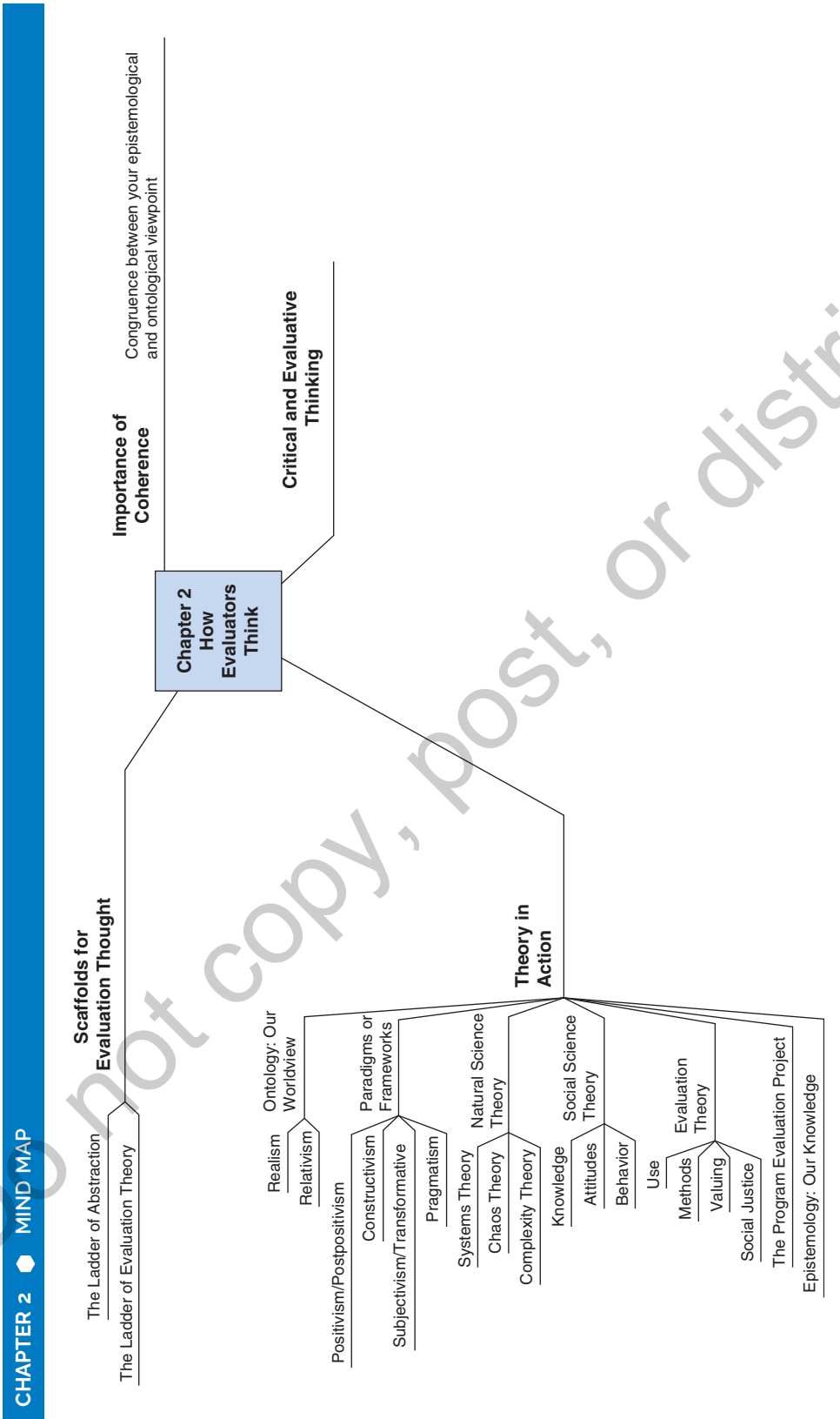




HOW EVALUATORS THINK

LEARNING OBJECTIVES

1. Distinguish between concrete and abstract ideas.
2. Explain the difference between the two worldviews that can be held by evaluators.
3. Describe four paradigms prevalent in evaluation along with the perspective on meaning, preferred methods, and critical thinking processes employed in each.
4. Describe some ways that the Socio-Ecological Model is useful in an evaluation context.
5. Identify four possible branches on the Evaluation Theory Tree and give an example of a theorist on each branch.
6. Explain why the assumptions and methodological decisions in any evaluation need to have coherence.



BEV'S PHILOSOPHY OF MORNINGS

I am not a morning person. Ask me to stay up late and accomplish something and I am your person. A few years ago, I came across a book called *Miracle Morning* (Elrod, 2017). The philosophy of this book is simple: the first 10% of your day can dictate and influence how the rest of your day goes. I wanted to try this partly because I have never looked at mornings favorably. Once the sun went down, I came alive. I needed to change my philosophy. The book hypothesized that if you make a set list of activities in the morning, every morning, the quality of your day will be improved. For six minutes (1 minute each), you have Silence, Affirmations, Visualization, Exercise, Reading, and Scribing. I tried this and noticed the benefits the first day.



Source: iStock.com/bruev

It is not only important to have a morning plan but also to have a philosophy about the activities planned. I changed the way I looked at my mornings, took a different approach and it worked. As a result, I was more focused and more energetic. By the end of the week, I was looking forward to my morning!

INTRODUCTION

Bev's morning story gives us a simple example of why having a philosophy is important as a guide to life. This chapter explores how evaluators view the world, how they gain knowledge, and what they do with that knowledge when planning an evaluation project. Evaluators use many terms, sometimes interchangeably, to describe their field. This can be muddy and confusing and often leaves us asking more questions that we had when we started. For this reason, we have gone back to first principles and asked one question, "How do evaluators think?" The answers may be surprising. The work of evaluators is embedded in a long history of scientific thought. How they see the world affects the work they produce.

The chapter is filled with metaphors, from ladders and amoebas to trees and puzzle pieces, all simple devices to illustrate the different ways in which evaluators think about evaluation, social science, and the nature of society. We will distinguish between concrete and abstract thinking, will look at four **paradigms** that underpin different approaches to evaluation, and will review how differing assumptions and frames of

reference influence how evaluation theorists think. Whatever their perspective, how they view the world influences how they design, collect, and interpret data. As we often cite evaluation theorists in our work, it is important to understand their philosophical stance to see if it fits with our own. Being able to interpret the through-line of evaluation thought in both your work and that of others will make you an agile and critical thinker.

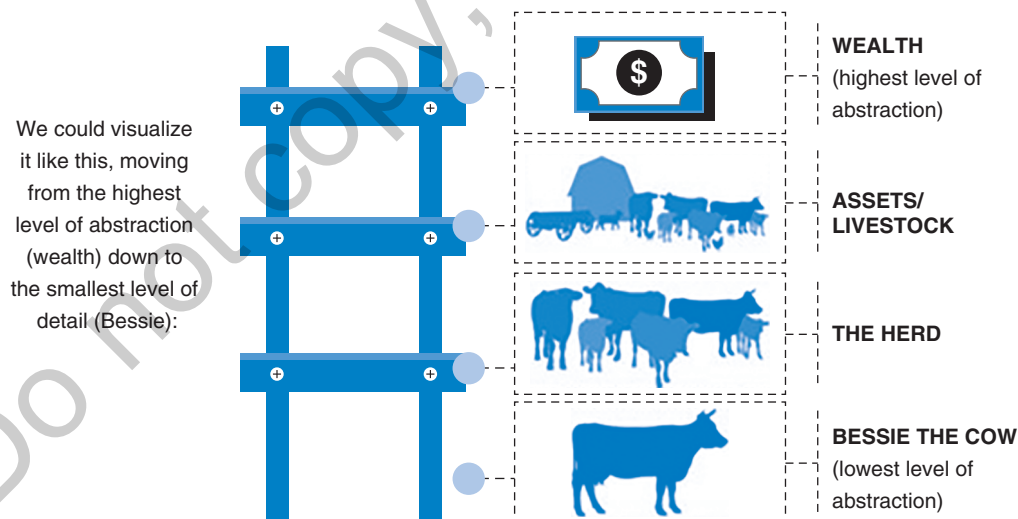
You will enjoy the resources in this chapter including the Spotlight on Equity, which presents systems thinking in rural medical care, our expert Dr. Donna Mertens, Key Terms, Main Ideas, Critical Thinking Questions, Student Challenges, and Additional Readings and Resources.

SCAFFOLDS FOR EVALUATION THOUGHT

The Ladder of Abstraction

Let's start with a simple and fun metaphor created by S. I. Hayakawa (1939) who wrote a book called *Language in Thought and Action*. It has been used widely to distinguish between very concrete things at the bottom of the ladder, like your favorite Netflix movie, and some very abstract ideas at the top, like the entertainment industry. For our purposes, we use an adaptation of Hayakawa's Ladder (Figure 2.1) to describe Bessie the Cow (Freishtat & Leipzig, 2019).

FIGURE 2.1 THE LADDER OF ABSTRACTION



Source: Freishtat, R., & Leipzig, A. (2019). Information acrobat: Climbing the ladder of abstraction.

To climb the Ladder of Abstraction, we start on the ground beside the physical cow. We know she is made up of atoms and molecules. We can actually touch her and give her a pat. As we start to climb the Ladder of Abstraction, we discover that her name is Bessie, a word that stands for her and is a kind of short-cut that incorporates all her particular characteristics. As we climb a little further, we find out that she is part of the “Herd” and shares similar features with the other herd members. Higher up, we learn that she is part of “Livestock” because she shares certain traits with pigs, chickens, goats, and other domestic animals. She represents an asset for the farmer, meaning that she is salable, like all the other assets on the farm. Finally, at the top of the ladder, we see that she represents the concept of wealth. Her monetary value is the only thing that remains at this high altitude. It is a pretty cold and abstract place up there, and we cannot tell that you are referring to Bessie at all when you say the word “wealth.”

As we start to climb the ladder, note that we know many details and facts about Bessie. The higher we climb, the fewer details we can carry. Similarly, the more abstract our thinking, the fewer details we can provide. What level should we choose? Of course, the answer is, “it depends.” If we want to talk about theory, we will probably find ourselves quite high up the ladder. If we want to talk about a specific evaluation project, we may find yourself quite close to Bessie.

What is the point of this simple metaphor? Evaluators need to be agile thinkers, able to climb up or down the Ladder, depending on the type of information we need. Do we need something that is more abstract? Climb up. If we need something more specific, climb down. People often talk past each other because they are operating on different levels of abstraction. When we ask questions, we need to aim them at the right level so that we get the information we need. Communicating well requires that we understand the level of abstraction that others are using and make sure that these abstractions are derived from the same specific information (Leviton, 2015, p. 241). If we ask a question that is either too general or too specific, we may get the wrong answer. We could end up looking like this unfortunate creature (Figure 2.2, p. 40).

The Ladder of Evaluation Theory

The great evaluation thinker, Michael Scriven, noted that “evaluation is a very young discipline—although it is a very old practice” (Scriven, 1996, p. 395). There has been much discussion about its theoretical and philosophical foundations. In their enthusiasm for the topic, however, evaluation theorists and writers have interpreted and reinterpreted their assumptions and language to explain their worldview until the field of evaluation has been littered with conflicting terminology. Such terms as paradigms, concepts, theories, frameworks, approaches, types, methods, and models are often used interchangeably. One evaluator’s theory is another’s framework and a third’s model. It can be difficult to parse out actual meaning.

FIGURE 2.2 THE COW AND THE LADDER



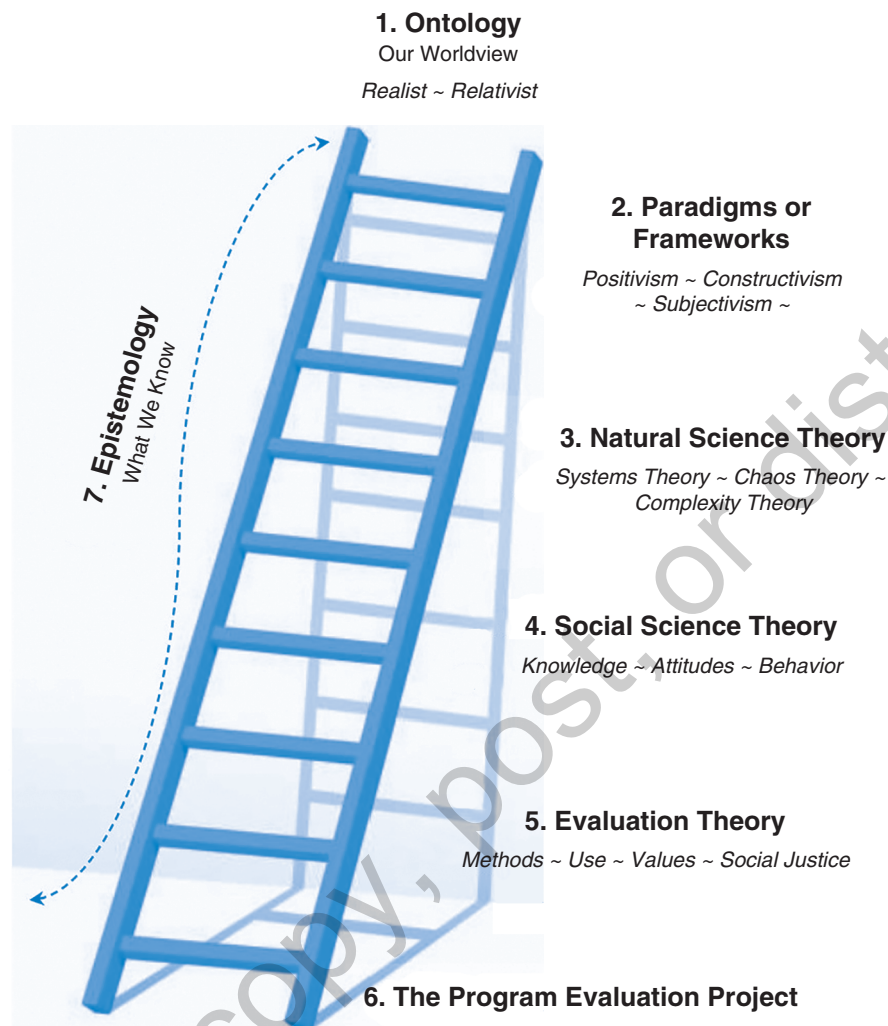
Source: Scottish Society for the Prevention of Cruelty to Animals, Scottish SPCA.

Why should we worry about philosophy? The answer is simple. If you do not understand the principles and assumptions embedded in your work, your entire evaluation study can be compromised. Muddy thinking leads to unclear study findings that are weak or indefensible. Our goal as evaluators is to affect positive social change and so we want to have the most impact we can. To do this, we need to be clear about our own worldview and know exactly why we choose the methods we do.

The literature on the philosophy of science is confusing, inconsistent, and sometimes, as Crotty (1998) suggests, even impenetrable. For this reason, we have taken a simplified approach, borrowing Hayakawa's metaphor to create our own Ladder—a serviceable and utilitarian tool to use as a scaffold. See the Ladder of Evaluation Theory (Figure 2.3, p. 41).

This visual shows how various theories are used in program evaluation and how they relate to each other. You could start anywhere on the ladder and climb up or down, but for simplicity's sake, we will start at the top which represents your worldview. Ontology, or your view of reality, is the most abstract level. We will climb down slowly, exploring and clarifying each idea until at last, we reach ground level, where any specific program evaluation study is located. Finally, we will step back to admire the overall epistemology or philosophical scaffold we have assembled, made up of the ideas or concepts we have selected that fit best together and provide the strongest foundation for our study.

FIGURE 2.3 LADDER OF EVALUATION THEORY



Source: istockphoto.com/frans_kuz

THEORY IN ACTION

Ontology: Our Worldview

What we know about the world is our **ontology**. Since Parmenides, who developed this term, philosophers have worried about this fundamental question, “What is reality?” How confident can an evaluator be about the phenomenon we are studying? Who decides what is real? What do we do with conflicting ideas of reality? Our most basic assumptions are of an ontological nature because we must decide if the reality we are investigating is external to the individual (i.e., real substances of the natural universe such as particles, atoms, and

genes) or if it is a product of their consciousness (i.e., a subjective or experiential interpretation of reality).

Realism

Realism holds that one reality exists, and humans experience it as “the truth” independent of who is viewing it (Moon & Blackman, 2014). Beliefs about reality include:

- There is only one truth
- Truth does not change
- Observation and measurement discover truth
- Truth can be generalized to other situations.

Relativism

Relativism means that reality exists in the mind of the individual and that each of us creates our own version of reality (Moon & Blackman, 2014). Beliefs about reality include:

- There are multiple versions of reality
- Context shapes reality
- Truth does not exist independent of meaning
- Truth evolves and changes based on experience
- Truth cannot be generalized.

Thus, ontology ranges from realism, where only one reality exists, to relativism, where multiple versions of reality exist side by side.

Let’s look at an example. When you see a red traffic light, your instinct is probably to stop your car. We know this is true based on our own experience; however, it may not always be the case. In Italy, for example, a place sometimes known for its cheery dysfunction, Severgnini (2006) explains:

When many Italians see a stoplight, their brain perceives not prohibition (Red! Stop! Do not pass!). Instead, they see a stimulus. OK, then. What kind of red is it? A pedestrian red? But it’s seven in the morning. There are no pedestrians about this early. That means it’s a negotiable red; it’s a “not-quite-red.” So we can go. Or is it a red at an intersection? What kind of an intersection? You can see what’s coming here, and the road is clear. So, it’s not a red, it’s an “almost red,” a “relative red....” And what if it’s a red at a dangerous intersection with traffic you can’t see arriving at high speed? What kind of question is that? We stop, of course, and wait for the green light. (pp. 9–10)

Severgnini suggests that Italians may see a red light as negotiable. Americans do not.

The National Community Anti-Drug Coalition Institute (2009) uses a phrase to help communities understand themselves: “But why, but why here” (p. 22)? These questions focus on why problems, such as teen drinking, occur in their community. The question pushes to understand what it is about the local area that contributes to a high incidence. For example, in one community, alcohol vendors may take a casual approach to checking identification. In another, adults may buy alcohol for youth or condone drinking behaviors (e.g., parents hosting teen drinking parties). In a third community, parents may work actively to support a “dry grad,” an all-night party which follows the formal convocation, banquet, and dance where no alcohol or drugs are allowed. Each community has its own unique reality, and it is likely that each approach will produce different outcomes.

Paradigms or Frameworks

We categorize our knowledge in a paradigm or framework to help us interpret the world. Thomas Kuhn (1962), an influential twentieth century science historian, described a paradigm as a coherent tradition of scientific research (p. 11). It attracts a community of practitioners toward shared rules and standards for scientific practice, drawing them away from other schools of thought. Their views reinforce each other’s views, and their shared preconceptions and assumptions shape the research approach that they endorse, such as (Kuhn, 1996):

- What to observe
- What questions to ask
- How to structure these questions
- How to conduct an experiment
- What outcomes to expect
- How to interpret the results.

Over time, however, a paradigm’s discrepancies and anomalies begin to accumulate. Eventually, a thought revolution occurs, and a new theory emerges that appears to be more relevant than the old one. Kuhn described this as a **paradigm shift** or scientific revolution (Kuhn, 1962, p. 90). An example of a contemporary paradigm shift in medicine is the transition from clinical judgment to evidence-based medicine (Eddy, 1990).

Four paradigms are prevalent in evaluation literature: positivism, constructivism, subjectivism, and pragmatism (Crotty, 1998; Moon & Blackman, 2014). Refer to [Table 2.1](#) on page 44. Each paradigm represents a different set of assumptions about the nature of reality, and each has preferred evaluation methods and ways of thinking critically. Evaluators generally feel more comfortable in one specific paradigm and tend to design their work from that perspective. Over time, however, and depending on the nature of their work, evaluators can move from one perspective to another.

TABLE 2.1 ■ FOUR PARADIGMS FOR EVALUATORS

Positivism/Postpositivism	Constructivism	Subjectivism/Transformative	Pragmatism
Only one reality.	Multiple realities.	Meaning exists within the individuals studied.	It depends.
The evaluator discovers reality through observation and experimentation.	The evaluator solves problems through interaction with participants.	The evaluator uses the participants' perspective to interpret reality.	The evaluator focuses on the outcomes and consequences of a specific problem.
Methods: Quantitative.	Methods: Qualitative	Methods: Mixed	Methods: All methods appropriate to the situation
Critical thinking: Deductive	Critical thinking: Inductive	Critical thinking: Reflective	Critical thinking: Abductive

Source: Crotty, M. (1998). *The foundations of social research: Meaning and perspectives in the research process*. SAGE Publications; Moon, K., & Blackman, D. (2014). A guide to understanding social science research for natural scientists. *Conservation Biology*, 28, 1167–1177.

In [Figure 2.4](#), p. 45 we present our model, Miranda, who demonstrates how an individual perceives reality:

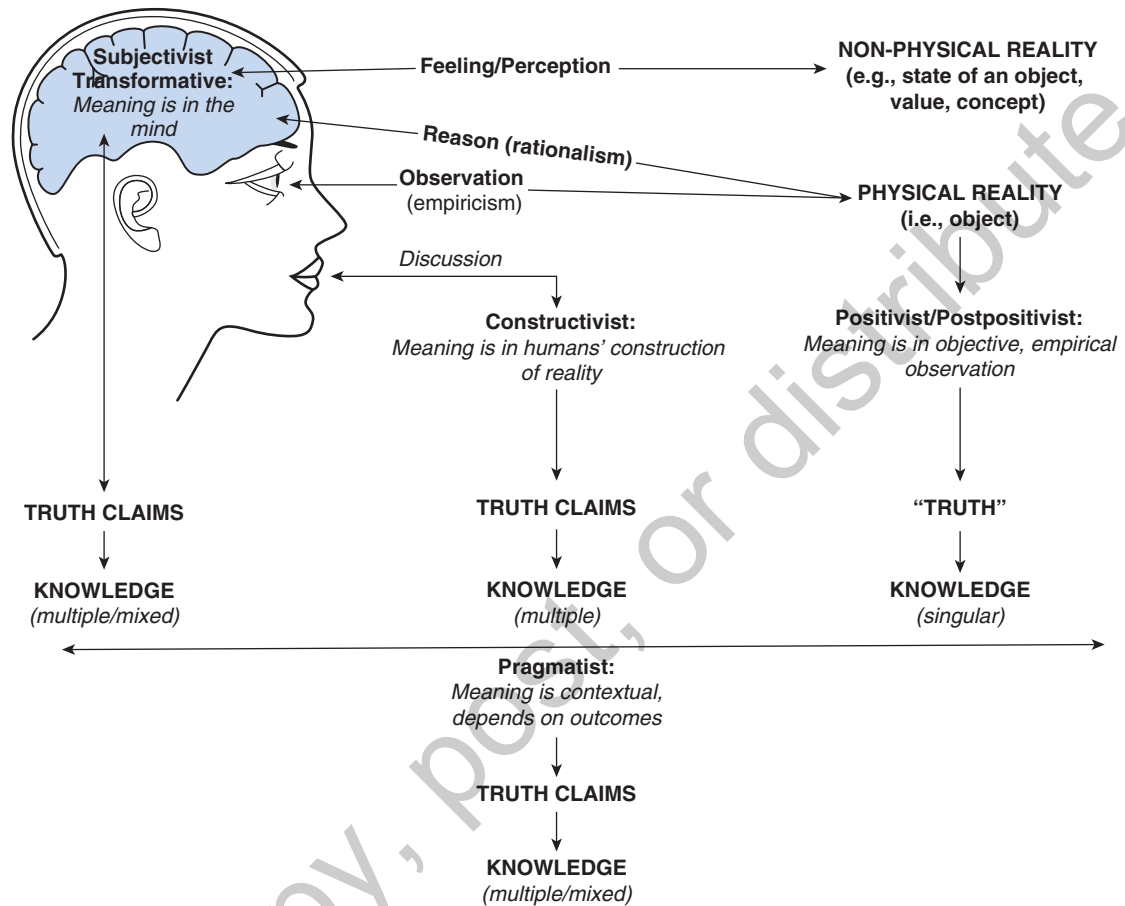
- If she is a positivist or postpositivist, she relies on observation and reason to arrive at a singular truth. She tends to use quantitative methods.
- If she is a constructivist, she relies on discussion and interaction with stakeholders and participants to jointly construct a shared version of reality. She tends to use qualitative methods.
- If she is a subjectivist and uses the transformational paradigm, she focuses on feelings, perceptions, and values, usually of marginalized groups. She relies on understanding multiple versions of reality simultaneously. She tends to use mixed methods.
- If she is a pragmatist, she is open to any of these paradigms, depending on the circumstances, and is likely to use at least two in combination.

Let us explore the four paradigms.

Positivism/Postpositivism

Positivism, or more recently postpositivism, is close cousins with realism, logical positivism, objectivism, and empiricism. Positivists see meaning as existing within an object, independent of whoever is observing it. For them, reality exists outside of the individual mind. As Trochim (2006) explains, “the goal of knowledge is simply to describe the phenomena that we experience. The purpose of science is simply to stick to what we can observe and measure. Knowledge of anything beyond that, a positivist would hold, is impossible” (para. 4). Thus, knowledge comes from experimentation. The universe is deterministic and operates by laws of cause and effect. We postulate theories and test

FIGURE 2.4 REALITY, MEANING, TRUTH CLAIMS, AND KNOWLEDGE ACCORDING TO VARIOUS EPISTEMOLOGIES



Source: Adapted from Moon, K., & Blackman, D. (2014). A guide to understanding social science research for natural scientists. *Conservation Biology*, 28, 1167–1177. [istockphoto.com/juliawhite](https://doi.org/10.1007/s11771-014-0177-1)

them. If we find that our theory does not fit the facts we uncover, we can revise our theory to predict reality better next time.

Since the mid-twentieth century, there has been a shift toward a broader interpretation of positivism. **Postpositivism** uses multiple methods because any one method is prone to error. While the positivist believed the goal of science was to uncover the truth, the postpositivist believes that “the goal of science is to hold steadfastly to the goal of getting it right about reality, even though we can never achieve that goal!” (Trochim, 2006, para. 6). While positivists rely on quantitative methods (e.g., surveys, biometrics), postpositivists can use both quantitative and qualitative methods (e.g., interviews, focus groups, case studies) as valid approaches. Both the positivist and postpositivist believe in empiricism—observation and measurement are the core of the scientific endeavor. This is embodied most closely in the scientific method.

The Scientific Method. Grounded in observation, experimentation, and replication, the **Scientific Method** is based on studying cause–effect relationships in the natural world. To understand the human condition, scientific thought moved away from philosophy and mysticism toward observation and nature. A Muslim scholar, Ibn al-Haytham (c. 965 to c. 1040), wrote the *Book of Optics* in which he described a scientific process grounded in observation, experimentation, and replication. Later, his writings were translated into Latin and influenced European Renaissance thinkers like Galileo (1564–1642), Descartes (1596–1650), and Newton (1642–1726) (United Nations Educational, Scientific and Cultural Organization, 2015). The process became known as the scientific method. It has been the gold standard for studying cause–effect relationships in the natural world ever since and continues to permeate scientific thought today.

Although procedures can vary from one field to another, in general, the steps are shown in [Figure 2.5](#), p. 47.

The value of objectivist research is its focus on validity (applicability of the research in other contexts) and reliability (consistency of results obtained) (Moon & Blackman, 2014).

Foundational ideas in the positivist/postpositivist paradigm include (Patton, 2002, 2015):

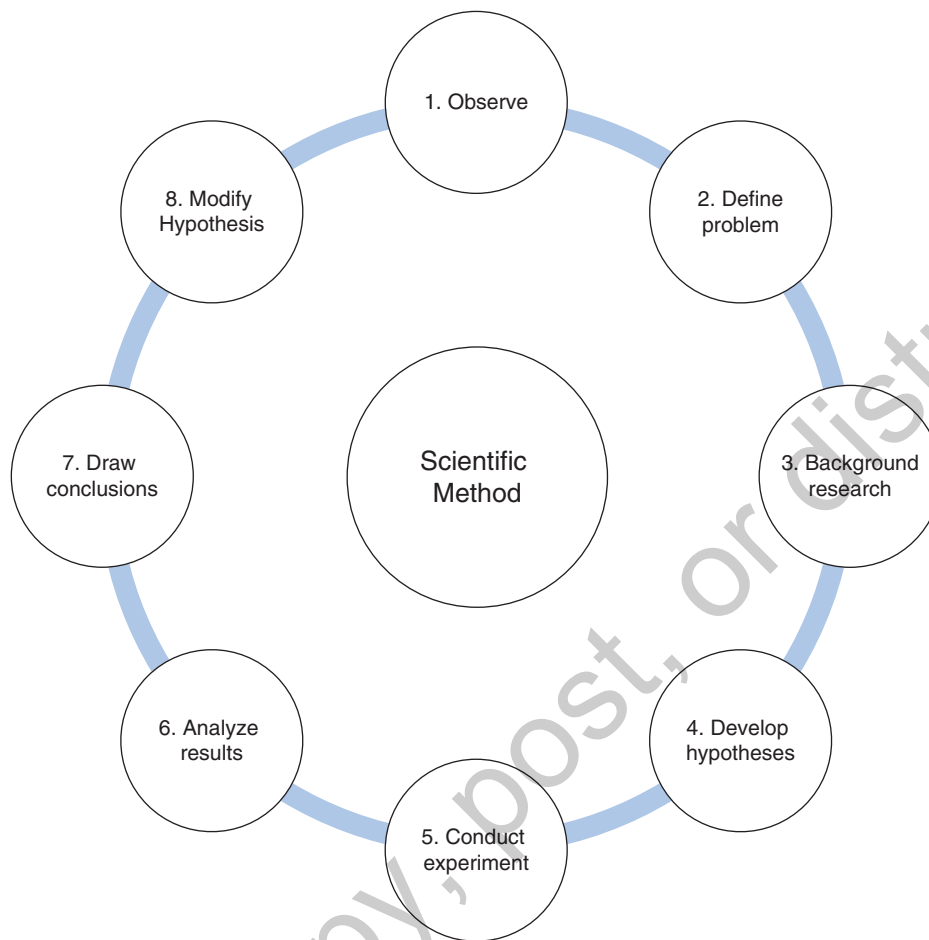
- A real, observable, predictable world exists; researchers seek correspondence with it.
- Where scientific credibility is paramount, it is the dominant perspective of policymakers and commissioners of evaluation.
- Its focus is on validity, reliability, and objectivity.

Positivists ask such questions as (Patton, 2015):

- What is going on in the real world?
- What can we establish with some degree of certainty?
- What are plausible explanations for verifiable patterns?
- What is the truth insofar as we can get at it?
- How can we study a phenomenon so that our findings correspond to the real world as much as possible?

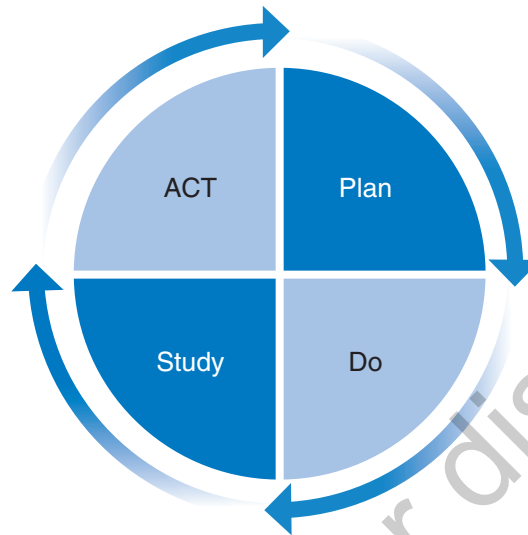
Plan-Do-Study-Act (PDSA) Cycle. Often used as an improvement strategy in organizations, the PDSA Cycle is an outgrowth of the scientific method. Moen and Norman (2009) tracked the evolution of scientific thought from the work of Galileo (1564–1642) to the modern day and related it to the work of statistician and management consultant, Edward Deming in the 1950s. Deming created it when he worked with Japanese auto industry leaders, and many credit his ideas about processes as one of the drivers for the Japanese postwar economic miracle

FIGURE 2.5 THE SCIENTIFIC METHOD WHEEL



(The W. Edwards Deming Institute, 2019). He taught that using appropriate management principles could increase quality and decrease costs by reducing waste, rework, and staff attrition, and by increasing customer loyalty (Halwes, 1998). Originally known as the Deming Wheel, the PDSA Cycle is still common today, particularly in healthcare (Taylor et al., 2014). See [Figure 2.6](#), p. 48 it offers “a continuous and ongoing effort to achieve measurable improvements in the efficiency, effectiveness, performance, accountability, outcomes, and other indicators of quality in services or processes...” (Riley et al., 2010, p. 5). Data are collected throughout the cycle to determine if a new intervention is more effective than previous practice and conclusions are used by management to inform changes in policy and behavior. We provide more details about the PDSA Cycle in Chapter 5 as a midcycle program evaluation method.

FIGURE 2.6 ■ THE PLAN-DO-STUDY-ACT CYCLE



Source: PDSA Model courtesy of The W. Edwards Deming Institute®.

Deductive Thinking. Positivists tend to use deduction as their preferred method of critical thought. The researcher examines the known facts and, like puzzle pieces, tries to fit them into a preexisting framework such as a hypothesis or a conceptual framework. Quantitative inquiry employs deduction (see Figure 2.7).

FIGURE 2.7 ■ DEDUCTIVE THINKING



Source: thenounproject.com/Wes Breazell

In this graphic, you can see that something is missing, likely the answer to the research question or hypothesis posed. The data that the researcher collects either prove or disprove

(i.e., it fits or it doesn't fit) the research hypothesis. Note how the research is not really interested in the surrounding puzzle pieces but has a clear focus on finding the missing piece.

Milkie and Warner (2011) used deductive reasoning to study the mental health of first graders ($N = 10,700$) who had different classroom environments. The authors based their hypothesis on earlier research and a theoretical understanding of these issues. The authors believed that negative classroom or system characteristics contributed to children's emotional and behavioral issues. These characteristics include lack of basic school supplies, heating, interprofessional relationships of the teachers in the school, excessive paperwork demands on administrators, perceived low-academic standards, and negative behavioral situations of student's peers. The mental health outcomes studied included, for example, attention span, ability to persist at a task, argument frequency, fights, impulsivity, ability to form friendships, display of sensitivity toward others, anxiety, and low self-esteem. In this example, the authors posed a general idea (social science theory) of the way the world works based on their own ideas. Reading the literature, they designed a study to examine the hypotheses, and supported or rejected them based on the analysis of their findings.

The scientific method continues to dominate the fundamental and natural sciences. We rely on scientists to produce findings that are the result of objective, standardized, and replicated experimentation. Historically, evaluation was also dominated by the scientific paradigm; however, sometimes this resulted in evaluations that did not address the issue in question (Mertens, 1999, p. 4). As life is both complex and messy, so programs that attempt to address these tough issues can lead evaluators to avoid or ignore what is deemed to be unmeasurable. In cases like these, other paradigms may be more appropriate.

Constructivism

Constructivism is a paradigm that rejects the idea of objective truth. It is closely linked to constructionism and social constructionism. Meaning is found in the interaction between the subject and the object. Humans construct knowledge as they engage with and interpret the world, so different individuals can construct different meanings about the same object.

For example, Kant challenged the empiricists' claim that research must include measurement (Mertens & Wilson, 2019). The positivist view acknowledges an independent reality and a single universal truth, but social scientists, in trying to understand social phenomena, often work with individuals who hold different, and sometimes conflicting, views of the truth. Rawls (1971) suggested that nobody has privileged access to the truth; all have equal standing, which brings us to the idea of lived experience. A constructivist tries to understand meaning from the perspective of the people who have had the experience (Schwandt, 2000). The value of constructivist research is in generating a contextual understanding of a defined problem, beneficial to stakeholders who need to design contextually relevant responses to social problems.

The key ideas in the constructivist paradigm include (Lincoln & Guba, 1980; Patton, 1990):

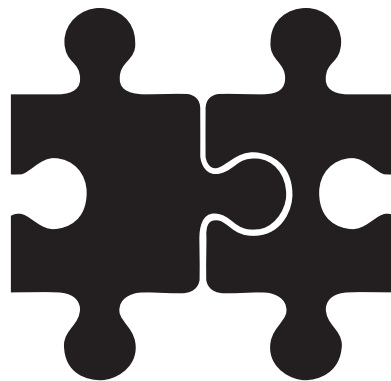
- Human perception is a constructed version of reality shaped by culture and language; therefore, we study multiple realities differently.
- Truth is a matter of consensus, and phenomena are only understood within a particular context.
- Stakeholders have different program experiences; one perspective is not more valuable than another despite power differentials.
- Scientists are bound by socially constructed and consensually validated knowledge but can still function in a positivist paradigm.

Constructivists ask questions like these (Patton, 2015):

- How have the people in this setting constructed reality?
- What are their reported perceptions, “truths,” explanations, beliefs, and worldviews?
- What are the consequences of this construction for their behaviors and for those with whom they interact?

Inductive Thinking. The type of critical thinking most often used by constructivists is induction. The researcher starts with a puzzle piece or a relevant response to a broad research question. Then they look for other puzzle pieces, or responses, until they find one or more pieces that fit the first one (see [Figure 2.8](#)).

FIGURE 2.8 ■ INDUCTIVE THINKING



Source: [thenounproject.com/Marcio Duarte](http://thenounproject.com/Marcio-Duarte)

By searching and comparing, the researcher slowly starts to develop some other pattern, and slowly, piece by piece, they build connections that develop the story, theory, or description. Induction is frequently used in qualitative inquiry (Mayan, 2009).

Allen, Kaestle, and Goldberg (2010) investigated the process of how young men know and understand menstruation. Evaluators conducted focus groups with 23 young men and then generated a theory based on the qualitative data (i.e., written text of the focus group discussions). Based on this inductive approach of collecting, comparing, and finding connections, the authors developed a theory that young men receive information about menstruation from their sisters.

Subjectivism

Sometimes called the transformative paradigm (Mertens, 2020; Mertens & Wilson, 2019), those who practice **subjectivism** find meaning within the individuals being studied and then interpret the phenomenon using the participants' perspective. As Mertens and Wilson (2019) explain, the transformative paradigm holds

the belief that the lives and experiences of diverse groups of people are of central importance in the evaluation to address issues of power and justice. [It] focuses primarily on eliciting and understanding viewpoints of marginalized groups and interrogating systemic power structures through mixed methods to further social justice and human rights. (p. 548)

Unlike positivism, which looks for a single truth and tries to avoid external influences, and constructivism, which focuses on multiple perspectives and lived experience, subjectivism actively incorporates values, culture, power dynamics, and social justice into its design. As a result, it is an appropriate paradigm for evaluators who are working with marginalized groups so that they can listen to the voices of the oppressed and also address larger systemic issues.

There are many approaches to subjectivism. Main branches include (Magana, n.d.):

- Critical theory—focuses on how injustice and subjugation shape people's understanding and experience of the world
- Queer theory—focuses on sexual orientation
- Indigenous and postcolonial perspectives—focus on challenging existing colonial attitudes, institutions, and practices and understanding problems in their sociocultural context
- Feminist inquiry—focuses on the importance of gender in human relationships and social processes.

The evaluator attempts to understand individuals' knowledge, interests, purposes, and values (Schwandt, 2000). The focus of the inquiry is often on power inequities, the impact of privilege, and how the consequences of these affect social justice. It places central importance on the experiences of marginalized groups (Mertens, 1999):

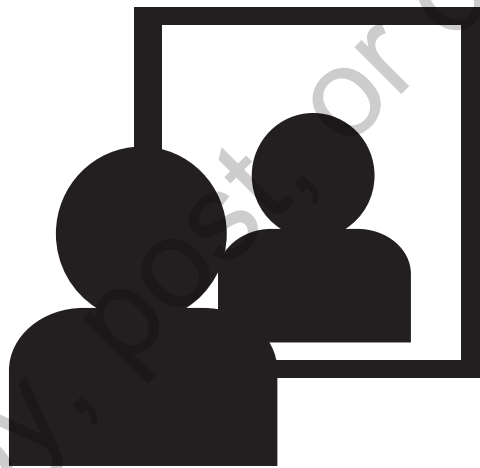
such as women, ethnic/racial minorities, people with disabilities, and those who are poor. The evaluator who works within this paradigm consciously analyzes

asymmetric power relationships, seeks ways to link the results of social inquiry to action, and links the results of the inquiry to wider questions of social inequity and social justice. (p. 4)

With its ideological lens, subjectivism avoids claiming open-mindedness in favor of the emergent theory. Thus, it is at the opposite end of the research spectrum from positivism. The value of subjectivism or a transformative view is that it reveals how an individual's experience shapes their perception of the world (Moon & Blackman, 2014) and provides essential insight into their behavior.

Reflective Thinking. Critical thinking for the transformative evaluator involves self-reflection and dialogue (see [Figure 2.9](#)).

FIGURE 2.9 REFLECTIVE THINKING



Source: thenounproject.com/bmijnlieff

The researcher assesses personal assumptions and positions of privilege, becomes critically aware of how he/she perceives, understands, and feels about the world, and reformulates those assumptions to permit a more inclusive, discriminating, permeable, and integrative perspective (Mezirow, 1990). Reflection leads to dialogue, respect, and collaboration. The transformative evaluator believes that reality exists in the groups' voices.

To ground the research, transformative evaluations use both qualitative and quantitative data collection methods in what is known as a mixed methods approach. The evaluation goes through several cycles of data collection, generally organized in a sequential fashion so that data obtained in the first cycle are used to inform the questions of the next cycle. Sometimes, however, two data collection initiatives are launched simultaneously. (See Chapter 9 for more information on mixed methods.)

Silka (2009) studied Laotians who lived in Massachusetts and became ill after fishing in a contaminated lake. The researchers trained the Laotians to become part of the research team. They conducted blood tests and collected interview data. Their findings were presented at fishing festivals to the Laotian community, providing education on the effects of contaminated fish on their health. Together, they were then able to devise new fishing methods. By using a transformative approach, the team was able to develop relationships with the Laotians, so that their reality could be incorporated into both study outcomes and solutions.

Pragmatism

Pragmatism aligns closely with much of evaluators' work because of its focus on the practical application of ideas and problem-solving in the real world. However, rather than focusing on truth and reality, pragmatism is grounded in consequences and results. Thus, evaluators can use a variety of appropriate methods without being committed to any specific philosophical position.

Pragmatism originated in the United States in the late nineteenth century. It was founded by psychologist William James, philosopher Charles S. Pierce, educator John Dewey, and sociologist George Herbert Mead, all members of the Chicago school of thought. They proposed that consequences determine the meaning of actions and beliefs. Morgan (2014, pp. 26–27) explained three other essential elements of pragmatism:

1. In judging actions, we consider the situations and contexts in which they occur.
2. Actions link to consequences in ways that are open to change.
3. Actions depend on worldviews that are socially shared sets of beliefs.

Rather than emphasizing the nature of reality, pragmatists focus on the outcomes of action. Thus, pragmatism becomes a helpful paradigm for evaluators who work mainly in the realm of program results.

A question for pragmatists is:

- What difference would it make to act one way rather than another?

Decision making becomes a critical link between belief and action (Morgan, 2014, p. 28) and so for evaluators, a pragmatic approach:

accepts that quantitative, qualitative, and mixed research are all superior under different circumstances [original italics] and it is the researcher's task to examine the specific contingencies and make the decision about which research approach, or which combination of approaches, should be used in a specific study. (Johnson & Onwuegbuzie, 2004, pp. 22–23)

Rather than using the scientific method, Morgan (2014, pp. 29–30) suggests Dewey’s five-step problem-solving framework to reflect on a problem and arrive at an appropriate action. For example, if an evaluator is trying to decide what methods to use in a specific program evaluation, their thinking process might look like.

Step 1 Recognize the problem.

Step 2 Based on existing beliefs, search for a possible solution.

Step 3 Speculate about and suggest a solution or set of actions (a leap of faith).

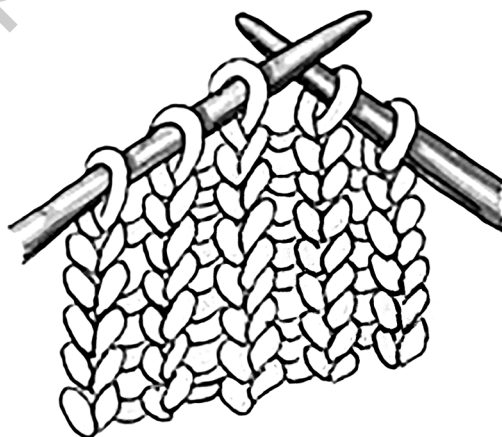
Step 4 Assess the tentative solution, reflect on its potential effectiveness.

Step 5 Take action.

Looking at Dewey’s Steps 3 and 4, for example, you can consider possible explanations, best guesses, speculations, conjectures, and hypotheses; then review, reframe, and revisit until you eventually arrive at the most plausible interpretation (Mayan, 2009; Patton, 2015).

Abductive Thinking. The critical thinking that pragmatists tend to use is known as abduction. It is a kind of in-between type of thinking, not entirely deductive, not fully inductive, but a combination of the two. In a mystery novel by Siger (2010), a police chief inspector used what he called “knitting” to go back and forth between what he knew and the gaps he needed to fill (Figure 2.10). He knit back and forth through his ideas until he could create meaning out of the facts. This knitting back and forth is an excellent way to think about abduction, going from induction to deduction, from framework to data, and back again.

FIGURE 2.10 ■ ABDUCTIVE THINKING



Source: istockphoto.com/Katsiaryna Pleshakova

When we use abduction, we usually have incomplete information, and we search for the likeliest explanation. For example, the thought process that your doctor uses to generate a medical diagnosis is a result of abductive reasoning. Two systems of thinking are at play at each stage of clinical reasoning: “intuitive, fast, and almost unconscious thinking,... and slower, analytical, effortful thinking” (Canadian Medical Protective Association, 2021, para. 2; Croskerry, 2018). This dichotomy of thinking methods is similar to that explained by Kahneman (2011), in which System 1 thinking is fast, intuitive, emotional, and easier, and System 2 thinking is slower, more deliberative, more logical, and more difficult. Abductive thinking incorporates both.

Silver Blaze is a famous story involving the great detective, Sherlock Holmes (Doyle, 1894). It tells about the curious incident of the dog in the nighttime and provides an example of abduction. Holmes solved the mystery by observing what did **not** happen. The fact that the dog did not bark when you would expect it to led Holmes to conclude that the evildoer was not a stranger but rather was someone the dog recognized. That is why the dog did not bark. Holmes drew his conclusions from a fact that did not occur, or a negative fact. The dog did not bark in the night (the result) and moving backward from there, he asked, “Why?” He was able to conclude that the dog knew the intruder (the cause).

Natural Science Theory

Evaluators are eclectic by nature and have often found useful parallels to their work in the physical and biological sciences. Here are three examples of how the natural sciences have influenced evaluation thought: systems theory, complexity theory, and related to the complexity theory, chaos theory.

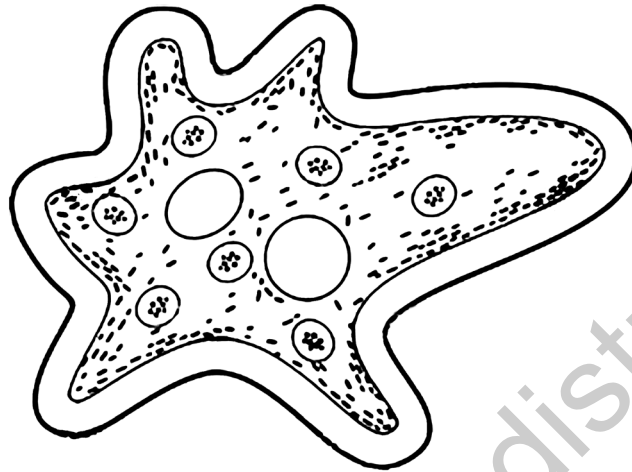
Systems Theory

An interdisciplinary study of natural systems is attributed to the work of biologist von Bertalanffy (1968). His General Systems Theory looks at the interrelated, interdependent parts of either a natural or artificial entity. A system has boundaries that distinguish it from its environment, and it adapts as needed to protect itself and preserve its purpose. A change to one part of the system affects other parts as well.

Many fields have adopted **systems theory**, including biology, ecology, organizational theory, management, and program evaluation. While in the past, science tried to explain phenomena by reducing them to independent units, von Bertalanffy noted that the interaction between a phenomenon’s parts is dynamic, as is the relationship between the phenomenon itself and its environment, all moving parts, all dynamic, all changing.

Think, for example, of a one-celled freshwater amoeba (Figure 2.11, p. 56) swimming in a soupy pond. To survive, it feeds on algae and bacteria and eliminates waste, thus creating

FIGURE 2.11 SYSTEMS THEORY AND THE AMOEBA



Source: istockphoto.com/mariaflaya

its perfect internal system. By eating the algae, the level of algae on the pond's surface allows water plants to survive through photosynthesis (Berry, 2016). As the plants oxygenate the water, other species have the oxygen they need to survive as well.

Von Bertalanffy described systems as being either closed or open (Frye & Hemmer, 2012). A closed system is one where nothing either enters or leaves the system. We find it very hard to think of any system today that does not interact with its environment. Gail was reminded of this when she visited the Eastern Orthodox monasteries in Meteora, Greece, perched high on remote rocky precipices. They were originally founded by monks who chose a solitary life for contemplation and prayer. However, they found that their survival depended on a rope and pulley system to haul up food and building supplies, as well as new members to their order. The late Fred Boissonnas, a famous Italian photographer, recorded his experience accessing this failed attempt at a closed system (see Figure 2.12, p. 57).

It was not possible to live in a completely closed system and today, we find beautiful monasteries on those rocky outcrops, built piece by piece by the monks and their successors. Thousands of tourists visit them every year, thus contributing to the larger system, the Greek economy. For breathtaking images, see Taylor (2018). It seems, then, that all systems are open to some extent.

An example of a large open system is the public health system (Figure 2.13, p. 58) comprised of “public, private, and voluntary entities that contribute to the delivery of essential public health services within a jurisdiction” (Centers for Disease Control and Prevention, 2020, para. 1).

FIGURE 2.12 ■ PROOF OF THE FAILURE OF A CLOSED SYSTEM

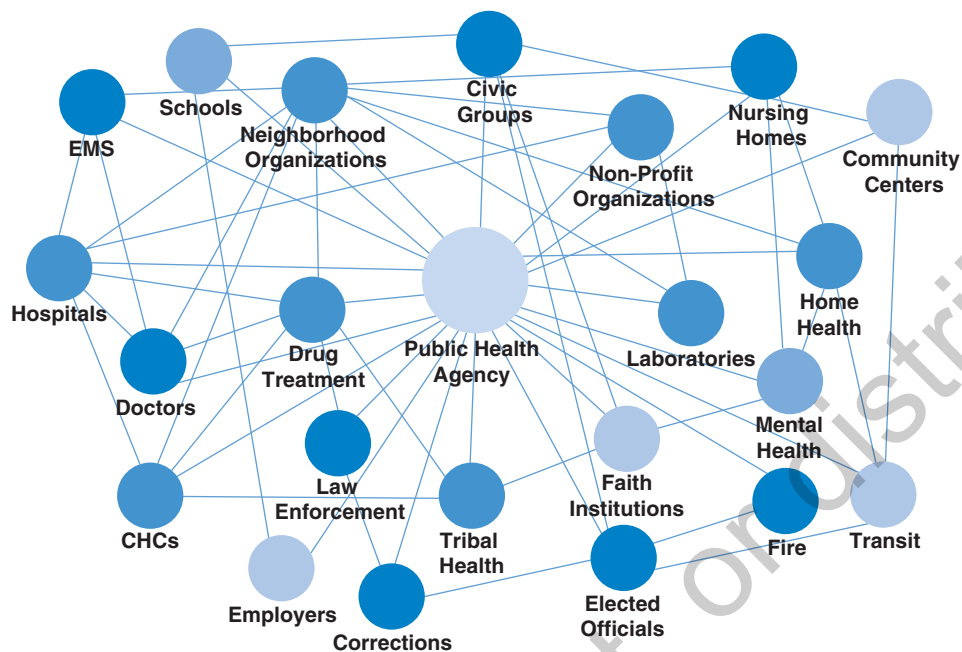


Source: thesprotia-news.blogspot.com/Meteora, ascent of Fred Boissonnas with basket

Each part of a system interacts internally with the others, but also interacts with the external environment, made up of clients, funders, stakeholders, the government, and the public. A change in any one of those internal or external components will result in a change in the others.

In early January 2020, the World Health Organization (WHO) announced a mysterious coronavirus-related pneumonia. On January 21, 2020, CDC confirms the first United States coronavirus case. By January 31, 2020, the WHO issues a global health emergency, and it became a public health emergency (American Journal of Managed Care Staff, 2021). In hindsight, there was not one aspect of the global system that has not been impacted: schools, hospitals, grocery stores, business, airlines, churches, the economy, and more. Systems operate best when the individual players in the system are familiar with other parts of the system before a problem occurs, when the system's true function is tested. In chaotic or emergent situations, the players know each other, have created agreements, and understand the functions and roles each other plays, so that in a near seamless manner they can work together to ensure the public's health.

FIGURE 2.13 THE PUBLIC HEALTH SYSTEM NETWORK

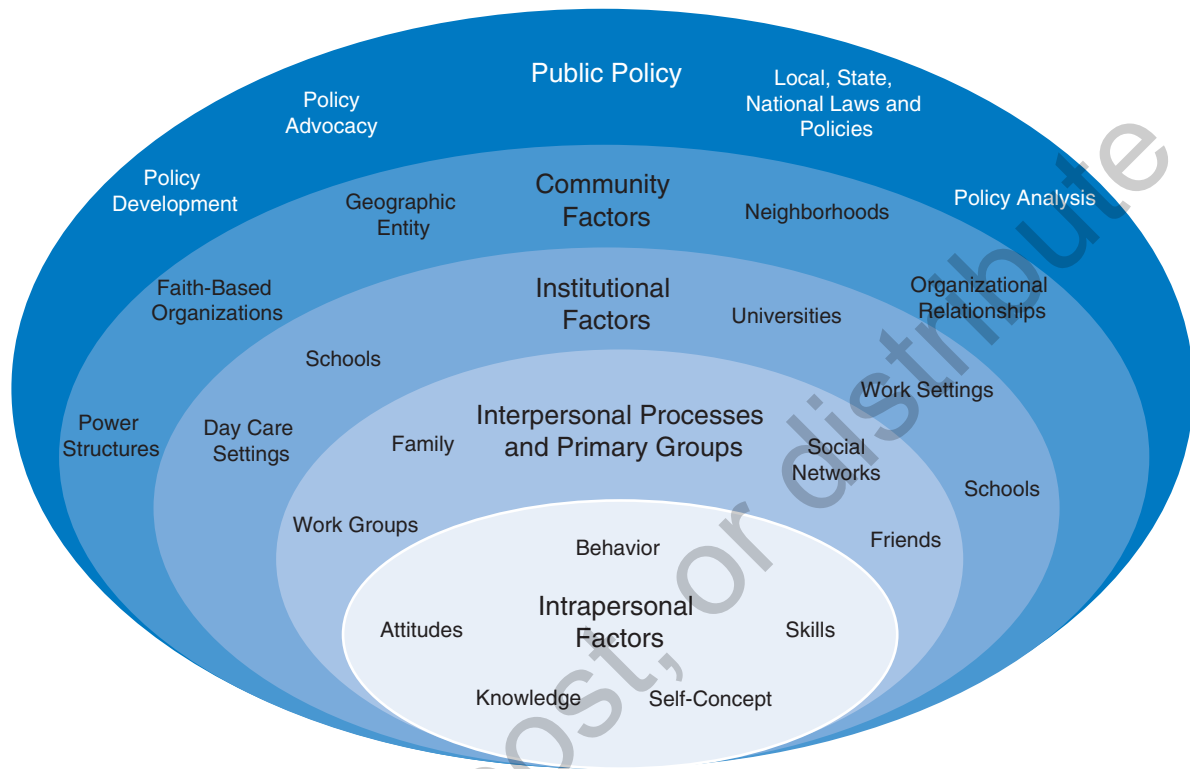


Source: Centers for Disease Control and Prevention. (2020). Original essential public health services framework: The public health system.

The Socio-Ecological Model (SEM). One way to view the interactions between people and their environments has been captured by the Socio-Ecological Model (SEM). Bronfenbrenner (1977) first introduced the model in the late 1970s to understand human development, especially to see how children were influenced by the interactions and relationships of an expanding set of systems. McLeroy et al. (1988), in their seminal article, described a layered framework which “serves to direct attention to both behavior and its individual and environmental determinants” (p. 354). Figure 2.14, p. 59 illustrates their model with the intrapersonal (individual), interpersonal, institutional factors, community factors, and public policy elements. Table 2.2, p. 59 provides examples of each level of the SEM.

Many in the social sciences have adapted the SEM to illustrate various social issues, such as the CDC which has adapted the framework for various health promotion activities (e.g., violence prevention, healthy college campuses, geriatric preventive health, and colorectal cancer prevention) (Kilanowski, 2017). The model can reflect the multiple factors that influence the health system, enabling an examination of the barriers and facilitators at each level.

FIGURE 2.14 THE SOCIO-ECOLOGICAL MODEL



Source: Adapted from McLeroy, K. R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly*, 15(4), 351–377.

TABLE 2.2 EXAMPLES OF THE SOCIO-ECOLOGICAL MODEL FOR ENVIRONMENTAL SUSTAINABILITY

Level	Example
Intrapersonal Factors	<ul style="list-style-type: none"> Teach children the everyday items that are recyclable.
Interpersonal Processes and Primary Groups	<ul style="list-style-type: none"> Post a toolkit to teach families in the community how to reduce, reuse, and refuse in their home.
Institutional Factors	<ul style="list-style-type: none"> Create a company-wide environmental plan to reduce waste in purchases.
Community Factors	<ul style="list-style-type: none"> Provide recycling centers in neighborhoods or special bins at houses.
Public Policy	<ul style="list-style-type: none"> Tax incentives for donating clothing, making energy improvements to homes, buying electric vehicles, using mass transit, bicycling to work Eliminating or reducing the use of single-use products.

Next, you will read about systems evaluation in rural and remote areas.

SPOTLIGHT ON EQUITY

Systems Evaluation in Rural and Remote Areas

The health outcomes of people living in rural and remote areas are poorer relative to those of people living in urban areas. One major contributor is the lack of proximity to healthcare facilities. For people living in rural states like Wyoming and Montana and provinces like Saskatchewan and the Yukon, it is not uncommon for the nearest healthcare facility to be over an hour away. Early intervention is vital to good outcomes for a heart attack or stroke. Without immediate intervention, such as CPR, brain damage begins to occur in 3 minutes and after 10 minutes most patients will not survive. In a rural emergency response situation, the situation is exacerbated by transport times as the distance to the nearest hospital with lifesaving expertise and equipment can be significant.

Dr. Ralph Renger and his evaluation team addressed these issues in their evaluation of a rural healthcare response system by using systems thinking to enhance their evaluation. It was a particularly informative approach because the ability to get a patient as quickly as possible to a definitive care center (or heart hospital with qualified staff and equipment) depended on the coordination of many moving parts (Renger et al., 2020). People, equipment, agencies, and technology all had to be coordinated for patients experiencing an emergent heart condition to survive. Finding efficiencies in the system could save minutes that could make the difference between life and death. The team chose system evaluation theory (SET) to guide their analysis. SET consists of three guiding principles derived from both system theory and evaluation theories about use, value, and methods (Renger, 2015, p. 17). SET processes reflect these principles and have three main steps: define the system, evaluate system efficiency, and evaluate system effectiveness.

Step 1. Define the system. Before evaluating the system, it must be defined. This involves determining system boundaries, subsystems, processes, relationships, feedback mechanisms, attributes, inputs, and goals. Like building a jigsaw puzzle, the evaluators began by defining the boundary for the study, then filling in the pieces, and then determining how the pieces interacted. They talked with state leaders to identify a goal for the cardiac system. “Getting the patient to definitive care in the shortest time possible” became their shared goal. Then they asked the leaders to identify system elements, and five subsystems were confirmed:

- **Dispatch:** managed by law enforcement.
- **Volunteer Emergency Medical Services:** the first responders to the scene with the local knowledge needed to locate the rural caller and administer CPR.
- **EMS Service with paramedics:** able to administer CPR and lifesaving heart medication.
- **Critical access hospital:** the closest rural hospital, unlikely to have the expertise and equipment needed.
- **Definitive care facility:** with cardiac expertise and equipment, located in an urban center.

The evaluators then worked with representatives of each subsystem to map their standard operating procedures (SOPs).

Step 2. Evaluate system efficiency. To evaluate system efficiency, the evaluators needed to assess the feedback mechanisms and the extent of goal alignment among the subsystems. They conducted a sudden cardiac arrest (SCA) simulation, mimicking an actual event in which people and resources were deployed. The simulation revealed that SOPs were not executed with fidelity. To understand the discrepancies, they applied systems thinking, exploring system wastes, reflex arcs, feedback loops, and cascading failures to identify corrective actions to improve efficiency.

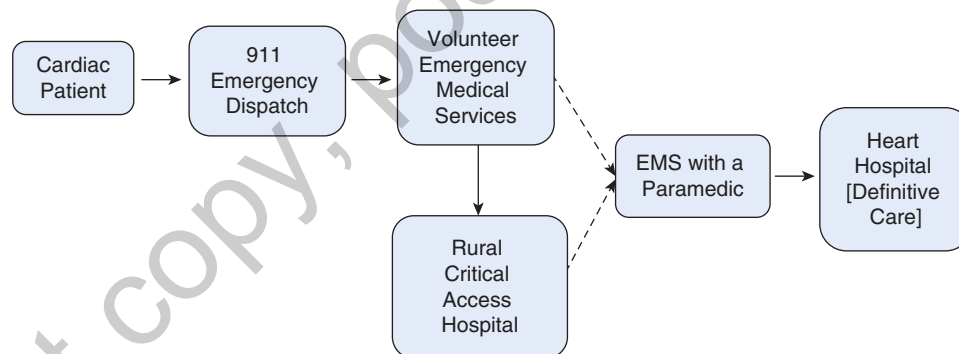
- They found that system waste occurred when the EMTs and the hospital staff both called the same volunteer list to find available drivers. Using a table-top exercise of an evolving cardiac scenario, they identified the overlap and were able to modify procedures.
- Looking at the reflex arc, they identified areas and organization levels used to address issues were higher than necessary. For example, they found that urgent communication was delayed because two different dispatching agencies patched together the responding, volunteer-driven ambulance, and the intercepting ambulance that would take the patient on to the urban hospital. The solution was to ensure that ambulance drivers could communicate with each other directly through a dedicated statewide emergency network.

- They reviewed feedback loops and found that dispatchers, who were law enforcement trained, provided inadequate details about a patient's condition. These individuals now receive training on how to triage the call so that more meaningful information can be passed on to the rest of the system.
- Finally, cascading failures were examined for their impact on the larger system. The simulation had revealed that the heart hospital did not prepare for the patient until they received an EKG. However, the EMS could not send the EKG because there were often connectivity gaps on the road. This resulted in a time delay cascading through the system. The response was to map areas of connectivity on the route so that EMS

personnel could proceed there to transmit the vital patient data (Figure 2.15).

Step 3. Evaluate system effectiveness. Having looked at all these elements working together at different levels, the evaluators could now consider the overall emergent system's effectiveness. They were able to synthesize the results of subsystem interactions and interdependencies to determine the emergent system property of lives saved. By streamlining response coordination, over 250 additional lives were saved in a three-year period. As Renger commented, "When the evaluation approach fits the problem and the context, then evaluation results are more meaningful and will have more utility."

FIGURE 2.15 ■ KEY SUBSYSTEMS



This Spotlight on Equity is based on the contribution of Dr. Ralph Renger.

Chaos Theory

Chaos Theory is a branch of mathematics that looks at dynamical systems highly sensitive to initial conditions. Edward Lorenz was a meteorologist who researched weather and climate predictability. One day when he was running some numbers in a

program simulating weather patterns, he discovered that by rounding off one variable of the 12 he was using in his model (e.g., temperature, wind speed) from 0.506127 to 0.506, this tiny alteration drastically transformed the whole pattern of the prediction (Dizikes, 2011).

This property has been captured in mathematics to denote a sensitive dependence on initial conditions, where small differences in the initial conditions of a chaotic system are persistently magnified because of the dynamics of the system (Lauterborn, 2003). This esoteric mathematical finding eventually resulted in the idea of chaos theory which has changed our classical understanding of nature. Unpredictability and nonlinearity have replaced long-standing ideas of determinism and certainty. Lorenz rather imaginatively suggested that even the flap of a butterfly's wing might cause a tornado, and this idea became known as the Butterfly Effect (Chodos & Ouellette, 2003).

As you will probably continue to check the weather forecast every morning, it is heartening to know that Lorenz's work has led to forecasting improvements (Dizikes, 2011):

wider data collection, better modeling, and “the recognition of chaos” in the weather, leading to what’s called ensemble forecasting. In this technique, forecasters recognize that measurements are imperfect and thus run many simulations starting from slightly different conditions; the features these scenarios share form the basis of a more reliable “consensus” forecast. (para. 26)

Chaotic behavior exists in natural systems like weather and climate but can occur spontaneously, as in road traffic or mob behavior (Safonov et al., 2002). A faucet that drips with an irregular pattern could be called chaotic because the predictability of the next drop is difficult to determine (Shaw, 1984). We recognize chaotic behavior in such different areas as irregular heartbeats, epilepsy disorders, riverbed erosion, career uncertainty, and leadership (Burns, 2002; Dizikes, 2011; Iasemidis & Sackellares, 1996; Kumar & Hegde, 2012; Law et al., 2014). Nevertheless, it is still possible to study chaos. There are underlying patterns, constant feedback loops, repetition, and self-organization within this apparent randomness.

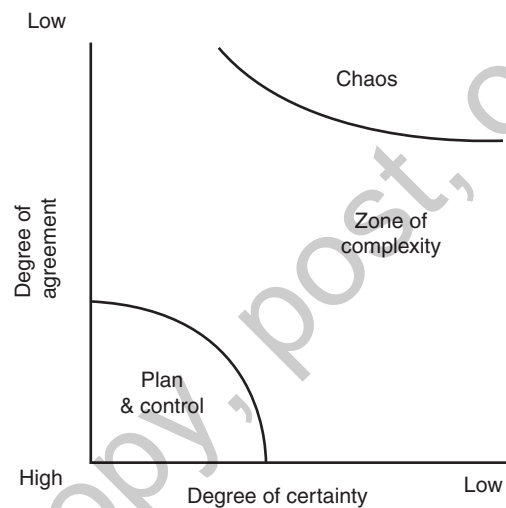
A small defiance on the part of Rosa Parks in Montgomery, Alabama, has often been cited as an example of chaos theory in action (Rettie, 2016). She refused to give up her bus seat to a white man and her subsequent arrest launched a bus boycott involving thousands. Her action became a catalyst that fueled the civil rights movement, led to the desegregation of buses, and raised international awareness about racism in the United States (Čirjak, 2020). She is considered the mother of the civil rights movement. Small changes can indeed have large consequences.

Complexity Theory

An offshoot of chaos theory is **Complexity Theory** which examines unpredictable and nonlinear situations in the fields of management, organizational studies, economics,

and evaluation. To maintain a balance between flexibility and stability, organizations, communities, and groups respond to turbulence and uncertainty by becoming Complex Adaptive Systems where individuals' actions may not be totally predictable yet to make things more complex, their actions are also interconnected with those of others in the organization (Henry, 2014, para. 3). Examples of complexity theory include how bees swarm and how the stock market operates. Concepts like nonlinearity, emergence, dynamic interaction, adaptation, interdependency, and coevolution have developed to understand what is perceived to be unpredictable. Complexity theory explains why plans, aims, and strategies may not work (Henry, 2014; para. 6) and has changed the way we look at organizational behavior, relationships, and interactions. Henry (2014) provides a useful diagram that compares certainty and agreement (Figure 2.16).

FIGURE 2.16 ■ CERTAINTY-AGREEMENT DIAGRAM



Source: Henry, H. (2014, November 3). Complexity theory in nursing. *Independent Nurse*.

Henry describes what happened when a dilapidated community health clinic in Greater Manchester was closed by the Primary Care Trust (PCT) (a branch of the British National Health Service). Health services were moved two miles away and as a result, many residents were unable to access it due to health and transportation issues. Although several alternatives were explored, they failed. The old clinic fell into disrepair, becoming an eyesore and a risk to youth who played on the roof. The residents decided to do something about it themselves. They formed a partnership with service providers, conducted a needs assessment, and held a public meeting. For the first time, residents were able to voice their distress and feelings of being ignored. When the PCT decided to sell the building, the residents approached the community pharmacist who then approached a physician he knew who wanted to open a teaching practice and was looking for a good location.

Together they purchased the clinic site and redeveloped it into a new health center. As Henry comments, “communities on the edge of chaos are fertile ground for complexity theory to work.”

Complexity concepts apply to the social change encountered by evaluators. Patton’s (2011) **developmental evaluation** approach allows the evaluator to use systems thinking to conduct evaluation under conditions of complexity. The evaluator can track emergent and changing realities, illuminate varying perspectives, and feed back meaningful findings in real time to support the dynamics of innovation (p. 7).

Social Science Theory

As well as natural science theories such as these, social sciences, such as anthropology, sociology, psychology, and public health, have developed theories that often influence evaluators’ work. To understand the principles that shape social behavior, evaluators often turn to generalizable and verifiable knowledge found in the social sciences. Donaldson and Lipsey (2006) encourage evaluators to use social science research and empirically based theory to:

- Support initial needs assessment and program design
- Assess the likelihood that a program will accomplish specific objectives
- Guide measurement and design decisions.

For example, deterrence theory, social learning theory, attachment theory, cognitive dissonance theory, social capital theory, public choice theory, and transaction costs theory have all provided explanatory power to evaluation studies. These theories address phenomena related to social programs and the social conditions they intend to improve. They are also helpful in an evaluation context because they can describe practical strategies for program implementation, provide benchmarks for success, offer suggestions about evaluation approaches that others have used, and inform evaluation measurement and design. As Donaldson and Lipsey point out, they can reveal lessons learned about what works or does not work. Evaluators use these theories to measure changes in such areas as knowledge and learning, attitudes and beliefs, and behavior change as well as to understand culture.

Knowledge

One of the earliest approaches (1959) used in evaluation was Kirkpatrick’s Four Levels of Learning Evaluation Model. The four levels of training evaluation are Reaction, Learning, Behavior, and Results (Mind Tools Content Team, n.d.). The theory has been used by human resource professionals and evaluators for many years to determine the effectiveness of training programs.

Let us say a health researcher wanted to learn more about Geographical Information Systems (GIS) to create maps out of data. This researcher enrolled in a GIS course and afterward did the following:

- Reported that the training was engaging, informative, and the instructor was easy to follow and helpful (Reaction)
- Was tested on her knowledge of GIS terms and concepts and received a grade of a 90%. She then demonstrated her GIS mapping skills in a project at work (Learning)
- Held a mini-training program for her colleagues and taught them GIS techniques (Behavior)
- Wrote several grant applications for funding new programs and received all three grants because of her ability to do GIS analysis (Results).

A simple analysis indicated that the Return on Investment for the GIS course was high because the cost of the GIS course was a fraction of what she received in funding (Cost-effectiveness).

Attitude

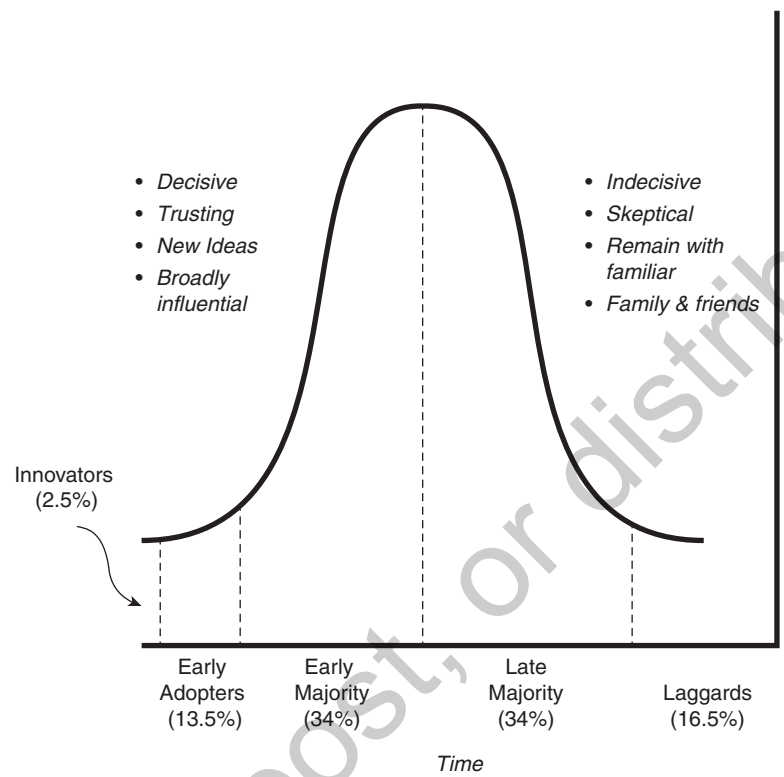
Everett Rogers popularized the Diffusion of Innovation theory in 1962, and it has been frequently used in evaluation. Innovation characteristics, communication channels, time, and social systems all affect the uptake of the innovation (Rogers, 2003). Adopters fall into several categories, depending on the likelihood of their adopting an innovation. They include innovators (i.e., those most likely to adopt the innovation), early adopters, early majority, late majority, and laggards (i.e., those who never will or not likely to adopt the innovation). See [Figure 2.17](#), p. 66 for an illustration of the categories of individuals and the pace at which they adopt interventions.

Valenti and Rogers (1995) studied rural farming practices from the 1920s to the 1940s using the Diffusion of Innovation theory. They found that while hybrid corn was more profitable for the farmer, only a few planted it in the 1920s (i.e., Innovators and Early Adopters). It was not until 1941 that the new strain was adopted by almost the entire community (i.e., Late Majority, Laggards). Through personal interviews with the farmers, they found that there was “considerable farmer reluctance to adopt this highly profitable innovation, a resistance that was very gradually overcome as a farmer talked with his neighbors who already were satisfied adopters” (Valenti & Rogers, 1995, p. 249).

Behavior

A social science theory that is quoted extensively in the behavior change literature is the Transtheoretical Model (Prochaska & Velicer, 1997) (see [Table 2.3](#), p. 66).

FIGURE 2.17 ■ DIFFUSION OF INNOVATION THEORY



Source: Based on Rogers, E. (2003). *Diffusion of innovations* (5th ed.). Simon and Schuster.

TABLE 2.3 ■ STAGES OF THE TRANSTHEORETICAL MODEL OF BEHAVIOR CHANGE

Stage	Definition
1. Precontemplation	Not intending to take action.
2. Contemplation	Intend to take action in the next 6 months. See pros, but focus on cons too and this could cause indecisiveness. Can remain in this phase.
3. Preparation	Intend to take action soon (next month). Have a plan of action.
4. Action	Specific behaviors have occurred in the last 6 months.
5. Maintenance	Trying not to go back to old behaviors. Risk for relapse. Increased confidence.
6. Termination	Zero temptation and 100% self-efficacy.

Source: Prochaska, J., & Velicer, W. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion*, 12(1), 38-48.

An excellent example of behavior change relates to the risk of falling in older people. Zimmerman et al. (2000) designed a fact sheet for physicians, using the Transtheoretical Model (i.e., Stages of Change). Regarding behavior change, the authors state:

Behavior change is seen as a dynamic process involving both cognition and behavior that moves a patient from being uninterested, unaware, or unwilling to make a change (precontemplation); to considering a change (contemplation); to deciding and preparing to make a change (preparation); to changing behavior in the short term (action); and to continuing the new behavior for at least 6 months (maintenance).
(para. 1)

Physicians using the model could measure patients' stage (e.g., precontemplation, contemplation) at the beginning and end of the intervention. Patients' medications could be reviewed to determine if a change was needed. They could do prescribed exercises, change areas in their home that were conducive to falls, and attend a fall prevention program. The physicians were provided a list of conversation topics to use with patients that matched their stage of change. For example, a patient in the precontemplation stage might believe that falling is just part of getting old. The physician could respond with a statement about falling being common but that it can be prevented by simple changes in attitudes and behavior. Another patient, who might be in the preparation stage, might be ready to move to action about preventing falls. The physician could make a referral to a specialist who assists with balance, eye exams, and proper footwear.

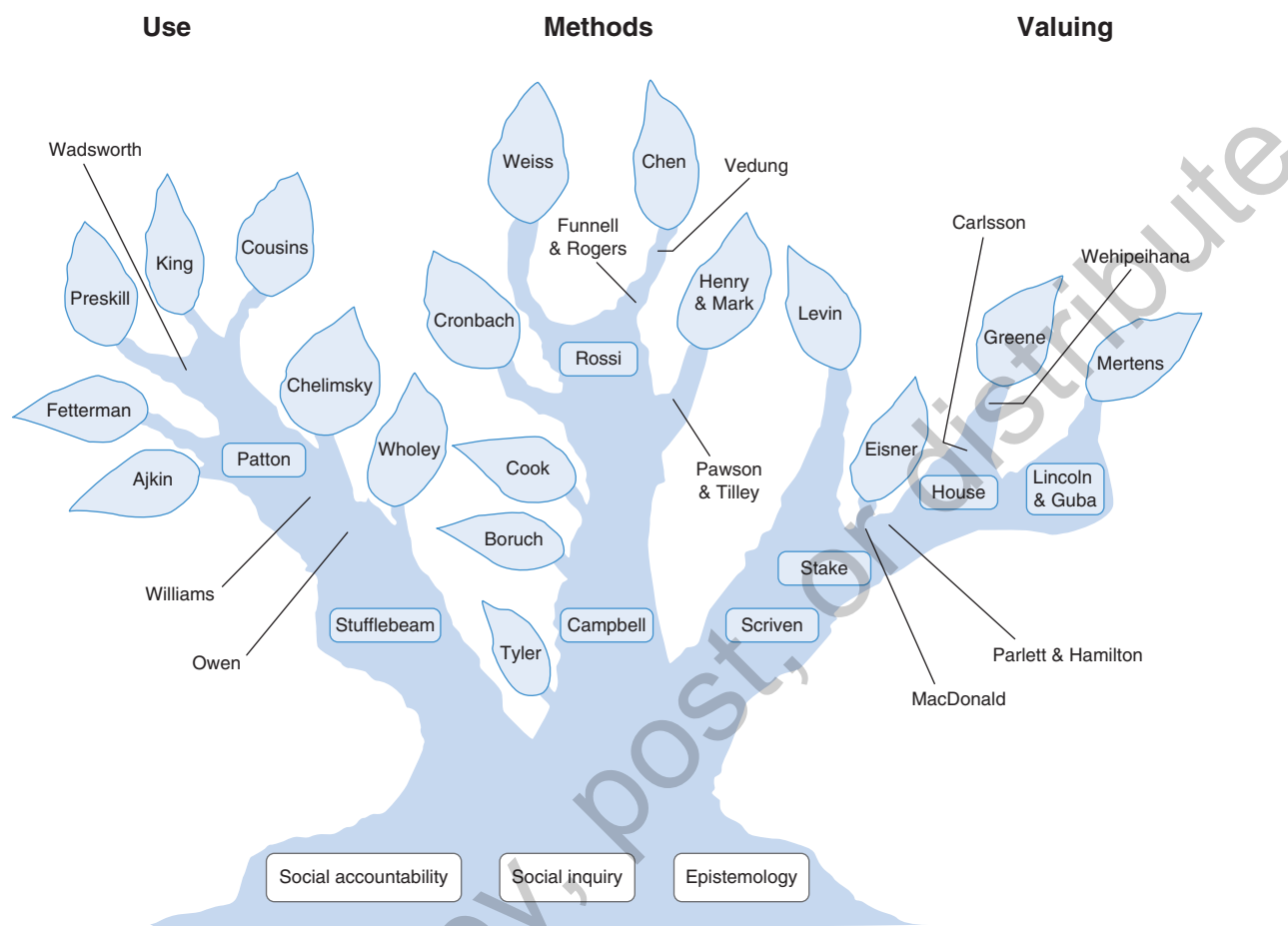
Evaluation Theory

Evaluators have developed many valuable ways to tackle evaluation problems. However, we have a secret to share. There is no “real” evaluation theory! Donaldson and Lipsey (2006) reveal “a confusing mix of concepts related to evaluators’ notions about how evaluation should be practiced, explanatory frameworks for social phenomena drawn from social science, and assumptions about how programs function or are supposed to function” (p. 57). Evaluation theories are approaches, models, or exemplars, that describe what a good evaluation entails from the perspective of the specific theorist. These are not empirical theories (Alkin, 2004, p. 5). They do not predict an outcome (i.e., involve a hypothesis). Indeed, for positivists, they do not meet the precepts of the scientific method. Newcomers to evaluation find the multiplicity of terms littering the evaluation landscape daunting and confusing.

Nevertheless, prominent evaluation writers and thinkers are currently referred to as *theorists* and so we will simply go with common practice.

Shadish (1998) called evaluation theory “central to our professional identity...and perhaps most important, it is what makes us different from other professions” (p. 1). Alkin and Christie created a metaphor to classify these evaluation thinkers and their ideas (Alkin, 2012; Alkin & Christie, 2004). Known as the Evaluation Theory Tree, it has offered students a much-needed way out of the muddle (Figure 2.18, p. 68).

FIGURE 2.18 THE EVALUATION THEORY TREE: USE, METHODS, AND VALUING



Source: Alkin, M. C. (2012). *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed.). SAGE Publications.

In the metaphor, the tree trunk represents the foundation of the evaluation understanding established between an evaluator and a client. The tree's three roots provide the foundation for evaluation work. Social accountability is an important motivator for evaluation, systemic social inquiry reflects the need for sound research methods, and epistemology describes the nature of the knowledge on which the evaluation is based (Alkin, p. 11). The tree has three branches, used to organize the different types of evaluation and the theorists who promote them.

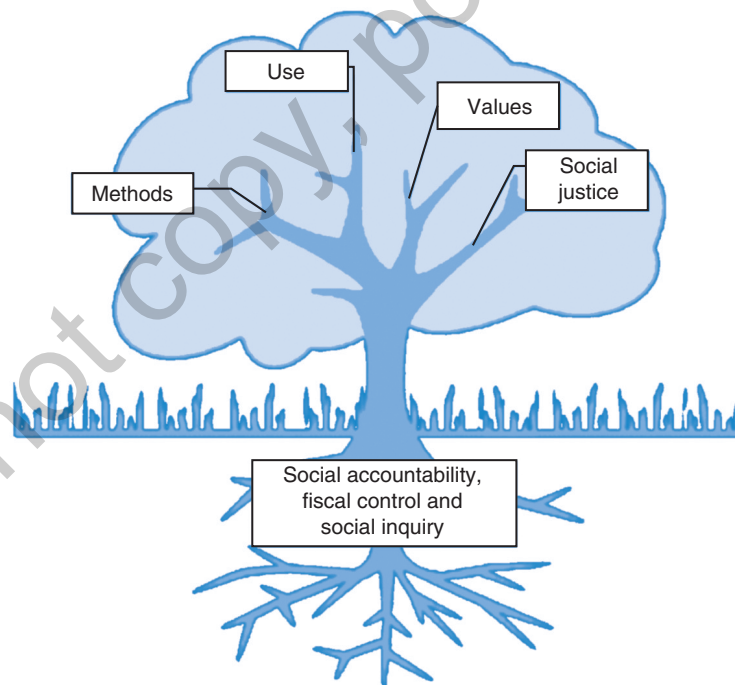
- Use theorists focus on how stakeholders go through decision making and how they are going to use the information. They focus on any appropriate method helpful to stakeholders.
- Methods theorists look at obtaining generalizability and constructing knowledge. They focus on quantitative methods and measurement.

- Values theorists stress the importance of placing value on the data we collect. They use multiple perspectives and focus on qualitative or mixed methods.

One criticism is that this tree primarily reflects the work of White Western evaluation theorists and is not inclusive of “evaluation theorists who are feminists, people of color, persons with disabilities, members of the lesbian/gay/bisexual/transgender/queer or questions (LGBTQ) community, communities in economically poor countries, or members of Indigenous groups” (Mertens & Wilson, 2019, p. 40).

However, the tree metaphor has become a powerful tool for students and practitioners alike and scholars continue to add to its value. For example, Carden and Alkin (2012) modified it to include methods from low- and middle-income countries, including indigenous ones. Recently, Mertens and Wilson (2019) added a critical fourth branch on Social Justice as issues of race, identity, diversity, and marginalization continue to gain prominence. Social Justice theorists have an inclusive perspective that focuses specifically on human rights and the viewpoints of marginalized groups. Typically, they use mixed methods (see Figure 2.19).

FIGURE 2.19 ■ EVALUATION THEORY TREE: METHODS, USE, VALUES, AND SOCIAL JUSTICE



Source: Mertens, D., & Wilson, A. (2019). *Program evaluation theory and practice: A comprehensive guide*. New York, NY: Guilford Press.

Some of the many theorists suggested for the Social Justice branch include those listed in [Table 2.4](#). As awareness grows, the evaluation literature is becoming much more inclusive.

Deliberative Democratic Evaluation	Feminist	LatCrit	Indigenous African	African American/CRT
Barry MacDonald	Kathryn Sielbeck-Bowen	Dolores Delgado Bernal	Bagele Chilisa	Stafford Hood
Saville Kushner	Sharon Brisolera	Lilia Fernandez	John Bewaji	Rodney Hopson
Ernest House	Cynthia Dillard	Tara Yosso		Asa Hilliard
Kenneth Howe	Denise Seigart	Daniel Solórzano		Aaron Brown
Jennifer Greene	Bessa Whitmore	Disability/Deaf Rights	Native American	Leander Boykin
Katherine Ryan	Saumitra SenGupta	Donna Mertens	Marie Battiste	Reid E. Jackson
Human Rights	Sharlene Hesse-Biber	Martin Sullivan	Cheryl Crazy Bull	James Scheurich
Thomas Schwandt	LGBTQ	Carol Gill	Joan LaFrance	Gloria Ladson-Billings
Donna Mertens	Jeffrey Todahl	Raychelle Harris	Maori	Henry Frierson
Marco Segone	Sarah Dodd	Heidi Holmes	Fiona Cram	Veronica Thomas
Karen Kirkhart			Linda T. Smith	

Source: Mertens, D., & Wilson, A. (2019). *Program evaluation theory and practice: A comprehensive guide* (2nd ed.). Guilford Press.

Interestingly, the four branches align with the four paradigms (Mertens & Wilson, 2019) (see [Table 2.5](#)).

Paradigm	Branch of Evaluation Theory Tree
Positivism	Methods
Constructivism	Values
Subjectivism/Transformative	Social Justice
Pragmatism	Use

Many influential evaluation theorists located on these branches are discussed in this book but there are many more, and like real trees, the branches continue to grow and have off-shoots. New theorists emerge all the time.

The Program Evaluation Project

You can see how hundreds of years of philosophy and scientific thought can influence an evaluation project. How we view the world and interpret reality significantly impacts the way we design, collect, and interpret our data. Knowing this background will make you an agile thinker, help you assess your own and other's evaluation thought, and guide you to a strong design for your next evaluation project.

The rest of this book focuses on evaluation projects and how evaluators structure their work. Part 2 examines how evaluators structure their work. It explores how different program phases can be associated with different types of evaluation. In Part 3, we look into what evaluators actually do in their evaluation projects, the types of questions, tools, data collection strategies, data analysis techniques, and reporting and communication processes. Finally, in Part 4, we look at evaluation use in the real world and take on some of the global issues we must incorporate in our work.

Epistemology: Our Knowledge

At last! We have reached the bottom of the Ladder of Evaluation Theory but wait—just a bit more exercise. Climb back up to the top of the ladder, keeping your program evaluation project in mind, and double-check the consistency of your philosophical stance. As with Bessie the Cow, we now see how our project relates to the broader and more abstract philosophical context of evaluation. By understanding our **epistemology**, we can select with confidence the type of evaluation that is needed, including the best design, the most important data to collect, how to analyze it, and how to report it effectively.

Importance of Coherence

Coherence occurs when the parts of something fit together in a “natural or reasonable way” (Cambridge Dictionary Online, n.d. b, para. 2). Mayan (2009) states “Methodological coherence will ensure congruence between your epistemological and ontological viewpoint, your theoretical position/perspective, the method you choose, your research question, and so on” (p. 13).

Evaluation coherence means that the assumptions and choices the evaluator makes regarding purpose, design, methodology, analysis, and interpretation are aligned. Sabarre (2018) argues that:

Many practitioners design evaluations around methodology. However, I argue a more holistic approach starts with theory before methodology. Evaluators should first consider the purpose of the evaluation to determine its theoretical foundation, and then develop evaluation questions to inform methodology. (para. 2)

Evaluators can work from different paradigms, depending on the study, but most tend to be pragmatists, working with the end in mind (i.e., the client's information needs). However, we must avoid the trap of becoming magpies. Although handsome birds, they are known for stealing things, especially shiny objects. As Patton (2002) suggests,

we can pick and choose methods “without giving a thought to their philosophical traditions.”



Source: iStock.com/bestdesigns

Morse (1999) recommends ensuring methodological coherence by doing an armchair walk-through. The evaluator sits quietly, thinking through the methodological trajectory of their study, and reviews their ontology, paradigm, and theory. In other words, they climb down the Ladder within the context of their specific evaluation study and make sure that their epistemological approach fits together.

So, for example, you cannot take a realist perspective, work within the objectivist paradigm, and do intensive interviews with five participants. How could you draw conclusions and arrive at a single “truth”? Each of the five would have their own worldview. You cannot consider yourself a constructivist and then use a prepost standardized survey with a large population as your only research method. Using a rigorous quantitative design, how can you capture the varying lived experiences of this large group? If you are a pragmatist, you may want to use mixed methods, with both quantitative and qualitative components. You will need to follow approved methodological approaches for each component and then follow the mixed methods guidelines for sequencing and data integration. In each paradigm, you need to stay true to the methodological requirements.

The evaluator must confirm the relevance and fit of the choices made. As Sabarre (2018) concludes: “Focusing on theory at the onset of a project ensures the process (i.e., stakeholder involvement, methodology, data collection, analysis, reporting) is intentional, purposeful, and more useful for the client” (para. 2). If you are faithful to the paradigm you have selected, the quality of your work will shine through.



Source: Chris Lysy/FreshSpectrum.com

Critical Thinking and Evaluative Thinking

In this chapter, we introduced the many layers of evaluation thought, theory in action, and the importance of coherence. Evaluators must engage in **critical thinking**. Paul and Elder (2008) state that “a well cultivated critical thinker: raises vital questions and problems, formulating them clearly and precisely; gathers and assesses relevant information, using abstract ideas to interpret it effectively and comes to well-reasoned conclusions and solutions...” (p. 4). Critical thinking helps us appraise complex problems in society, study data, and arrive at workable conclusions.

More specifically, evaluators use **evaluative thinking** which aligns with critical thinking. Buckley et al. (2015) suggest, “**Evaluative thinking** is critical thinking applied in the context of evaluation, motivated by an attitude of inquisitiveness and a belief in the value of evidence, that involves identifying assumptions, posing thoughtful questions, pursuing deeper understanding through reflection and perspective taking, and informing decisions in preparation for action” (p. 378). Preskill and Boyle (2008) explain that it is “about getting people in organizations to look at themselves more critically through disciplined processes of systematic inquiry ... about helping people ask these questions and then go out and seek answers” (p. 148). Evaluative thinking allows us to continue our inquiry, questioning our assumptions about causes and effects, considering potential solutions, and judging the pathways to change.

Now we move on to look at evaluation in action, starting with program logic. First, though, let’s take a moment to meet Dr. Donna Mertens, the evaluation theorist and researcher who added the social justice branch to the Evaluation Theory Tree.

EXPERT CORNER

Dr. Donna Mertens



Donna Mertens, Professor Emeritus, Gallaudet University, specializes in research and evaluation methodologies designed to support social, economic, and environmental transformation. Her most recent books include *Mixed Methods Design in Evaluation*, *Program Evaluation Theory and Practice*, and *Research and Evaluation in Education and Psychology*. She is a past President of AEA and a founding Board member of the International Organization for Cooperation in Evaluation (IOCE).

1. *What caused you to update the Evaluation Theory Tree with a new branch on social justice? To what extent has it been accepted so far?*

I have worked with members of many marginalized communities, e.g., deaf people, Indigenous communities, underrepresented racial/ethnic groups, and women. The consistent message they shared was that they were either not

being represented in evaluation studies, or they were inaccurately represented. Given the dire consequences of inequities in society, such as lack of access to health care, education, and meaningful employment, I think evaluators have a moral responsibility to design their work so that it consciously addresses issues of human rights and social justice. The Evaluation Theory tree included the branch of values; however, there was nothing explicit about the value of social justice. My hypothesis is that if we explicitly position ourselves as supporting the furtherance of social justice, then we are more likely to see our results used for that purpose. Thus, I added that branch and it has been accepted by many evaluators who share this value.

2. *Transformative evaluation is a term that is becoming widely used in the context of working with marginalized people, particularly in the international development community. How applicable is it in North America?*

The international community is abuzz with the word “transformation,” especially as it applies to the achievement of the UN’s SDGs or Sustainable Development Goals (United Nations Development Goals, n.d.). I have recommended the application of a transformative lens to evaluations that are used to address the societal transformations needed to achieve the SDGs. North America is not immune to the presence of injustices. We need only to look at headlines in the newspapers to see the incidences of violence against immigrant groups, shooting of unarmed Black men, inequities in resources in many communities of color, lack of access for people who have a disability or are deaf, or the disappearance of Indigenous peoples to convince us that we live in an unequal world. I have worked with communities in 85 countries; I can assure you that the transformative lens for evaluation has applicability in all countries, including North America.

3. *What evaluation strategies would you recommend that are inclusive and supportive of all voices in an evaluation context?*

I recommend a multistage mixed methods design that is informed by the assumptions of the transformative paradigm if we are to legitimately include and support all voices in evaluation studies. This begins with the identification of relevant stakeholders in the particular context. Evaluators should start with such questions as: Who is included? Who is excluded? What needs to be done to respectfully include all stakeholders? What are the cultural issues of relevance in this context? How can we build relationships with the full range of stakeholders that are culturally respectful? These should be followed by a contextual analysis to identify the cultural, historical, and political factors that are relevant. Based on this analysis, interventions and evaluation strategies that are culturally responsive can be developed and tested. Evaluators should examine the effectiveness of programs, as well as the quality of relationships as that will determine the sustainability of an intervention.

4. *Finally, do you have any words of wisdom for those of our readers who are just beginning to explore the field of evaluation?*

Ask yourself this question: “When I leave this world, will I look back on my contributions and say I have contributed to making a more just world?” While communication with our wider community is essential to grow our understanding of how to conduct better evaluation, the inclusion of a transformative lens has the potential to increase all our efforts to create a more just world.

Key Terms

Chaos theory 61	Evaluative thinking 73	Realism 42
Coherence 71	Ontology 41	Relativism 42
Complexity theory 62	Paradigm shift 43	Scientific method 46
Constructivism 49	Paradigms 37	Subjectivism 51
Critical thinking 73	Positivism 44	Systems theory 55
Developmental evaluation 64	Postpositivism 45	
Epistemology 71	Pragmatism 53	

The Main Ideas

- Using the Ladder of Abstraction, we begin with ideas at the lowest level of abstraction (concrete) and move upward to ideas that are increasingly abstract.
- An evaluator must understand the theoretical principles that underpin their study, or their work may be compromised.
- What we know about the world is our ontology. Are you a realist or a relativist?
- Paradigms set the foundation for different approaches to evaluation. They help evaluators understand their perspective on reality: positivism/postpositivism (one reality); constructivism (multiple realities); subjectivism/the transformative paradigm (meaning lies within the individual); and pragmatism (it depends).
- The socio-ecological model provides a framework to study the different layers of influence on a program, including the individual, interpersonal groups, institutions, community, and public policy. Evaluators work across these levels to identify barriers and facilitators and to determine outcomes.
- The Evaluation Theory Tree separates evaluation theorists into various categories depending on their perspective. The branches include use, methods, valuing, and social justice.
- Evaluation aims for philosophical coherence. Our perspective reveals the assumptions and the choices we make regarding purpose, design, methodology, analysis, and interpretation.
- Evaluators should think critically, continually ask questions of hard problems, and seek solutions that are not readily apparent. Critical thinking in an evaluative context is known as evaluative thinking.

Critical Thinking Questions

1. Which of the paradigms most closely aligns with your beliefs? For example, are you a positivist, objectivist, constructivist, subjectivist, or pragmatist? Do you find yourself in between two of the paradigms or solidly within one of them?
2. Think about the field in which you hope to practice one day. What paradigm is prevalent in this field? Do you agree or disagree with this focus? How does this paradigm impact the types of evaluations conducted in your field? To what extent do they have a positive impact?
3. Can you remember a time in your life when things did not go as planned and became, instead, chaotic, and unpredictable? What was your response? What would you do differently next time?

Student Challenges

1. **Critically Thinking about your Life.** A well-rounded evaluator thinks critically about their life. You would raise questions about challenges you have, gather information to support a needed change, and arrive at a conclusion about the needed direction that is warranted. If you could change one thing about your life that would exponentially improve your life, what would that be? Identify the one thing you want to change and write a paragraph providing details. When does it happen? Why does it occur? Who is involved? Where does it occur? What precipitates it? What happens after it? What small intervention or strategy, if done, might change the outcome? Try a strategy. Reflect on your selection and write one sentence each day to explain your thinking for 30 days. What outcomes have you seen?
 2. **Views of Theorists.** Select one of the theorists mentioned in this chapter. Find their three most recent articles and read them. Identify their important ideas. Do your views align with theirs? Why or why not?
 3. **Realist and Relativist Approaches.** Find two articles describing program evaluations in a program area you know well, one using a realist perspective and quantitative methods and the other using a relativist perspective and qualitative or mixed methods.
- Your Tasks:
- a. Describe the methods used in each article.
 - b. Observe any differences about the types of conclusions drawn in each one.
 - c. Decide if a qualitative approach would have yielded different conclusions in the realist article and, similarly, if a quantitative approach for the relativist topic would have elicited different conclusions as well.
 - d. Determine if you can draw any conclusions from this comparison.
- Evaluation Journal Examples:
- *The Canadian Journal of Program Evaluation (free access to all their issues)*
 - *New Directions in Evaluation*
 - *The American Journal of Evaluation*
 - *Evaluation and Program Planning*
 - *Evaluation & the Health Professions*
 - *Assessment & Evaluation in Higher Education*

Additional Readings and Resources

1. Buescher, T. (1984). Thinking through the evaluation process: An interview with Dr. Joseph Renzulli. *Journal for the Education of the Gifted*, 7(1), 3–11. Several terms from Chapter 2 are used by this author, including pragmatist, evaluation plan, reductionist, evaluation questions.
 2. Connor, T. (2018, October 10). *Deductive, inductive, and abductive reasoning*. <https://medium.com/10x-curiosity/deductive-inductive-and-abductive-reasoning-c508e6b43097>
Connor has a well-written blog post on deductive, inductive, and abductive reasoning. He posits that to influence others, you need to design your arguments to make a more straightforward case for your position.
 3. Dlugan, A. (September 15, 2013). *The ladder of abstraction and the public speaker*. <http://sixminutes.dlugan.com/ladder-abstraction/>
This website guides you to becoming an effective speaker and using the ladder of abstraction to frame your presentation delivery.
 4. Jones, M., Verity, F., Warin, M., Ratcliffe, J., Cobiac, L., Swinburn, B., & Cargo, M. (2016). OPALesence: Epistemological pluralism in the evaluation of a systems-wide childhood obesity prevention program. *Evaluation*, 22(1), 29–48.
The authors explain how they apply epistemology to the evaluation research of an obesity prevention program. They look at the intervention at the individual, family, community, and policy levels.
 5. McLeroy, K., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly*, 15(4), 351–377.
The authors underscore the importance of visualizing interventions and changes within society through an ecological lens. The article outlines methods for focusing on intrapersonal factors, interpersonal processes, institutional factors, community factors, and public policy.
- Not all programs should be based on how individuals can change but on the best intervention strategies shown through evidence to cause change.
6. Moon, K., & Blackman, D. (2014). A guide to understanding social science research for natural scientists. *Conservation Biology*, 28, 1167–1177. The authors provide a link from the natural scientists to the research world indicating that conservation problems are commonly social problem (p. 1). They discuss the ontological and epistemological positions of the natural scientist and the way these positions can influence how they set up their designs, methodology, interpretation, and conclusions.
 7. Mowles, C. (2014). Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship. *Evaluation*, 20(2), 160–175. The author reviews simple, complicated, and complex problems and addresses when best to choose complexity in a program evaluation.
 8. Esbensen, F., & Matsuda N. (n.d.). *Changing course: Preventing gang membership* (T. Simon, N. Ritter, R. Mahendra, eds.). Office of Justice Programs and Centers for Disease Control and Prevention. <https://www.ojp.gov/pdffiles1/nij/239234.pdf>
This book was a joint venture by the Department of Justice and the Centers for Disease Control and Prevention. It details why gang membership is important to prevent and the attraction for youth to gangs, aligns the attraction to gangs and the prevention of them to the socio-ecological model, and provides a wealth of information as to how individual, relationships, community, and societal factors either promote or discourage joining gangs. The authors give information about what interventions should be employed and how the identification process of gang membership works.

Do not copy, post, or distribute