make an immediate decision whether to accept or refuse the gift. On one hand, and in some cultures, a gift conveys a great deal of respect and is a sign that the relationship is valued by the giver. If there is a problem, it may rest with the evaluator who may not trust the program director’s motives. On the other hand, when gift giving is intended to favorably influence the evaluator’s judgment, acceptance of the gift is ethically problematic. In this case, the evaluator must immediately decide, in real time, his or her course of action—accept or decline the project director’s gift.

Ethics in evaluation at the individual level is not only restricted to the behavior of the evaluator. It is also related to the rights, responsibilities, and behaviors of various evaluation stakeholders. For example, program staff, clients, sponsors, and other relevant stakeholders (e.g., users, community members) have an ethical responsibility for acknowledging unique entry and access issues, as well as particular community habits, customs, and values that determine suitable conduct in the evaluation process (Yarbrough, Shulha, Hopson, & Caruthers, 2011).

Over the past three decades, the evaluation profession has undergone major developments that have served to highlight the importance of ethics in the discipline. Probably most notable among these include the development and multiple subsequent revisions of both the AEA’s *Evaluators’ Ethical Guiding Principles* and the Joint Committee on Standards for Educational Evaluation’s *Program Evaluation Standards* (discussed in the next section). Additionally, there have been published books on program evaluation ethics (e.g., Morris, 2008; Newman & Brown, 1996), research on evaluation ethics, and numerous journal articles and book chapters on the topic. Through publication of the Ethical Challenges section of the *American Journal of Evaluation*, the AEA has sought to promote evaluators’ understanding of ethical challenges, enhance evaluators’ sensitivity to the ethical dimensions of their work, and translate the *Evaluators’ Ethical Guiding Principles* and the *Program Evaluation Standards* into everyday practice.

Sources of Ethical Thinking

Most individuals have a set of fundamental beliefs or principles that guide their ethical behavior. However, people differ on the sources they draw upon for ethical decision making. Newman and Brown (1996) point to several sources as guidelines for evaluators' ethical thinking. As outlined in the following activity, these include evaluator intuition, past experience, observations of behavior among colleagues, personal values and beliefs, and ethical rules as presented in ethical codes.

Reflect and Discuss

Evaluator Sources for Ethical Thinking

Consider the five sources of ethical thinking described as follows and the examples provided. Now, provide additional examples of instances when an evaluator should avoid or rely on one or more of these sources while working in the field. Reflect on and discuss your responses with others.

- *Intuition:* An evaluator may have an intuitive feeling that something is wrong when a program director demands a summary of an interview with a particular staff member.
- *Past experience:* Because of a previous bad experience, the evaluator makes sure that the evaluation contract explicitly states the purpose and scope of the evaluation, who has input into the evaluation planning, expected deliverables, and how the results will be used.
- *Observations of or consultation with colleagues:* An evaluator may observe a colleague's use of a particular methodology when conducting focus groups with Indigenous populations that would be most responsive to use in another evaluation. In yet another instance, an evaluator might consult with trusted colleagues about an ethical dilemma in order to get the perspective of other evaluators before taking action.
- *Personal values and beliefs:* Evaluators use their personal values, visions, and beliefs to make a decision about the right course of action to take when mistreatment of clients was observed while visiting the program site for the purpose of meeting with the program staff.
- *Rules and codes:* Evaluators refer stakeholders to guidelines and standards provided by the AEA's *Evaluators' Ethical Guiding Principles* and the Joint Committee on Standards for Educational Evaluation's *Program Evaluation Standards* to justify their behavior.

Source: Adapted and updated from Newman and Brown (1996).

Any one, or a combination, of these sources can be appropriate (or inappropriate) under certain circumstances. For example, evaluators can have “educated intuition,” making them more sensitive and alert to the nuances of potential ethical conflicts if they have thought through conflicts before either in simulations or in real life (Newman & Brown, 1996). However, overreliance on one particular source can become problematic. Evaluators who, for instance, primarily rely on intuition or past experience to guide their ethical thinking might inappropriately generalize what was seemingly ethical behavior in a previous setting to a different context that might render that behavior problematic. Evaluators who exclusively use their own personal values and beliefs to make ethical decisions may be on ethically shaky grounds, particularly when they are

from the dominant group and working in a setting that primarily serves persons of color and other marginalized groups such as people with disabilities and the elderly. For example, sometimes the values, beliefs, and experiences of the **dominant group** are used as the yardstick, point of reference, norm, or standard by which persons from the dominant group judge the values, beliefs, and behaviors of marginalized groups. Since values and beliefs are influenced by individuals' cultural background, the use of a dominant standard by the evaluator can be inappropriate and lead to faulty conclusions. For example, as shown in the following case study, the fact that an instrument has demonstrated validity in a dominant setting does not automatically make it ethical to use that same instrument in another setting.

Case Study

Moving Beyond Past Experience

In the Centers for Disease Control and Prevention's (CDC, 2014, p. 16) *Practical Strategies for Culturally Competent Evaluation*, the following scenario is presented:

An existing validated instrument was piloted as part of an evaluation that assessed risk factors related to heart disease and stroke. Some of the items in the instrument dealt with sensitive issues (e.g., cultural eating practices, cultural perceptions of attractive body images, cultural

views on prescribed medications). Respondents were offended by some of the items, which they viewed as racial stereotypes. The inappropriate items led evaluators to conclude that participants would be reluctant or refuse to complete the evaluation protocol. Consequently, the evaluation team members discussed these issues, which resulted in a revised protocol for culturally appropriate communication and the subsequent revision of the data collection instrument.

Cultural Competence as an Ethical Imperative

Evaluators oftentimes work in settings that represent a very different cultural context than their own. The updated AEA (2018b) *Evaluators' Ethical Guiding Principles* highlight the need to mindfully and proactively attend to diversity, equity issues, and common good as prerequisites for ethical practice. Evaluators have an ethical obligation to be culturally competent and to create an inclusive climate in which everyone invested has an opportunity to fully participate in the evaluation process. In the updated preface to the *Evaluators' Ethical Guiding Principles*, a **culturally competent evaluator** is described as one who

draws upon a wide range of evaluation theories and methods to design and carry out an evaluation that is optimally matched to the context; the evaluator reflects the diverse values and perspectives of key stakeholder groups. (AEA, 2018b, p. 1)

The AEA (2011) has also issued a *Public Statement on Cultural Competence in Evaluation*. This statement challenges evaluators to deepen their self-awareness and sensitivity in terms of their own cultures and those of others and to acquire the necessary skills to bridge the cultural gaps between themselves and those in the evaluation context.

Cultural competence in evaluation is a reflective activity requiring evaluators to achieve and maintain a high degree of self-awareness and self-examination to better understand how their own cultural backgrounds and life experiences can serve as either assets or limitations in the conduct of an evaluation. Specific examples of ethical practice related to cultural competence offered by the AEA in its *Public Statement on Cultural Competence* include the following:

- Use approaches that are appropriate to the context; for example, verbal consent can be used in communities with oral traditions, high levels of concern about privacy, or low levels of literacy.
- Engage issues of culture directly, respectfully, and fairly when collecting data, making interpretations, and forming value judgments.
- Incorporate ways to make findings accessible to all stakeholders, including forms of communication beyond written texts and the use of languages other than English.
- Consider unintended consequences when reporting findings; for example, in some cultural contexts, participants in evaluations who are proud of their accomplishments may want to forgo anonymity and have their names attached to their stories. While this may be appropriate in some instances, in other situations the identification of participants may infringe on the rights of people who have not given informed consent.

Reflect and Discuss

Self-Exploration

- Consider diverse cultural experiences that have led you to further “self” exploration. What did you learn about yourself that you were not consciously aware of prior to this self-exploration?
- After reviewing the AEA’s (2011) *Public Statement on Cultural Competence* (available at www.eval.org/ccstatement) in full, discuss how exploring the “self” can improve one’s work as an evaluator.

In **culturally incongruent** settings, ethical issues can arise relative to the evaluator’s respect (or lack thereof) for local customs, values, and belief systems. Three types of ethics are particularly relevant in the evaluation of programs serving nonmajority and culturally diverse populations. These include procedural ethics, situational ethics, and relational ethics (Ellis, 2007; Guillemin & Gillam, 2004). **Procedural ethics** involve those mandated usually by institutional review boards to ensure that the study’s procedures adequately address the ethical concerns of informed consent, confidentiality, right to privacy, freedom from deception, and protection of participants from harm.

While institutional review boards certainly have tremendous value and are entrusted with ensuring the ethical conduct of research, scholars vigorously pushing for use of antiracist methodologies often cite problems with traditional institutional review boards’ ethics review procedures based in positivism, which separates thought from action and subject from object and assumes that research can and should be value free (Chavez, Duran, Baker, Avila, & Wallerstein, 2003). Further, it has been argued that institutional

review boards generally give emphasis to assessing risks to individuals without paying attention to risk to communities (Minkler, 2004), a condition that has ethical implications for evaluations focusing on marginalized communities. Institutional review boards and the protection of human participants are covered in more detail in Chapter 12.

Situational ethics, or ethics in practice as it is sometimes referred to, include concerns involving the day-to-day unpredictable, often subtle, yet ethically important periods that arise while conducting research and evaluations. Situational ethics often reveal vulnerability, and the evaluator must decide what to do and how far to probe the situation at hand. **Relational ethics**, another type of ethics relevant to the evaluator, situates ethical action explicitly in relationships recognizing and valuing mutual respect, dignity, and the connectedness between the researcher and the researched and between researchers and the communities in which they live and work (Ellis, 2007).

Ethical Dimension of Racial Bias

Race is only one of many cultural constructs. However, it is a powerful one in American society and one that deserves special attention. Race, as a socially constructed phenomenon, continues to differentially shape the allocation of power and distribution of benefits and burdens among groups within the United States. When considering the prominent ills (e.g., poverty, crime, education gaps, health disparities) of the country, race is always a factor in the equation. Over two decades ago, Patton (1999) questioned how the lens of race shapes and affects evaluators' understanding and actions. This question still has relevance today.

Racism can be a complex and destructive force in the evaluation context. Evaluators have an ethical obligation to eliminate, or at least mitigate, racial (and other) biases in their work. In 2018, Thomas, Madison, Rockcliffe, DeLaine, and Lowe called for evaluators to use their power and privilege to advance a more equitable society by "calling out" racial biases, policies, and practices and unmasking power inequities and outcomes of people in programs within the context of their racialized environments. Racial bias, particularly as it relates to African Americans, is due in large part to their unique historical position in the United States, and it can be present in both evaluators and evaluation stakeholders such as funders, policymakers, program staff, and others.

Biased conceptions of race can hinder an evaluator's ability to evaluate culturally and socially different worlds and realities. This, in turn, can prevent an evaluator from rendering ethical, honest, and fair evaluations. As discussed in Chapter 1, implicit bias, or unconscious attitudes and stereotypes, can affect evaluators' understanding, decisions, and actions in a manner that lends itself to questionable ethical conduct. Such biases are activated involuntarily and without an individual's awareness or intentional control. Many people do not deliberately discriminate against others, but when they offer preferential treatment to those who are like them, whom they know, and/or whom they like or feel most comfortable with, the resulting outcome can be unconscious bias and discrimination against those who are different. Individuals typically fail to recognize the harm that implicit favoritism of in-group members causes to members of the social out-groups (Sezer, Gino, & Bazerman, 2015). Evaluators, like other professionals and laypersons, are not immune to implicit biases. In the article "Evaluation and the Framing of Race," House (2017, p. 188) pointed out that

a common error of evaluations is that the conclusions of the study don't fit the data on which they're based. This is more likely when the evaluations involve minorities. We need to weigh the empirical evidence carefully against the conclusions.

Inferences from data to conclusions are particularly susceptible to bias because they're based on more background knowledge than just the study.

Evaluators should be prepared for racial biases (House, 2017). As House and others point out, majority evaluators are often poorly informed and even misinformed about minorities and thus ill equipped to evaluate projects serving communities of color. As a result, various scholars have called for use of evaluation approaches, such as those that are culturally responsive, contextually responsive, culturally competent, and transformative and that emphasize inclusiveness, pluralism, and better understanding of minority cultures on the part of majority evaluators (e.g., Frierson, Hood, Hughes, & Thomas, 2010; Hood, Hopson, & Frierson, 2005, 2015; Hopson, 2009; Mertens, 2009; Samuels & Ryan, 2011; Thomas & Stevens, 2004; Thompson-Robinson, Hopson, & SenGupta, 2004). House (2017) calls for (majority) evaluators to check their own predispositions, arguing that no white person growing up in this country can be entirely free of racial framing. He also stresses that evaluators should check the work of their colleagues for such dispositions. Evaluators can be “critical friends” and look for biases that evaluators might harbor in their beliefs, dispositions, and behavior given that such biases can significantly affect evaluation findings (House, 2017, p. 187). This is indeed an ethical responsibility for evaluators.

The following case study describes an example of how evaluators noticed and highlighted racism inherent in a particular project. Even though Mathison (1999) admits that racism likely did not diminish as a consequence of the evaluation, she contends that at least the evaluation resulted in a range of stakeholders discussing the issue.

Case Study

Raising the Issue of Racism in Evaluation of a Program

Mathison (1999) describes how external evaluators working with a group of doctoral students to evaluate an after-school program for inner-city teens was able to raise the issue of racism. She contends that it was much easier for the external evaluator, in contrast to an internal evaluator, to raise issues of racism that were subtle and insidious within the program.

Mathison's Description of the Case

The program focused on a transition-to-work program and, as such the [inner-city African American] teens were responsible for finding their own internships within the organizations, in this case a museum. This organization provided a vast array of potentially valuable experiences for them [the teens], but for the most part they either floundered about looking for work

Source: Mathison (1999, p. 33).

or worked in menial jobs, such as cleaning the cafeteria. Underlying the program's inability to provide positive, meaningful work experiences for these teens was a deficit model of inner-city, African American teens—one that presumed they weren't very able, by definition had psychological problems, and couldn't really be counted on in things that mattered. The program itself was ethically flawed. As external evaluators we could much more easily raise these issues of racism than could an internal evaluator. We were not seen as having a vested interest (although the staff did feel betrayed by us for raising these issues, and the teens were grateful for being given an opportunity to reveal their perceptions) and were presumed by our expertise to be doing a fair job.

Ethical Sensitivity and Dilemmas

Ethics in evaluation is not simply related to doing the right thing or wrong thing, but is also related to ethical sensitivity. One of the early decisions that an evaluator may be confronted with is when a particular configuration of conditions, circumstances, and available choices should be framed and addressed as an ethical problem (Duggan & Bush, 2014). **Ethical sensitivity** includes an evaluator's ability to recognize and respond to the ethical dimensions of diverse evaluation contexts. Evaluators with ethical sensitivity continuously anticipate and seek to address ethical dimensions in their work rather than be surprised by them or simply ignore them.

In practice, ethical sensitivity in evaluation involves three aspects: the evaluator's awareness of or ability to determine whether a situation related to the evaluation or the **evaluand** involves an ethical issue, the evaluator's ability to identify the particular ethical value(s) underlying the issue under consideration, and the evaluator's awareness of the intensity of the issue.

First, the evaluator must be able to determine whether a situation related to the evaluation or the evaluand involves an ethical issue (i.e., awareness). This is a critical skill since oftentimes ethical issues go unnoticed because of their complexity or because of a particular cultural lens or unconscious bias on the evaluator's part. There is a large body of research arguing that unethical behavior often stems from actions that individuals do not even recognize as unethical (Banaji, Bazerman, & Chugh, 2003; Chugh, Bazerman, & Banaji, 2005; Sezer et al., 2015; Tenbrunsel & Messick, 2004). Individuals have an "illusion of objectivity" by acting against their ethical values without conscious awareness of such behaviors (Bazerman & Tenbrunsel, 2011; Epley, Caruso, & Bazerman, 2006). For example, an evaluator may fail to realize that the perspectives of the most vulnerable program participants are being ignored by the funder or project director. Here, ethical insensitivity may occur due to the evaluator's prejudices and biases toward the relatively powerless stakeholder group and favoritism toward the more powerful stakeholders. Evaluators can maintain what Bazerman and Tenbrunsel (2011) refer to as "ethical blind spots" as a result of seeing the world in a way that obscures one to the fact that a wrong or an unethical action (e.g., ignoring the perspective of the powerless) is occurring. In such instances, evaluators fail to see their own biases and subsequently fail to detect ethical dimensions in certain situations as a result. The activity that follows summarizes three sources of ethical blind spots evaluators might face: implicit bias, temporal distance, and failure to notice others' unethical behavior.

Activity

Ethical "Blind Spots" in Evaluation

Sezer and colleagues (2015) delineated various sources of ethical blind spots. After a review of these sources of ethical blind spots, summarized as follows, organize into small groups. Then discuss them as sources of evaluator blind spots. Give specific examples of how these sources might affect evaluators' sensitivities and course of action in practice.

- *Implicit biases*: These include attitudes or stereotypes that affect individuals' understanding, actions, and decisions in an unconscious manner and can lead those individuals to act against their ethical values.

- *Temporal distance*: Individuals overestimate the extent to which they will behave ethically in the future; therefore, temporal distance from decisions with ethical dimensions can be a source of unintentional ethical behavior.
- *Failure to notice others' unethical behavior*: Certain factors lead people to ignore the unethical behaviors of others including self-serving biases (e.g., if unethical behavior benefits them), outcome bias, the presence of intermediaries, and gradual erosion of ethical behavior; actions that produce negative outcomes are perceived as more unethical than similar actions that produce positive outcomes.

Another aspect that Sezer et al. (2015) discussed was the *slippery slope effect* or the gradual deterioration of ethical behavior. Here, individuals are more likely to justify small ethical indiscretions than major ones; however, over time, as they justify more and more, they can be led to justify even big indiscretions. When faced with abrupt and large dilemmas (rather than those that gradually increase), individuals are less likely to be unethical.

There are strategies that an evaluator can use to overcome ethical blind spots and be able to better identify ethical issues in real time. Among these are **active listening** (i.e., hearing a speaker and avoiding premature judgment, asking questions, reflecting understanding, clarifying information by restating a paraphrased version of the speaker's message, and summarizing the conversation), imagining the perspectives of others, and practicing **cultural humility** and openness to others' points of view.

A second aspect affecting evaluators' ethical sensitivity is their ability to identify the particular ethical value(s) underlying the issue under consideration. After gaining awareness of a potential ethical issue, evaluators need to reflect by asking themselves what ethical values are being compromised. In other words, an evaluator must first recognize there is an event to react to and then define an event as having an ethical dimension. Identify any ethical issues that might be facing the evaluator in the case study that follows.

Case Study

The Compromised Evaluator?

In an effort to become the long-term evaluator of a five-year project, an evaluator presents the project in a *slightly more* favorable light in the Year I Annual Report than in her original draft report submitted to the project administrator.

What ethical violations are present in this case, and how might such violations actually hurt (not help) the project being evaluated?

The third aspect affecting evaluators' ethical sensitivity is their awareness of the intensity of the issue. Evaluators do not treat all ethical issues the same. A decision must be made whether the ethical dimension is significant or important enough for the evaluator to do something about it. Ethical intensity often plays a role in how

an evaluator will act. Intensity pertains not only to the awareness of an ethical issue but also to how important such issues are to various stakeholders. Thomas M. Jones (1991) argues that ethical decision making in organizations is generally a function of six factors, each of which we view as relevant to program evaluation:

- *Magnitude of consequences*: the total harm or benefit that participants or the evaluator can derive from an ethical action
- *Social consensus*: agreement among evaluators or the evaluator, client, and other stakeholders whether a behavior is “good” or “bad”
- *Probability of effect*: the chances that something will happen and result in harm to others
- *Temporal immediacy*: the time between the act and the consequences the act produces
- *Proximity of effect*: the social, psychological, cultural, or physical distance of the decision maker (e.g., evaluator) from the beneficiary or victim (e.g., program stakeholders) of the course of action
- *Concentration of effects*: the number of people affected or how much an action affects the average person (pp. 374–378)

All six factors represent characteristics of the ethical issue itself and are expected to have interactive effects. Jones (1991) theorizes that if any factor increases, it is generally expected that the overall level of intensity will also increase, and vice versa, assuming all remaining components are constant. For example, depending on the extent to which evaluators believe that their unethical behavior will result in an immediate impact (i.e., temporal immediacy), they are more or less likely to engage in such behavior.

Sources of Ethical Dilemmas

There are multiple sources of ethical dilemmas for the evaluator. Mathison (2007) identified three sources of ethical dilemmas in evaluation that are not mutually exclusive. One source includes *ethical issues that arise from doing the evaluation*. These, she argues, are manifested in various ways such as delivering evaluation findings or reports that (a) are laundered to omit negative findings, (b) exaggerate successes and positive findings, (c) are suppressed altogether, (d) are released belatedly so they are no longer relevant, and (e) are prematurely released or leaked to the public. Reducing complex evaluation findings into sound bites can become an ethical issue when they are misleading. For example, based on some positive findings from the evaluation, a program administrator can post flyers or hold press conferences claiming “our treatment works.” This fails to communicate the complexity of the findings that “treatment worked” only for certain groups of participants and not others.

A second source of evaluation dilemmas identified by Mathison (2007) includes *ethical issues that are created by the evaluator*. Relevant examples include the evaluator’s (a) personal or financial interest in the evaluand (i.e., conflicts of interest), (b) lack of knowledge or skill in the evaluation technique or method being used, (c) lack of

cultural competence and sensitivity (such as lack of knowledge and respect for local culture and values), (d) ideological positions or values that can bias the evaluation outcome, and/or (e) propensity to deliver positive evaluations to increase job security. When evaluators make promises that they cannot deliver such as agreeing to totally unrealistic timelines just to secure a contract, another ethical dilemma arises as a result of their behavior. Ultimately, ethical evaluators have a responsibility to be honest, independent, impartial, credible, and accountable for their work and to the public. They also have the ethical responsibility to respect the rights, dignity, and diversity of participants; to do them no harm; and to maintain their dignity and confidentiality.

The third source of ethical dilemmas identified by Mathison (2007) includes *ethical issues that do not arise from the conduct of the evaluator or from doing the evaluation, but instead exist within the context of the evaluand and are discovered when planning or conducting the evaluation*. When such an issue is uncovered by the evaluator, it unexpectedly places the evaluator in an ethical dilemma. Examples include uncovering (prior to any data collection activity) that program administrators are engaging in illegal activities (e.g., theft) and malfeasance (e.g., misappropriation of program funds). Uncovering program activities that are knowingly harmful to clients or to public health or safety (e.g., poor food-handling practices) can also pose an ethical dilemma for the evaluator. Here, evaluators must determine whether there is an ethical imperative to “blow the whistle” on the activity in order to protect the public (Mathison, 2007).

It is noteworthy to mention that evaluators do not always agree whether a particular situation represents an ethical dilemma. Many times, the lines between ethical and unethical ways of responding can be blurred or ambiguous. In such situations, the evaluator must make a judgment call, which, of course, is done through his or her own cultural lens. In ambiguous situations, the course of action taken will be a function of how the evaluator interprets the situation. Morris and Jacobs (2000) had a national sample of evaluators respond to a case vignette in which the evaluator assembles a widely representative advisory group for a project but does not actively involve group members in the evaluation process. Their findings indicate that 39 percent of the sample regarded the evaluator’s failure to involve stakeholders actively in the advisory group as “definitely” or “probably” ethically problematic; 49 percent of the sample believed that the evaluator’s behavior was “definitely not” or “probably not” ethically problematic; and 12 percent were “unsure.” Thus, one evaluator’s ethical dilemma may be viewed by another evaluator as a political problem, a philosophical disagreement, or a methodological concern (Morris, 2008).

Handling Ethical Dilemmas

Evaluators are frequently faced with ethical dilemmas at some or all parts of the evaluation process. An **ethical dilemma** occurs when the evaluators have uncertainty about the proper or right thing to do because there is conflict between two or more valid and morally acceptable options such that making one choice prevents selection of the other. The complexity of an ethical dilemma arises out of a situational conflict or paradox between two possible ethical imperatives, in which obeying one would result in transgressing another. For example, focusing on the common good might result in failing to reveal malfeasance that could result in the elimination of a program and job

loss for community-based staff who were not at fault. Here, there is not a definitive, clear correct response—as may become evident to readers after reviewing the dilemma in the following case study.

Case Study

Revising the Evaluation Report

This case was taken from work reported by Morris and Jacobs (2000). It involves a request for an evaluator to tone down the negatives of a report in order to make the program appear more flattering. The evaluation's sponsor and primary client is a philanthropic foundation that is the major source of funding for the program. Review the scenario and answer the five questions posed at the end of the case.

Scenario

An evaluator has recently shared the draft of a final report with the director of the program being evaluated. After reviewing the draft, the program director asks the evaluator to tone down one section of the report that describes some operational problems within the program. The director believes that the findings in this section, although accurate, are presented in a way that could cause readers to overlook the overall success of the program's implementation.

The evaluator reexamines the draft and concludes that the findings on operational problems have been reported in a fair and balanced fashion. Nevertheless, the evaluator wishes to be responsive to the director's concerns. The evaluator revises the section in question, mainly by deleting a number of harshly worded quotes

concerning operational difficulties that were voiced by interview and survey respondents.

What is the ethical course of action?

Imagine that you are the evaluator referred to in this case. First, identify why this situation poses an ethical dilemma for the evaluator.

Answer the following questions, adapted from Newman and Brown (1996, p. 52), that will assist you (as the evaluator) in deciding how to respond to the potential ethical dilemma being posed in the case:

1. What are the consequences of the evaluator's choice? What would happen, for example, if every evaluator made the same decision?
2. What duties and obligations do evaluators have to themselves, the funder, project stakeholders, and society at large?
3. What would be just or fair in this situation?
4. What would be the caring response or course of action? Is that the ethical response? Justify your position.

There is no perfect solution when faced with ethical dilemmas since these situations require the evaluator to make a decision that requires placing certain ethical values over others. In theory, acting in an ethical manner may seem quite simple—that is, just do the right thing! But, in practice, identifying an issue, making decisions, and acting in ethical ways is not so straightforward for evaluators. In the following activity, readers are provided an opportunity to identify potential ethical issues for an evaluator working in the field.

Reflect and Discuss

You Didn't Hear It From Me!

Identify the potential ethical issues the evaluator faces in the activity that follows.

The Situations

During a confidential interview with a disgruntled (white) female staff member, an evaluator was told of an (alleged) incident of inappropriate sexual behavior by a program administrator (middle-aged white male) toward program clients (mostly poor women of color). The disgruntled staff member, while clearly wanting the conversation to remain confidential, stated to the evaluator, "I'm just saying; but you didn't hear it from me."

In another situation, during several confidential interviews, an evaluator learned that program administrators may have falsified the program's accountability reports. However, no one who made these allegations wanted to go on record.

Questions for Discussion

- Which potential ethical dilemmas exist for the evaluator in the two situations described?
- Should the evaluator do anything? If so, what, and why?

Both of these scenarios require the evaluator to weigh the principles of the common good against promises of confidentiality made to the interviewees. Given the possible hidden agendas and complexities embedded in these cases, the evaluator should examine the situations from multiple perspectives, reflecting on whether the situations represent an ethical, legal, or professional problem or a combination of the three. Unless evaluators are forced to take an immediate course of action, they should pause to seek out different points of view and review and troubleshoot options with a more experienced, knowledgeable, and culturally competent colleague for this particular setting while remaining open-minded and reflective. Evaluators who are not culturally competent and who do not know and respect the unique cultural values operating in the evaluation context might inadvertently use culturally insensitive and incongruent methods that damage, instead of support, the community under study. This raises ethical issues related to the value of doing no harm.

Reviewing case studies, such as those found in *Evaluation Ethics for Best Practice: Cases and Commentaries* (Morris, 2008), is a useful way to help evaluators think about, analyze, and organize their thinking about real-life ethical dilemmas that they may face when conducting evaluations. Reflecting on evaluation cases, and discussing them with others, can better prepare evaluators for effective evaluation practice by developing the understanding, skills, and confidence necessary to confront ethical dilemmas in a thoughtful and coherent manner.

Ethics and Conflicts of Interest

A major concern that has serious ethical ramifications is conflict of interest. A **conflict of interest** refers to a set of conditions in which professional judgment concerning the primary interest (i.e., the evaluation) might be influenced by a secondary competing

interest such as financial gain (Tobin, 2003). Evaluators must always be concerned with actual and potential conflicts of interest and must deal with them openly and honestly so that they do not compromise the evaluation process and results. Conflicts of interest in evaluation are inevitable, and they emerge in and affect all groups of stakeholders, including evaluators (Yarbrough et al., 2011).

A conflict of interest, in particular, occurs when two or more competing or contradictory interests relate to an activity by an individual or an institution. In evaluations, conflicts of interest include situations in which financial or other personal considerations may compromise, or have the appearance of compromising, an evaluator's judgment in conducting the evaluation and/or reporting the findings. When evaluators are caught in conflicts of interest, biases often distort findings (House, 2016). Conflicts of interest are distinct from bias, inasmuch as conflicts of interest occur when evaluators' judgment concerning their primary interest (i.e., the production of valid and useful evaluations) is clearly influenced by some secondary and competing interest. Here, conflicts of interest exist regardless of whether the evaluator's judgment and behavior can be demonstrated to have adversely influenced the evaluation. Instead, the conflict exists simply as a condition of the evaluator having competing interests (Tobin, 2003).

Currently, there are two major conflict of interest areas generally considered in the evaluation field that can negatively impact the ethical integrity of the evaluator and the resulting evaluation. These include financial conflicts of interest (e.g., monetary arrangements with sponsors) and professional conflicts of interest (e.g., personal friendships, professional relationships). In evaluation, conflicts of interest extend beyond simple personal or financial interest; they can also occur when different individuals or groups try to influence when evaluations are commissioned, which purposes and questions are addressed, who can serve as evaluators or evaluation staff, when data are collected, which methods are used, who can provide or later access information, who has primary access to findings, and how findings are interpreted (Yarbrough et al., 2011). Examples of conflicts of interest in evaluation include

- any personal benefit the evaluator (or the evaluator's spouse, child, etc.) might gain in a direct or predictable way from the developments of the program or projects the evaluator is reviewing or asked to review in the future;
- any previous involvement the evaluator has had with the program or projects he or she has been asked to review, such as serving on the advisory board or having an undisclosed relationship with the program administrators or staff; and
- financial interest held by the evaluator (or the evaluator's spouse, child, etc.) that could be affected by his or her evaluation.

In addition to financial and professional conflicts of interest, we propose a third conflict of interest area; that is, **cultural conflict of interest** is about evaluator power, status, knowledge of, and identification with the goals and values of the dominant culture and how these factors become a secondary yet competing interest, with the primary interest of the individuals and communities under consideration. This secondary interest might impact the evaluator's questions, methods, measures, definitions of success, and interpretations. Evaluators, particularly those from the dominant culture, are at risk of experiencing a cultural conflict of interest because they think the world is the way they see it or, if not, it ought to be that way. The evaluators may be well intentioned but still demonstrate cultural arrogance or lack of respect. Here, they need to work to educate themselves about relevant cultures in the evaluation context.

A cultural conflict of interest may be unconscious and thus not recognized as a conflict by the evaluator. Such unrecognized conflicts can result in ignoring important perspectives and rendering some important (yet marginalized) stakeholders and communities as relatively invisible.

Ethical Challenges and Dilemmas Across the Evaluation Process

Scholars have researched and written extensively on ethics and evaluation (e.g., Barnett & Camfield, 2016; House, 2011; Mathison, 1999, 2007; Morris, 2008, 2015; Newman & Brown, 1996). Morris and colleagues (Morris, 2015; Morris & Cohen, 1993) identify common ethical challenges evaluators face at various stages of the evaluation. The following case study provides a list of commonly reported ethical challenges faced by evaluators across various phases of the evaluation process.

Case Study

Ethical Challenges Commonly Reported by Evaluators

| | |
|---|---|
| <p>Entry/Contracting Phase</p> <ul style="list-style-type: none"> • A stakeholder has already concluded what the findings “should be” or plans to use the findings in an ethically questionable fashion (e.g., to support a decision previously made). • A conflict of interest exists. • The type of evaluation to be conducted is not adequately specified or justified. • A stakeholder declares certain research questions “off limits” in the evaluation despite their substantive relevance. • Legitimate stakeholders are omitted from the planning process. • Various stakeholders have conflicting expectations, purposes, or desires for the evaluation. • The evaluator has difficulty identifying key stakeholders. <p>Designing the Evaluation Phase</p> <ul style="list-style-type: none"> • The evaluator fails to gain acceptance of the overall design from all relevant stakeholders. | <ul style="list-style-type: none"> • The evaluator believes evaluation design is fundamentally flawed. • Insufficient time and resources are available to conduct a credible evaluation. <p>Data Collection Phase</p> <ul style="list-style-type: none"> • The rights or dignity of those providing data are compromised in some fashion (e.g., violations of confidentiality, anonymity, informed consent). • The evaluator discovers behavior that is illegal, unethical, or dangerous while conducting the evaluation. • The evaluator discovers staff incompetence. <p>Data Analysis and Interpretation Phase</p> <ul style="list-style-type: none"> • The evaluator fails to distinguish between findings and his or her opinions in data analysis. • Methodological choices highlight some findings while downplaying others of equal or greater importance. |
|---|---|

(Continued)

(Continued)

Communication of Results Phase

- The evaluator is pressured by a stakeholder to misrepresent findings.
- The evaluator is pressured by a stakeholder to violate confidentiality.
- Although not pressured to violate confidentiality, the evaluator is concerned that reporting certain findings could represent such a violation.

Utilization of Results Phase

- Findings are suppressed or ignored by the stakeholder.

- Disputes or uncertainties develop concerning ownership/distribution of the final report, raw data, and so on.
- Findings are used to punish the evaluator or someone else.
- Findings are deliberately modified by a stakeholder prior to release.
- Findings are misinterpreted by a stakeholder.
- Plagiarism/misrepresentation of authorship occurs.
- Information gathered for one purpose is used for another.

Source: Morris, M. (2015). Research on evaluation ethics: reflections and an agenda. In Paul R. Brandon (Ed.), *Research on evaluation*. *New Directions for Evaluation*, 148, 31–42.

Ethical Principles and Standards for Evaluators and Evaluations

One of the best ways evaluators can avoid and resolve ethical dilemmas is to know both what their ethical obligations are and what resources are available to them. In the evaluation field, there are explicit principles and standards for guiding the ethical behavior of evaluators and achieving quality evaluations. This section summarizes and discusses two major sources of guidance for evaluators: the AEA's (2018b) *Evaluators' Ethical Guiding Principles* and the *Program Evaluation Standards* developed by the Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2011). Both of these sources were developed in the United States and represent the longest-standing professional principles and standards in the evaluation profession. Many other regions of the world have also developed their own statements of standards. Notable examples include the Canadian Evaluation Society's *Guidelines for Ethical Conduct*, the African Evaluation Association's *African Evaluation Guidelines*, and the Australasian Evaluation Society's *Code of Ethics* (links to these guidelines are provided in the additional resources section at the end of the chapter). It is imperative that evaluators become intimately familiar with the *Evaluators' Ethical Guiding Principles* and the *Program Evaluation Standards* to be better positioned to understand how they should respond in the evaluation context in order to produce the most ethical and highest-quality evaluations possible.

The Evaluators' Ethical Guiding Principles

The AEA's *Evaluators' Ethical Guiding Principles* (referred to as the *Guiding Principles for Evaluators* until 2018) were first adopted in 1994 and have subsequently undergone multiple revisions. They are intended to proactively guide and inspire the ethical conduct of evaluators at all stages of the evaluation process. These guidelines build, implicitly and explicitly, upon the three principles (respect for people, beneficence, and justice) in the *Belmont Report*. The *Evaluators' Ethical Guiding Principles* stress that

it is the primary responsibility of the evaluator to initiate discussion and clarification of ethical matters with relevant parties to the evaluation. (See Appendix A for a full presentation of the *Evaluators' Ethical Guiding Principles*.)

There are five major *Evaluators' Ethical Guiding Principles*. Each of these ethical principles is accompanied by several directives or subprinciples to amplify the meaning of the overarching five principles and to provide guidance for their application. The five guiding principles, briefly described as follows, do not imply priority among them, but instead, priority will vary by situation and evaluator role.

- *Systematic inquiry*: Evaluators conduct data-based inquiries that are thorough, methodical, and contextually relevant. This principle focuses most directly on methodological decisions made during the evaluation, although it renders no judgments favoring some methodologies over others. There are six subprinciples under systematic inquiry.
- *Competence*: Evaluators provide skilled professional services to stakeholders. The principle of **competence** focuses on issues of the evaluator's education, experience, relevant expertise, cultural competence, and professional development. This guiding principle includes four subprinciples.
- *Integrity*: Evaluators behave with honesty and transparency in order to ensure the integrity of the evaluation. Here, evaluators must cultivate openness and full disclosure with stakeholders throughout the entire evaluation process. There are seven subprinciples under integrity.
- *Respect for people*: Evaluators honor the dignity, well-being, and self-worth of individuals and acknowledge the influence of culture within and across groups. At all times, evaluators must demonstrate respect in terms of their interactions with stakeholders (regarding ethnicity, class, gender, orientation, etc.), including not judging them; not discrediting them; ensuring that their views are faithfully recorded, as appropriate; and giving them due consideration in the evaluation process. This guiding principle includes four subprinciples related to the overarching respect for people principle.
- *Common good and equity*: Evaluators strive to contribute to the common good and advancement of an equitable and just society. Prior to the August 2018 *Evaluators' Ethical Guiding Principles* revision, this principle was labeled "responsibilities for general and public welfare." Because the revised principle places more explicit focus on **common good** and **equity**, it was renamed as such. There are five subprinciples under common good and equity.

The five *Evaluators' Ethical Guiding Principles* are not independent, but instead, they overlap in many ways. For example, being honest and transparent (*integrity principle*) overlaps with honoring the dignity, well-being, and worth of individuals (*respect for people principle*). Conversely, sometimes these principles will conflict, and so evaluators will have to choose among them. When this occurs, evaluators must use their own values and knowledge of the evaluation context to determine the appropriate course of action. The following case study involves the external evaluation of a health program and was developed in 2006–2007 by the AEA Ethics Committee Professional Development Task Force. The first author, Veronica Thomas, was a member of that task force, and this case has been used as part of a training package on the *Evaluators' Ethical Guiding Principles*.

Case Study

Application of the *Evaluators' Ethical Guiding Principles*

Read the following case example, keeping in mind the AEA's (2018b) *Evaluators' Ethical Guiding Principles*. Then, organize into small groups and discuss the case. Complete the worksheet and question at the end of the case.

Evaluation Context. The Health Care Collaborative program grew out of a multiyear effort funded in many sites by a national foundation. That initiative promoted local collaboration among health care providers and residents in poorly served or underserved neighborhoods. The Health Care Collaborative office uses trained residents as outreach health workers to raise health-issues awareness among residents and to give them options for accessing health care. Health care providers who are collaboration partners deliver a range of services to neighborhood residents. A local funding source supports the Health Care Collaborative, which has a program director, administrative staff, and a small network of outreach workers. The Health Care Collaborative Board of Directors consists of a small group of health care providers.

The Health Care Collaborative serves an economically challenged neighborhood in a small metropolitan area: Average income is one-third to one-half of its metro and national counterparts. The neighborhood is quite diverse along many dimensions, including age, household composition, sexual identity, education, religious preference, race, and ethnicity. The neighborhood has a large African American population, an increasing population of refugees from African and Eastern European nations during the past 20 years, and a rapidly growing Hispanic population in recent years.

Entry, Contracting, and Design. The Health Care Collaborative Board and local funders found that they needed more information than the program's reporting system alone could provide about how program participants viewed the Health Care Collaborative, how the staff viewed the program and the neighborhood, and how the program met or did not meet identified service needs. The funder provided \$20,000 for this purpose, and the Board established a one-year schedule for completing an evaluation. The funder and the

program director approached a local faculty member, an evaluator who also teaches evaluation, to ask for a proposal. The faculty member has previously served on the Health Care Collaborative Board. Discussions with the funder, the program director, and some members of the Board identified key expectations and constraints.

The faculty evaluator proposed a multimethod approach for a formative evaluation. The design included surveys of participants (brief), program staff, and other health care provider partners. The surveys would include questions about racial and ethnic identity. Selected program participants would be asked to keep journals and to participate either in a focus group or in an observed service delivery for a small group. Three focus groups were proposed: one for senior citizens; another for adult, nonsenior males; and a third for adult, nonsenior females. The Health Care Collaborative focus group participants would be offered a \$25 gift card for their time. The institutional review board's approvals would be obtained for informed consent to voluntarily participate in the evaluation.

A graduate student would do most of the data collection, under the evaluator's supervision. The student was fluent in Spanish and English, and this project would be the subject of the student's master's thesis. The evaluation's final product would be a presentation of results, in PowerPoint format, with the slides and notes delivered to the program director and funder.

Data Collection. The student administered the staff surveys in person. These surveys asked for how long the staff members worked with the Health Care Collaborative, what they did in the program, how they viewed the participants, and what difference the program made in the neighborhood. Surveys of other providers involved with the Health Care Collaborative were web-based. The questions concerned what kinds of interaction the providers had at the Health Care Collaborative, with whom, and how often; how that relationship affected both organizations; and what services the responder brought to resident-participants in the Health Care Collaborative.

The Health Care Collaborative staff administered surveys to program participants during ongoing program contact. The student also conducted a small number of interviews of people identified for their longevity in working with this particular neighborhood, and added open-ended historical questions.

The student observed both staff and participants in health care awareness sessions for small groups to better enrich the evaluator's and student's understanding of the program, its staff, and the participants. Participants' journals provided inspirational stories of their experiences in navigating the health care maze.

Data Analysis and Interpretation. From the surveys, some data were aggregated and reported descriptively (e.g., comparisons of the racial and ethnic composition of the Health Care Collaborative participants for the neighborhood). Scaling and cluster analyses were used to structure and analyze the results of the focus groups, and some journal entries and responses to open-ended questions from interviews also were analyzed.

All in all, the program served a disproportionate number of Hispanic adults (compared to the neighborhood's composition) and disproportionate numbers of people without health insurance and without other known ways to access health care. Participants and staff were very positive about the program and its value in their neighborhood and lives. The Health Care Collaborative program participants overwhelmingly credited the use of racially and ethnically diverse staff, from the neighborhood itself, as the main reason for the Health Care Collaborative's success.

Younger adults placed more concern on financial issues related to health care, compared with older adults. Hispanic participants in focus groups were all female, and most were unemployed. From all three focus groups, whether participants were treated fairly and had access to insurance and to health care was more important than waiting times or actually getting to appointments.

When the evaluator and student felt comfortable with their work, they shared draft findings informally with the program director, funder, and Board members—through in-person as well as telephone

conversations and through email. Some feedback was given and considered in reviewing those findings and in developing the final product.

Dissemination and Utilization of Results. The final evaluation briefing was delivered at a meeting of the Health Care Collaborative Board, to which the funder and some residents were invited. The funder could not make this meeting, accepted the electronic PowerPoint file, and asked no further questions. Only one resident—a regular attendee of Board meetings—was present for the briefing. Two or three questions were asked, more of apparent curiosity than any other cause or purpose. No future plans for the findings were discussed at this meeting.

The student completed the thesis based on this project, and it was very well received by the faculty committee. The evaluator adapted the evaluation for use in an advanced evaluation course for graduate students.

The student and evaluator also proposed a poster session focusing on the evaluation findings to an annual, national professional conference in their discipline. The proposal was accepted and a large poster developed, which covered the basics of the evaluation. Those who stopped to read and talk about the evaluation expressed admiration for its scope and methods.

As the evaluator, what are some things that you would do differently to better ensure that your actions are ethically defensible?

Case Study Worksheet to Be Completed

| Guiding Principle | Issues or Questions Raised Related to the Principle |
|---------------------------|---|
| Systematic Inquiry | |
| Competence | |
| Integrity | |
| Respect for People | |
| Common Good | |

Source: This case is republished with permission of the American Evaluation Association (with minor edits).

The Program Evaluation Standards

In addition to the *Evaluators' Ethical Guiding Principles*, the **Program Evaluation Standards** is another document that provides guidance and direction for those in the evaluation field. It includes much more specificity regarding what to do and not do in program evaluation than the *Evaluators' Ethical Guiding Principles*. Whereas the *Evaluators' Ethical Guiding Principles* are concerned specifically with the ethical conduct of the evaluator, the *Program Evaluation Standards* pertain to the quality of the evaluation. Initially established in 1981 by the Joint Committee on Standards for Educational Evaluation¹ with multiple editions since then, the *Program Evaluation Standards* provide guidance for improving evaluation quality and accountability. The *Program Evaluation Standards* contain 30 standards organized around five central attributes of evaluation quality. These quality attributes include (a) utility ($N = 8$ standards), (b) feasibility ($N = 4$ standards), (c) propriety ($N = 7$ standards), (d) accuracy ($N = 8$ standards), and (e) evaluation accountability ($N = 3$ standards). A full description of the 30 *Program Evaluation Standards* is provided in Appendix B. An overview of the five central attributes discussed in the *Program Evaluation Standards*, as adapted from Yarbrough et al. (2011), include the following:

- **Utility** standards are concerned with evaluation use, usefulness, influence, and misuse. Utility is supported by standards that provide guidance to increase the likelihood that the evaluation will have positive consequences and substantial influences such as contributing to stakeholders' learning, informing decisions, leading to improvements, or providing information for accountability judgments.
- **Feasibility** standards are intended to increase evaluation effectiveness and efficiency by ensuring that an evaluation is practical, efficient, and contextually viable. These standards highlight the logistical and administrative requirements of evaluations that must be managed, bring the world of possible evaluation procedures into the world of practical procedures for a specific evaluation, and serve as a precondition for other attributes of quality.
- **Propriety** standards support what is proper, fair, legal, right, and just in evaluations. These standards cover three overlapping domains: (a) the evaluators' and participants' ethical rights, responsibilities, and duties; (b) systems of laws, regulations, and rules that regulate the conduct of people and organizations, such as federal, state, local, and tribal regulations and requirements, institutional review boards, and local/tribal constituencies that authorize consent to work in and with respective communities; and (c) the roles and duties inherent in evaluation professional practice.

¹The Joint Committee on Standards for Educational Evaluation (JCSEE) is supported by 17 sponsoring organizations and has been a member of the American National Standards Institute (ANSI) since 1989. During its history, the mission of the JCSEE has remained to develop and implement inclusive processes producing widely used evaluation standards that serve educational and social improvement. To learn more about the history and organizational support of the JCSEE, visit www.jcsee.org.

- **Accuracy** standards seek to increase quality in data collection and analyses and to increase the truthfulness and dependability of evaluation representations, propositions, and findings by urging that evaluations strive for as much accuracy (i.e., validity, reliability, reduction in error and bias) as is feasible, proper, and useful to support sound conclusions and decisions in specific situations. Ignoring nondominant cultural perspectives and assuming that certain methodologies (e.g., experimental designs) are the only factor necessary for justified conclusions and decisions is a barrier to adherence to the accuracy standards.
- Evaluation **accountability** standards encourage adequate documentation of evaluations and a **metaevaluation** (evaluation of the evaluation) focuses on improvement and accountability for evaluation processes and products. Attention to accountability guides improvement during all phases of the evaluation, and it encourages reflection and a metaevaluative perspective in evaluators and evaluation users.

In his Voices From the Field interview, Michael Morris stresses that evaluators must uphold the *Evaluators' Ethical Guiding Principles* and the *Program Evaluation Standards* and resist pressure to act unethically. This sometimes takes, as he points out, consideration of potential ethical challenges during evaluation planning, moral courage, and just the willingness to do the right thing.

Voices From the Field

Michael Morris: Ethical Considerations in Evaluation

Evaluators must act with integrity and see themselves as more than just methodological technicians as they uphold the *Evaluators' Ethical Guiding Principles* and the *Program Evaluation Standards*. They should strive to understand the organizational and other cultures in which a project is embedded, because they cannot do justice to the evaluation without such an appreciation. Before the evaluation is designed and implemented, evaluators should consider the ethical challenges that might arise and find a way to introduce these topics into discussions with stakeholders during the contracting and negotiation phase, in addition to soliciting the stakeholders' concerns. Having mildly uncomfortable conversations with stakeholders early on can reduce the likelihood of having to engage in much more difficult interactions later in the project. Doing this will also enable the evaluator, at a later point, to bring stakeholders' attention back to those initial discussions,

increasing the chances that the latter will act in accordance with whatever guidelines had been agreed upon. For example, pressure to misrepresent or ignore (unflattering) findings is frequently encountered by evaluators. Early discussion of how to deal with potentially unwelcome results in the evaluation report is a worthy investment of everyone's time. Ultimately, moral courage is key for evaluators, particularly internal ones. Doing the right thing can put an evaluator at risk. Sometimes, however, the only reason for doing the right thing is that it is the right thing to do.

Michael Morris is emeritus professor of psychology at the University of New Haven and a former chair of the AEA Ethics Committee. He is the author of Evaluation Ethics for Best Practice: Cases and Commentaries (Guilford Press, 2008). Veronica Thomas interviewed Dr. Morris in the fall of 2019.

Evaluation Corruptibility and Fallacies

It is often said by numerous scholars and practitioners in the evaluation community that “evaluators must be able to speak truth to power.” Evaluation corruptibility and evaluation fallacies are two factors that can put an evaluator at risk of unethical decision making, jeopardizing evaluation quality, and, thus, an inability to “speak truth to power.” Fitzpatrick, Sanders, and Worthen (2004) use the term *evaluation corruptibility* to describe ways that evaluators may be convinced to go against ethical standards, thus engaging in ethical compromises or distortions. They point to five specific areas of evaluation corruptibility (Fitzpatrick et al., 2004, pp. 423–424):

- *Conflict of interest*: a willingness to twist the truth and produce positive findings due to conflict of interest or other perceived payoffs or penalties (such willingness may be conscious or unconscious)
- *Unsubstantiated opinions*: an intrusion of unsubstantiated opinions because of sloppy, capricious, and unprofessional evaluation practices
- *Prejudices and biases*: “shaded” evaluation “findings” as a result of intrusion of the evaluator’s personal prejudices or preconceived notions
- *Inducements*: obtaining the cooperation of clients or participants by making promises that cannot be kept
- *Not honoring commitments*: failing to honor commitments that could have been honored

To avoid corruptibility, evaluators must be transparent and disclose any relationships (e.g., previous organizational ties or ties with program staff) that might predispose them to bias or give the appearance of bias. Further, they should in no way profit from the outcome of an evaluation. Familiarity with and adoption of the *Evaluators’ Ethical Guiding Principles* and the *Program Evaluation Standards* can provide much-needed guidance for dealing openly with situations that can impact ethical decision making and quality evaluations.

House (1995) considered the issue of evaluator corruptibility from a different perspective than Fitzpatrick et al. (2004). He suggested that evaluators can have the best intentions and may not be corrupt, per se, but, at times, may have a misunderstanding about their responsibilities. House referred to these misunderstandings as *evaluation fallacies*. A fallacy is a mistaken belief based on unsound argument deriving from reasoning that is logically inaccurate. House (1995, pp. 29–30) identified five evaluation fallacies that can have negative ethical consequences:

- *Clientism*: the fallacy that doing whatever the client requests or whatever will benefit the client is ethically correct
- *Contractualism*: the fallacy that the evaluator must follow the written contract without question, even if doing so is detrimental to the public good
- *Methodologicalism*: the belief that following acceptable inquiry methods ensures that the behavior of the evaluator will be ethical, even when some methodologies may actually compound the evaluator’s ethical dilemmas

- *Relativism*: the fallacy that opinion data the evaluator collects from various participants must be given equal weight, as if there is no basis of appropriately giving less priority to the opinions of peripheral groups than to those of more pivotal groups
- *Pluralism/elitism*: the fallacy of allowing powerful voices to be given higher priority because the evaluator feels they hold more prestige and potency than the powerless or voiceless

Evaluator Role, Power, Politics, and Ethics

Ethical issues can arise centering on the evaluator roles, power imbalances between the evaluator and key stakeholders, and evaluator privilege. Politics can also have ethical dimensions that impact an evaluator's work. As discussed throughout this book, power and privilege are concepts that extend far beyond an individual evaluator or a particular evaluation. Frequently, relationships between the evaluator and stakeholders and between/among stakeholders are enthralled in power imbalances and hierarchical struggles. Hierarchical arrangements and power imbalances in the evaluation context exist long before the evaluator is on the scene since oppressive systems often shape the conception, design, and implementation of the program that the evaluator is tasked with studying. Evaluators are often asked to assess the effectiveness of social programs that are designed to yield a quick “magic bullet” fix to problems (e.g., racial achievement gaps, poverty) derived from years of racial and other oppressions (Thomas et al., 2018). In order to accomplish this, evaluators must develop a critical consciousness of how institutional, historical, and systemic forces limit and promote the life opportunities for particular groups. Instead of, for example, identifying delinquency, substance abuse, and violence as problems, evaluators should emphasize the root causes by examining the larger political, economic, and social forces that create persistent poverty, thus jeopardizing healthy development (Thomas et al., 2018).

In any given evaluation, the evaluator occupies multiple roles, including those of expert, knower, judge, and educator. For example, an evaluator can be an expert or program facilitator during the program implementation, a researcher when collecting and analyzing evaluation data, a judge during the reporting phase when making an assessment of program merit and worth, and an educator or advisor throughout the entire evaluation process. The roles that evaluators assume are generally all positions of tremendous power with opportunities to exercise that power in either ethically “just” or “oppressive” ways. Power in evaluation is more distributed toward the evaluator since it is, in fact, the evaluator who is studying others and generating knowledge (and not vice versa). While evaluators do not generally own the knowledge generated from their evaluations, it is still the case that their perspectives and interpretations are often privileged over those being studied in the evaluation context.

Interplay of Politics and Ethics

It is also worth distinguishing ethical considerations from political issues, although they are oftentimes closely intertwined. Ethical considerations relate to issues of right and wrong, good and bad, whereas the central focus of political issues relates to power and control. Political issues can undermine the integrity of an evaluation and certainly

have ethical ramifications by silencing voices and perspectives of the less powerful and rendering these individuals invisible. For example, politics is likely operating when an evaluator is only allowed to evaluate what project administrators or funders believe to be model or successful sites while more troublesome sites are hidden or excluded from consideration. This is a power play that has definite ethical implications related to excluding certain perspectives from consideration.

Power plays, or attempts to gain an advantage by using certain tactics to magnify one's influence or power, can be exhibited by the evaluator, as well as by various stakeholders. For example, power plays by those being evaluated (e.g., program staff) include denying the need for an evaluation, claiming the evaluation will take too much time away from their normal workload, and/or intentionally providing the evaluator with huge amounts of information so it is difficult to sort out what is relevant and what is not (International Program for Development Evaluation Training [IPDET], 2009). Power plays by the evaluator might include using the "experts know best" line, applying unstated criteria to decision making, and/or applying unstated values and ideological filters to the data interpretation (IPDET, 2009). Other stakeholders, such as community members, can also engage in power plays with ethical ramifications (see the following case studies for additional examples).

Case Studies of Political Power Plays in Evaluation With Ethical Ramifications

Political Power Plays Engaged in by Evaluatees

- Denying the need for the evaluation
- Claiming the evaluation will take too much time away from their normal workload
- Claiming the evaluation is a good thing, but introducing delaying tactics
- Providing the evaluator with huge amounts of information so it is difficult to sort out what is relevant and what is not
- Omitting or distorting information they are asked to provide so they do not look bad
- Coming up with new data at the end
- Arguing that the evaluation findings are irrelevant because things have changed

Political Power Plays Engaged in by Evaluators

- Using the "experts know best" line to exclude the perspectives of others

Source: Adapted from International Program Development Evaluation Training (2009).

- Insisting evaluations should only be quantitative in nature since statistics do not lie
- Not stating or shifting the measurement standards
- Applying unstated criteria to decision making
- Applying unstated values and ideological filters to the data interpretation
- Ignoring certain evaluation findings

Political Power Plays Engaged in by Other Stakeholders

- Giving their own conclusions to meet their own agenda
- Trying to get the media (or powerful others) to criticize (or praise) the organization being evaluated in order to sway opinion

SUMMARY

A critical task for evaluators in any evaluation is to identify issues, including those of an ethical nature, that might hamper the conduct of a fair, honest, and accurate evaluation. This chapter examined evaluation ethics and the quality standards that are expected to govern the behavior of evaluators and the outcomes of an evaluation. Evaluators must take necessary steps to equip themselves with the knowledge, skills, and dispositions to accomplish this goal. This means having the sensitivity to identify and deal with the ethical challenges in the evaluation context. This chapter highlighted some common ethical challenges and offered possible solutions. Special consideration was given to how conflicts of interest, cultural issues, racial bias, and political issues impact evaluation ethics. The origin of research ethics, why they are important, and the three ethical principles from the Belmont Report were discussed to provide readers with a foundation for better understanding current evaluation ethics. The AEA's *Evaluators' Ethical Guiding Principles* and the Joint Committee on Standards for Educational Evaluation's *Program Evaluation Standards*, although unable to cover every possible scenario that an evaluator might face, were discussed to provide a framework that gives guidance to evaluators.

In conclusion, the following is a set of reflective questions, adapted from Patton (2003, pp. 408–409), that evaluators can ask themselves to help them think through some ethical issues that might arise during their work.

- How will the evaluation contribute to society, the community, and/or the world?
- Why should individuals participate in your project? What are the benefits to them?
- How will you explain the purpose of the inquiry and methods to be used in ways that are accurate and understandable to those you are researching?

- In what ways, if any, will conducting this research or program evaluation put people at risk? (Consider psychological, legal, and political issues and the possibility of people becoming ostracized by others.)
- If you uncover controversial information, how should it be shared?
- What are reasonable promises of confidentiality that can be fully honored?
- What information can you *not* promise to keep confidential?
- What kind of informed consent, if any, is necessary for mutual protection?
- Who will have access to the data, and why?
- How will you and your respondent(s) likely be affected by conducting this research or program evaluation?
- Who will be the researcher or evaluator's go-to person(s) during the study regarding ethical issues that might arise?
- How hard will you press participants for data? Where will you draw the line?
- What ethical framework and philosophy informs your work and ensures respect and sensitivity for those you study, beyond whatever may be required by law?

In the final analysis, evaluators must use their own moral compass, in conjunction with the guidance of the profession's principles and standards, to take the most ethical and socially just course of action possible.

SUPPLEMENTAL RESOURCES

Practical Strategies for Culturally Competent Evaluation

www.cdc.gov/dhdsp/docs/cultural_competence_guide.pdf

Provided on the website of the Centers for Disease Control and Prevention (CDC), this document includes a crosswalk table in Appendix A, listing each of the *Program Evaluation Standards* in column 1 with

suggested strategies an evaluator can engage in to increase cultural competence relative to that standard in column 2. It also includes other appendices of resources and tools and tips for integrating cultural competence into evaluation.

The Belmont Report

www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm

Part of the U.S. Department of Health and Human Services Office for Human Research Protections, this website provides a link to the full *Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*.

“Human Subjects”

www.nsf.gov/bfa/dias/policy/human.jsp

This website of the National Science Foundation (NSF) has information concerning the basic principles of protection of human subjects as well as information about institutional review boards.

Protection of Human Subjects in Research

www.ed.gov/about/offices/list/ocfo/humansub.html

This U.S. Department of Education web page includes links to general information concerning human subjects in research and the regulations/legalities surrounding using human subjects in research.

Human Subjects Research (HSR)—CITI Program

<https://about.citiprogram.org/en/series/human-subjects-research-hsr/>

Human Subjects Research (HSR) basic content is organized into two courses: Biomedical (Biomed)

and Social-Behavioral-Educational (SBE). They are intended for anyone involved in research studies with human participants, or who have responsibilities for setting policies and procedures with respect to such research, including institutional review boards (IRBs). Additional modules of interest within HSR allow for exploration of several important topics and may be selected to meet organizational needs. HSR includes additional stand-alone courses for institutional/signatory officials, IRB chairs, and public health researchers, as well as a revised Common Rule course that covers the regulatory updates to the Common Rule. These courses were written and peer-reviewed by experts.

Web Links to Ethical Principles and Quality Standards

AEAs Evaluators’ Ethical Guiding Principles

www.eval.org

Joint Committee on Standards for Educational Evaluation’s *Program Evaluation Standards*

<https://jcsee.org/program/>

Canadian Evaluation Society’s *Guidelines for Ethical Conduct*

www.evaluationontario.ca/membership/standards-guidelines/

African Evaluation Association’s *African Evaluation Guidelines*

<https://afrea.org/the-african-evaluation-guidelines/>

Australasian Evaluation Society’s *Code of Ethics*

www.aes.asn.au/images/stories/files/membership/AES_Code_of_Ethics_web.pdf

Do not copy, post, or distribute