Social Work 3401 Coursebook

Social Work 3401 Coursebook

AUDREY BEGUN

Contents

Module 3

Module 4

Module 5

Preface & Orientation

Welcome to the online coursebook for SWK 3401 *Research & Statistics for Understanding Social Work Problems and Diverse Populations* course. There are several tools built into these chapters, designed to enrich and guide your learning.

- Material may be reviewed in three different formats: online (EPUB), downloaded (PDF), or as Mobilepocket (MOBI) format, which is better suited to small-screen devices.
- Embedded interactive "Stop and Think" exercises require internet connectivity, but each can also be downloaded for offline work. These exercises are presented to help you apply what you are reading, challenge yourself, prepare for quizzes, and have a little fun along the way.
- A list of key terms is presented at the end of the module coursebook to explain text appearing as ***highlighted in bold italics*** throughout the coursebook–in the interactive mode you can click on a highlighted word to jump to its explanation in the key terms section. Use the back arrow to return to where you were reading.
- **"Back to Basics"** textboxes appear in some chapters of some modules. These are designed as refreshers of some concepts that you may have learned in the past and that have applicability to the present material.
- **Green highlighted** boxes appearing in some chapters/modules refer you to learning activities in your online interactive Excel Workbook. These learning activities relate to content introduced in your main coursebook. The intent is for you to complete each Excel Workbook activity as you encounter it, then return to the coursebook.

Should you find it necessary to cite portions of the coursebook, here is the recommended citation: Begun, A.L. (2018). SWK 3401 *online coursebook: Research & statistics for understanding social work problems and diverse populations*. The Ohio State University, Open Educational Resources.

MODULE 1

Module 1 Introduction

Our first module allows us to become oriented to the importance of research and evidence in social work, as well as ethical practices in social work research. Exploring these topics prepares you to become engaged in locating, analyzing, and using evidence to inform your social work practice and to become engaged in social work research efforts, as well. Additionally, we borrowed a quote from Bannon et al. (2015) to inform development of the coursebook materials:

> *"The class operates upon the belief that learning is a product of performing quantitative analysis, not just reading about the process"* (p. 3).

Thus, throughout our online coursebooks we work with files and learning activities involving the Excel software program. These are noted in green boxes throughout the coursebooks; each green box refers you to a specific location in the online, interactive Excel Workbook.

Reading Objectives

After engaging with the Module 1 reading materials and learning resources, you should be able to:

- Recognize and analyze different types of knowledge that inform social workers' understanding of social work problems, diverse populations, and social phenomena;
- Identify principles of critical thinking;
- Distinguish between evidence from science, emerging science, pseudoscience, and opinion;
- Describethe roles of research in social work and in promoting

social justice;

- Identify the philosophical roots of different research methodologies that inform social workers' understanding of social work problems, diverse populations, and social phenomena;
- Explain important relationships between theory, research, and practice in understanding diverse populations, social work problems, and social phenomena;
- Analyze important human subject research principles and the ethical conduct of research;
- Locate sources of empirical evidence;
- Define key terms related to social work research and research ethics

Module 1 Chapter 1: Social Work and "Ways of Knowing"

This chapter is largely about social work knowledge—how social work professionals come to "know" what we know, and how that knowledge can be leveraged to inform practice. Thinking about what we know and how we come to know it is critically important to understanding social work values, beliefs, and practices. This means critically thinking about the sources and types of knowledge on which social workers rely, and the implications of relying on these different sources and types of knowledge.

What you will learn from reading this chapter is:

- The philosophical roots of different approaches to scientific knowledge on which social workers rely (epistemologies);
- Different types and sources of knowledge ("ways of knowing");
- Strengths and limitations of various "ways of knowing" social workers encounter and might apply in understanding social problems and diverse populations;
- Principles of critical thinking;
- Distinctions between science, pseudoscience, and opinions.

Thinking about Knowledge

The study of knowledge and knowing about the world around us called ***epistemology***, and represents one of the major branches of philosophy. Throughout much of human history, philosophers have dedicated a great deal of thought to understanding knowledge and its role in the human experience. Ancient Greek and western philosophers like Socrates, Plato, Aristotle, as well as those who came before and after them, made lasting contributions to the way we think about knowledge and its applications in daily life. A great deal of this western philosophical activity centered around building knowledge based on facts and provable "truths," rather than spirituality, mythology, and religion. These philosophical efforts contributed greatly to the evolution of rational thought, science, theory, and scientific methods that we use in our everyday and professional lives to make sense of the world around us. Much of the science on which we often rely to find answers to perplexing questions is derived from a philosophical position called ***positivism***. Positivism essentially involves adopting a stance where valid knowledge, or "truth," is based on systematic scientific evidence and proof: in order to be positive about knowing something, that something must be proven through scientific evidence. Scientific evidence, developed through a positivism lens, results from a rational, logically planned process identified as the scientific process. Similarly, ***empiricism*** is about proven knowledge, but contends that proof also may come through the application of logic or through direct observable evidence. In our SWK 3401 and SWK 3402 coursework you will be exposed to many of the methods that investigators apply in scientifically answering questions about social problems, diverse populations, social phenomena, and social work interventions. But science is not the only way of knowing, and it is important for social workers to understand the place science occupies in an array of ways knowledge is developed and applied.

Much of what is taught in the United States about knowledge and epistemology is somewhat constrained by traditions of western philosophy. Challenges to these constraints emerged in the literature during the 1990s under the headings of naturalistic epistemology, anthroepistemology, and ***ethnoepistemology*** (e.g., Maffie, 1990; 1995). An important contribution to our understanding of knowledge is an anthropological appreciation that knowledge is constructed within a cultural context. This is quite different from the positivist perspective concerning single, provable truths that are waiting to be discovered.

> "*Anthroepistemology views epistemology as an historically and contingently constituted phenomenon, the nature, aims, and province of which are to be understood in terms of the life context in which epistemology is organically rooted and sustained rather than in terms of divine imperative, rationality per se, or pre-existing epistemic facts or principles*" (Maffie, 1995, p. 223).

In other words, what we know, how we come to know it, and how we think about knowing all are influenced by the personal, historical, and cultural contexts surrounding our experiences. For example, consider what you "know" about deafness as a disability. This deficiency perspective comes from the cultural context of having lived in a hearing world. However, members of the Deaf community offer a different perspective: living within deaf culture and linguistic structures (e.g., using American Sign Language to communicate) conveys specific social and cultural implications for human development, behavior, thinking, and worldview (Jones, 2002). These implications are viewed the same way other cultures are viewed—as cultural differences when being compared, not as deficiencies, or "otherness." This diversity of experience perspective (rather than a disability perspective) could be applied to other topics, such as autism, reflecting human neurodiversity rather than a disorder.

Implications for Social Work.

One implication of this observation might be that social workers should expect worldviews held by individuals with vastly different lived experiences to differ, too. Well over 100 years ago, the philosopher and psychologist William James (1902), discussing the varieties of religious experience, observed:

> "*Does it not appear as if one who lived more habitually on one side of the pain-threshold might need a different sort of religion from one who habitually lived on the other?*" (p. 105).

In broader terms, social workers strive to understand diversity in its many forms. We appreciate that diverse life contexts, experiences, opportunities, and biology all interact in complex ways in contributing to diverse outcomes. These differences include differences in knowledge and understanding of the world—differences, not deficiencies. For example, we understand that the elements of a parenting education program delivered to two-parent, well-resourced, privileged, reasonably empowered families of one racial or ethnic background may be grossly inappropriate for families existing in a far different reality, facing very different challenges, very differently resourced or privileged, and responding to different experiences of discrimination, oppression, exploitation, threats of violence, and micro-aggression. The pre-existing knowledge different families bring to their parenting situations and their parenting knowledge needs differ significantly—even their parenting goals, approaches, and means of learning and developing knowledge may differ.

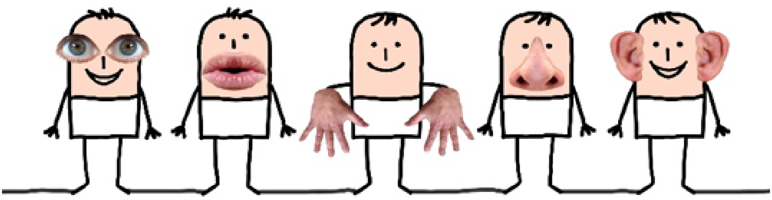
Social workers who adopt an anthropological, ethnoepistemology perspective are open to considering the beliefs of ordinary people around the world alongside those of leaders, academics, scientists, colleagues, and authorities. This line of thought encourages us to reflect on all epistemologies—wherever in the world they are practiced and by whomever they are practiced. Western philosophy, rational logic, and science are situated within this context—as one of the multitude of epistemologies that exist, not as the first or the most significant, but simply as one of many. These perspectives contribute to social work having a rich array of research methodologies available for understanding social problems, diverse populations, and social phenomena. A positivism/empiricism perspective contributes to many of our ***quantitative research*** methodologies; an anthropological, ethnoepistemology perspective underlies many ***qualitative research*** methods.

This philosophical background is relevant to social work education for several reasons:

1. Social work professionals engage with individuals from many diverse backgrounds and social contexts. This means that we encounter many different ways of knowing and great diversity of belief among the people with whom we interact. Colleagues, professionals from other disciplines, clients, community members, agencies, policy decision makers and others all have their own understanding about the world we live in, and these understandings often differ in critically significant ways. We cannot work effectively with others if we do not have an appreciation for this diversity of understanding, thinking, and believing.
2. To be effective in our interactions, social workers also need to understand and critically evaluate our own personal and professional epistemologies, and what we "know" about social work problems, diverse populations, and social phenomena. This idea fits into the social work practice mandate to "know yourself" (e.g., see Birkenmaier & Berg-Wegner, 2017).
3. Significant differences exist between the philosophies that underlie different scientific methodologies. Rational logic underlies quantitative methodologies and ethnoepistemology underlies on qualitative methodologies. These differences contribute to complexity around the qualitative, quantitative, and mixed methods (integration of qualitative and quantitative) choices made by investigators whose work helps us understand social work problems, diverse populations, and social phenomena.

Thinking about Different Sources and Types of Knowledge

Humans have many different ways of developing our knowledge about the world around us. Think about how young children come to understand the world and all kinds of phenomena they experience. First, they utilize all five senses to explore the world: vision, hearing, smell, taste, and touch. For example, a baby might hear food being prepared, see and smell the food, touch it and taste it.



Sensory evidence provides people of all ages with a great deal of experiential knowledge about the physical and social world.

Second, people engage in internal mental operations we call ***cognition***—they engage in thinking and problem solving—to create meaning from their experiences. To continue our example, it is through these internal mental experiences that babies develop their ideas about food. Some of their conclusions are effective, while their other conclusions are inaccurate; these guesses need to be revised through further experience and cognition. For example, this toddler learned (the hard way) to understand that paint is not food. These cognitions may take a few trials to accurately emerge.

Third, knowledge comes through basic learning principles involving reinforcement and punishment of behavior, as well as observational learning of others' behaviors. Young children often conduct repeated experiments to develop knowledge concerning the basic principles by which the physical and social world operate. For example, a young child might experiment with using "naughty" swear words under different conditions. The first few times, it might simply be a case of copying a role model (such as parents, siblings, peers, or television/movie/music performers). Based on the way the social world responds to these experimental uses of language, the child may continue to experiment with using these words under different circumstances. This behavior might get one response from siblings or other children (laughing, giggling, and “Oooohhhhh!” responses), a different response from a parent or teacher (a corrective message or scolding response), and yet a different response still from a harried caregiver (simply ignoring the behavior). We may consider this child to be an “organic” scientist, naturally developing a complex understanding about the way the social world works.

People of all ages rely on these multiple ways of knowing about the world—experiential, cognitive, and experimental. What additional ways of knowing are important for social workers to understand? Answers to this question can help us better understand the diverse people with whom we engage and can help us better understand our own sources of knowledge as applied in professional practice. By the way, thinking about our thinking is called ***metacognition***—having an awareness of your own processes of thinking and knowing.

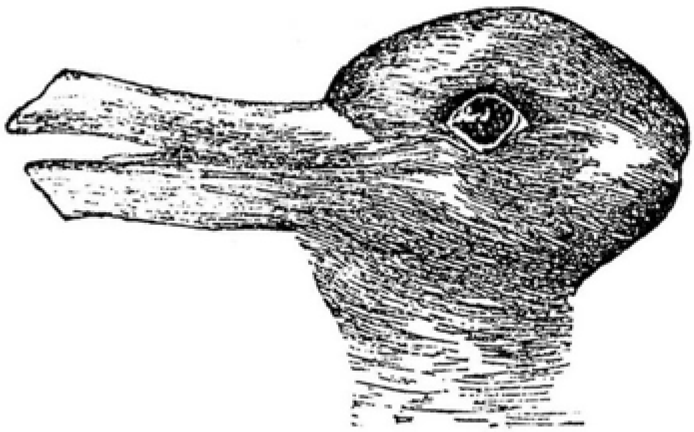
Ways of Knowing.

Textbooks about the Theory of Knowledge offer various ways through which human beings derive knowledge (see for example IB, n.d.; TOK network, n.d). These ways of knowing include:

- Sensory perception
- Memory
- Language
- Reason
- Emotion
- Faith
- Imagination and Intuition

Let's examine each of these ways of knowing in a little more detail, and consider their implications for knowledge in social work. It is helpful to think about how we might consider our use of each source of knowledge to avoid running "off into a ditch of wrong conclusions" (Schick & Vaughn, 2010, p. 4).

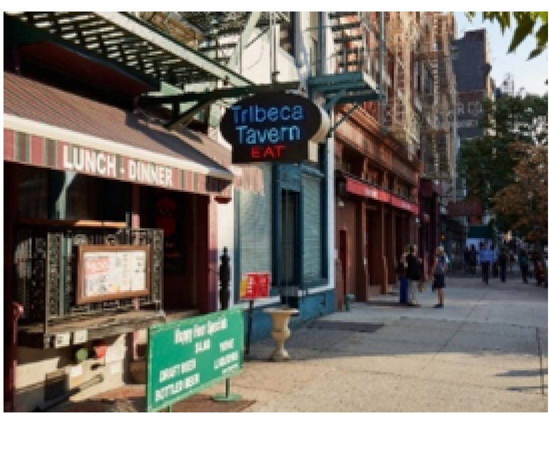
Sensory perception and selective attention. We have already considered the ways that individuals across the lifespan come to know about the world through engaging their five senses. In general, human beings rely on information acquired through their senses as being reliable evidence. While this is generally appropriate, there are caveats to consider. First, knowledge developed through our sensory experiences are influenced by the degree to which our sensory and neurological systems are intact and functional. There exists a great degree of variability in different individuals' vision, for example. We can reduce the variability to some degree with adaptive aides like eye glasses, contact lenses, adaptive screen readers, and tools for sensing color by persons who are "color blind." But, it is important to recognize that perception is highly individualistic. You may have played with optical illusions in the past .Or, you may recall the 2015 social media storm created by the "What color is this dress?" dilemma (#thedress, #whiteandgold, and #blackandblue). Or, for example, when you look at this image, what do you see?



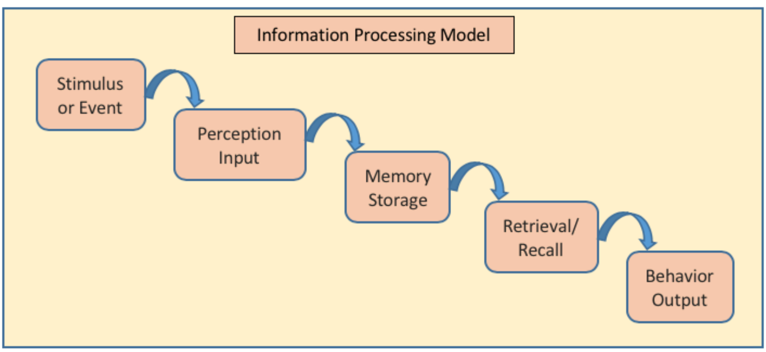
You might see a duck head looking to the left, a rabbit head looking to the right, both, or neither. In both the dress and duck/rabbit examples, different individuals perceive the same stimuli in different ways.

This raises the question: can we trust our own perceptions to represent truth? In our everyday lives, we rely heavily on our perceptions without thinking about them at all. Occasionally, we might experience the thought: "I can't believe my own eyes" or "I can't believe what I am hearing." Between these extremes, social workers often experience the need to confirm our perceptions and checking with others about theirs.

One reason why our perceptions might not match the perceptions of another person has to do with what we each pay attention to. Everyone practices a degree of selective attention . For example, as you are reading this material, are you aware of the feel of the room temperature, the sound of the air circulation, or the amount of light in the room? This information is all available to you, and you might have paid attention to them when you first sat down to read. But, you stopped being consciously aware of them as they remained steady and irrelevant to the task of reading. If any conditions changed, you might have noticed them again. But, it is human nature to ignore a great deal of the information collected by our sensory systems as being irrelevant. A source of individual difference is the degree to which they attend to different types of information. For example, as you ride through a neighborhood you might notice the location of different food or coffee establishments. Someone in recovery from an alcohol use disorder, taking the very same ride, might attend to the different liquor advertisements that could trigger a craving and warrant avoiding. You both experienced the same environment, but through selective attention you each came to know it differently.



Memory and memory bias. You might argue that memory is not really a way of knowing, that it is only a tool in the process of knowing. You would be partially correct in this argument—it is a tool in the process. However, it is not ONLY a tool. Memory is an integral and dynamic part of the human mind's information processing system. The diagram below represents steps in human information processing. The process begins with the sensory systems discussed earlier. But, look at what happens to those sensory perceptions next:



In terms of the information processing model, once something has been perceived or an event has been experienced by an individual (perception), the perception moves to the memory system. A perception first moves into short-term memory as something the person is actively aware of or thinking about. The short-term memory "buffer" is typically limited in space to around 5-7 items at a time and lasts for only about 18-30 seconds. After that, the memory either is cleared from the "buffer" zone and forgotten forever, or it is moved to longer term memory. If a memory is not converted into a long-term memory, it cannot be retrieved later.

The long-term memory system has a tremendous storage capacity and can store lasting memories. However, stored memories are relatively dynamic, rather than static, in nature. Memories are remarkably susceptible to change through the incorporation of new information. Human memory does not work like a video camera that can be replayed to show exactly what was originally experienced. Instead, recalling or retrieving memory for an event or experience happens through a process of recreation. In this process, other factors and information become woven into the result, making it a less-than-perfect depiction of the original experience. Consider research about the ways that additional (mis)information affects what individuals recalled about a story presented to them in a memory study (Wagner & Skowronski, 2017). Study participants reported "remembering" new facts presented later as having been part of the original story—these later facts became false memories of the original experience. False or erroneous memories were more likely when the original facts were scarce (versus robust), the new information seemed true (rather than questionably reliable), and the person was highly motivated to retell the story. These conditions provoked the human mind to fill in gaps with logical information.

The nature of a memory also may be (re)shaped by internally generated cognition based on our pre-existing understanding of the world—our thoughts. This process makes memories more relevant to the individual person but is a significant source of memory bias. We often use our own thoughts, ideas, cognitions, and other experiences to fill in knowledge gaps around what we remember from direct experiences. For example, you may have learned about the poor reliability of eyewitness testimony—what is remembered by the eyewitness is influenced by how the situation is understood, making it susceptible to distortion from beliefs, stereotypes, prejudice, and past experiences. This explains why convenience stores often have rulers painted on the doorframe: if a crime occurs, a shorter witness might recall a person's height as being tall, and a taller witness might recall the person's height as being short. The yardstick allows more objective, undistorted recall of the facts.

Memory reshapes itself over time, and sometimes people lose the connections needed to retrieve a memory. As a result, what we "think" we know and remember is influenced by other external and internal information.

Language. Like memory, language is both a product and producer of knowledge. Language is a created set of signs and symbols with meaning that exists because of conventional agreement on that meaning. Language helps describe what we know about the world, both for ourselves and to communicate with others. But, language also has the power to shape our understanding, as well. ***Linguistic***

determinism is about how the structures of language constrain or place boundaries on human knowledge.

Consider, for example, the limitations imposed on our understanding of the social world when language was limited to the binary categories of male and female used in describing gender. Our language being structured in this binary fashion leads individuals to think about gender only in this manner. Conventional agreement is extending to include terms like agender, androgynous, bigender, binary, cisgender, gender expansive, gender fluid, gender non-conforming, gender questioning, genderqueer, misgender, non-binary, passing, queer, transgender, transitioning, transsexual person, two-spirit (see Adams, 2017). With the addition of multiple descriptive terms to our language, it becomes easier to have a more contextual, complex, non-binary understanding of the diverse nature of gender. Language is one mechanism by which culture has a powerful influence on knowledge.

Reason and logic. Humans may come to "know" something through their own internal cognitive processing and introspection about it. Logic and reasoning represent thought processes that can result in knowledge. You may recall the important piece of reasoned logic offered by Rene Descartes (1637): “I think therefore I am.” This piece of reasoning was the first assumption underlying his building of an orderly system of “truth” where one conclusion could be deduced from other conclusions: ***deductive reasoning***. In deductive reasoning, we develop a specific conclusion out of assembling a set of general, truthful principles (see Table 1-1). The specific conclusion is only as good as the general principles on which it is based—if any of the underlying premises are false, the conclusion may be incorrect, as well. The conclusion might also be incorrect if the premises are assembled incorrectly. On the other hand, ***inductive reasoning*** involves assembling a set of specific observations or premises into a general conclusion or generalized principle (see Table 1-1). Inductive reasoning is a case of applying cause and effect logic where a conclusion is based on a series of observations. For example, estimating future population numbers on the basis of the trend observed from a series of past population numbers (Kahn Academy example from www.youtube.com/watch?time_continue=43&v=GEId0GonOZM). ***Abductive reasoning*** is not entirely inductive or deductive in nature. It generates a hypothesis based on a set of incomplete observations (inductive reasoning) and then proceeds to examine that hypothesis through deductive logic.

Table 1. Examples of deductive, inductive, and abductive reasoning from Live Science https://www.livescience.com/21569-deduction-vs-induction.html

Deductive Reasoning Example	Inductive Reasoning Example
Spiders have 8 legs. A black widow is a spider. Conclusion: black widows have 8 legs	Pulling a series of coins from a coin purse, 1st, 2nd, 3rd, and 4th is each a penny. Conclusion: the 5th will be a penny, too.
Abductive Reasoning Example	
Entering your room, you find that your loose papers are all over the place. You see that the window is open. Conclusion: the wind blew the papers around. But, you consider alternative explanations (another person came in and threw the papers around) and settle on the most probable/likely conclusion.	

A person does not actually need to witness or experience an event to draw logically reasoned conclusions about what is likely. For example, a social worker may be trained in the knowledge that persons with traumatic brain injury (TBI) in their histories have a greater likelihood than the general population to experience problems in family relationships, unemployment, social isolation, homelessness, incarceration, depression, and suicide risk (Mantell, et al, 2017). In meeting a client with a TBI history, the social worker might reason that it is important to screen for the risk of each of these problems. The social worker does not need to wait for these problems to be observed; the knowledge gained through reasoning and logic is sufficient to warrant paying attention to the possibility of their being present or emerging.

Sometimes a line of reasoning seems rational and logical, but when carefully analyzed, you see that it is not logically valid—it is fallible, or capable of being in error. A ***fallacious argument*** is an unsound argument characterized by faulty reasoning. Fallacious arguments or flawed logic can result in the development of mistaken beliefs, misconceptions, and erroneous knowledge. For example, in the 1940s to 1950s a mistaken theory of autism was offered in the professional literature (see Rimland, 1964): the “refrigerator mother” theory was that the low levels of warmth parents were observed directing to an autistic child caused the child’s autism symptoms. The logic flaw in this case was a very common error: drawing a false causal conclusion from a correlational observation. The observed parental behavior actually was a sensitive response to behavioral cues presented by the children with autism, not a cause of the autism. Thus, the “refrigerator mother” misconception was influenced by logical fallacy. Clearly, the knowledge gained through reasoning is only as good as the weakest link in the chain of ideas and logic applied.

Emotion and affect. Human feelings have a powerful influence on what we think and understand about the world around us. Another word for emotion is affect (not to be confused with effect, pronounced with an “a as in apple”). For example, the advertising, entertainment, and news media have long recognized that individuals more strongly remember information to which they have developed an emotional response. This explains why we see so many cute puppies, adorable children, attempts at humor, and dramatic disaster footage in advertising and news programming. This also is why we are more likely to remember the score of a game our favorite team played than the scores of other games. It is why we remember news events that "hit close to home" compared to other reported world events.

The strongest emotions we experience as human beings may have developed as ways to improve our individual chances for survival. For example, by avoiding experiences that elicit fear we may be avoiding danger and risk of harm. By avoiding that which we find disgusting, we may be avoiding that which could make us ill. A lot of questions surround the emotional world—the origins of emotions and what is innate versus learned, for example. One thing that research since the 1970s concerning affective education has demonstrated is that how we interpret our own and others' emotions (affect) has a strong influence on how we behave in different situations (e.g., Baskin & Hess, 1980). Because the ability to identify and label emotions (knowledge of your own and those of other people) has such a significant impact on interpersonal interactions and behaviors, a great deal of attention has been directed towards trying to understand individual differences in "emotional intelligence" or "EQ" (see Goleman, 1995).

Faith. Religious faith is only one aspect of how faith relates to knowledge. For some individuals, religious teachings provide them with knowledge surrounding certain topics. What a person knows through faith is considered to require no further analysis or proof—that is the nature of faith. For this reason, we can include other sources of knowledge derived through faith. This includes relying on knowledge attained through information shared by a mentor, expert, or authority. Identifying someone as an authority is intensely personal and individualized. Just as "beauty is in the eye of the beholder," perceptions of another person's expertise are relative. One person might have faith in the expert opinion shared by a parent, grandparent, sibling, or best friend. Another person might place greater faith in the opinions expressed by celebrities, sports figures, news anchors, or political leaders. Social work interns may place faith in the expertise shared by field instructors and course instructors. Social work clients often rely on the knowledge shared by others in their support or treatment group who may have experienced similar challenges in life.

Knowledge derived from experts and authority is not foolproof, however. First, there may be differences of opinion expressed by different experts. For example, legal counsel on different sides of a court case often engage expert witnesses who have diametrically opposed expert opinions. Jury members are then faced with the challenge of evaluating the experts' qualifications to decide which opinion is more reliable. Second, the knowledge held by experts and authorities is subject to the same sources of bias and limitation as any knowledge. For example, the expertise of social work supervisors was (at least partially) influenced by their supervisors' expertise, which was influenced by their supervisors' expertise, and so on. What if the original was wrong from the start? Or, more likely, what if that knowledge no longer applies to the current environment? In short, faith in the knowledge of authority needs to be tempered by awareness that knowledge from these sources is fallible.

> "*Arguments from authority carry little weight—'authorities' have made mistakes in the past and will do so again*" (Sagan, 1987, p. 12).

*Imagination and Intuition.*Sometimes what a person "knows" is not based on external events or objects, knowledge might be generated as new ideas or concepts completely through internal processes. Imagination, creativity, and intuition are special forms of cognitive activity, resulting in the creation of unique forms of knowledge. You have probably been exposed to examples of highly imaginative, creative art, music, film, dance, or literature. Sometimes the original insights we gain from these creations make them seem like sheer genius. We might say the same thing about a person's apparent cognitive leaps that we might call intuition: seeming to know something without being able to trace where the knowledge was derived; it comes from reasoning that happens at an unconscious level. These ways of knowing combine elements from other forms previously discussed, but in different ways: perception, memory, emotion, language, and reasoning, for example. This is what leads to that "aha" moment when a new idea is born.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=22

Culture and Ways of Knowing.

Earlier, you read that language is one way that culture has a powerful influence on knowledge. The influence of culture is not only in terms of what we know, but also in terms of how we learn what we know. The book *Women's Ways of Knowing: The Development of Self, Voice, and Mind*(Belenky, Clinchy, Goldberger, & Tarule, 1986) describes the results of a deep qualitative study of how women develop their views and knowledge. The women's development of knowledge was heavily influenced by their cultural experiences with authority, particularly in learning from family and schools. The authors included an analysis of five different ways of knowing, outlined below. These may very well express others' ways of knowing, based on their own experiences with majority culture, not only based on gender.

Silence. As a way of knowing, silence is an expression of "blind" obedience to authority as a means of avoiding trouble or punishment. The concept of silence refers to an absence of self-expression and personal voice. The authors indicated that individuals characterized by silence also found the use of words to be dangerous and that words were often used as weapons.

Received Knowledge. Carriers of this way of knowing rely on others for knowledge, even knowledge about themselves. They tend to learn by listening to others, lacking confidence in their own original ideas or thoughts. There is a heavy reliance on authority as a source of knowledge. As a result, thinking is characterized by an inability to tolerate ambiguity, and is concretely dualistic—everything is categorized as good or evil, us or them, and black or white with no shades in between.

Subjective Knowledge. Subjective knowledge is highly individualistic because personal experience, "inward listening" and intuition are major sources of knowledge. A sense of voice exists, expressed as personal opinion. The source of knowledge is internal, and the authority listened to is self—and this knowledge is more important than knowledge from external sources.

Procedural Knowledge. Objective approaches are applied to the process of acquiring, developing, and communicating procedural knowledge. In other words, careful observation and critical analysis are required. Knowledge can come from multiple external sources, but only after careful internal critical analysis of the arguments provided by those sources. There exists awareness that others and one's self can be wrong, and an effort to remove the impact of feelings from the process of objective analysis.

Constructed Knowledge. This is an integrative view where both subjective and objective strategies apply to knowledge. Ambiguity is tolerated well, as are apparent contradictions in what one knows. There exists a strong, constantly developing sense of self in relation to the external world. Constructed knowledge involves belief that knowledge is relative to context and frame of reference, and a person is responsible for actively trying to understand by "examining, questioning, and developing the systems that they will use for constructing knowledge" (see https://www.colorado.edu/UCB/AcademicAffairs/ftep/publications/documents/WomensWaysofKnowing.pdf).

Thinking about Critical Thinking

As we have seen, much of what humans "know" is susceptible to bias and distortion, often represents only partial or incomplete knowledge, and socially constructed knowledge varies as a function of the different societies constructing it. The concept of ***critical thinking*** describes approaches to knowledge and knowing that rely on disciplined, logical analysis, and knowledge that is informed by evidence. Critical thinking extends beyond simple memorization of facts and information; it involves the analysis and evaluation of information, leading to reasoned, "thought-out" judgments and conclusions. You might recognize elements of critical thinking in the procedural and constructed ways of knowing previously described. A person engaged in critical thinking expresses curiosity about the topic or issue under consideration, seeks information and evidence related to the topic or issue, is open to new ideas, has a "healthy questioning attitude" about new information, and is humbly able to admit that a previously held opinion or idea was wrong when faced with new, contradictory information or evidence (DeLecce, n.d.).

Carl Sagan (1987) argued that critical thinking is, "the ability to construct, and to understand, a reasoned argument and—especially important—to recognize a fallacious or fraudulent argument" (p. 12). To understand what a fallacious argument might be, we can examine the concepts of logic: a branch of philosophy dedicated to developing the principles of rational thought. Sagan's (1987) discussion of critical thinking indicated that it does not matter how much we might like the conclusions drawn from a train of reasoning; what matters is the extent to which the chain of logic "works." He also contended that untestable propositions are worthless—if you cannot check out the assertions, it is not worthwhile considering. However, Sagan also reminded readers that critical thinking and science itself exist in a state of constant tension between two apparently contradictory attitudes:

> *"an openness to new ideas, no matter how bizarre or counterintuitive they may be, and the most ruthless skeptical*

> *scrutiny of all ideas, old and new. This is how deep truths are winnowed from deep nonsense*" (Sagan, 1987, p. 12).

This quote describes the quality of a critical thinker mentioned earlier: simultaneously being open-minded and having a reasonable degree of skepticism about new information or ideas.

In terms of social work and social work education, being able to apply critical thinking skills, particularly "the principles of logic, scientific inquiry, and reasoned discernment" has been a curriculum expectation since the 1990s (Mathias, 2015, p. 457). The characteristics of critical thinking in social work include rational, reasoned thinking about complex, "fuzzy" problems that lack at least some elements of relevant information and readily apparent solutions (Mathias, 2015; Milner & Wolfer, 2014). Critical thinking in social work practice involves initially suspending judgment, then engaging in the process of generating relevant questions, considering assumptions, seeking information and divergent viewpoints, and applying logical, creative problem solving.

What is (Scientific) Evidence?

Strong arguments appear in the published literature concerning the importance of scientific reasoning in social work practice, and its significance to critical thinking and problem solving (e.g., see Gambrill, 1997; Gibbs & Gambrill, 1999; Gibbs et al., 1995). Francis Bacon is credited with presenting an approach to scientific method that is a foundation for science today (Dick, 1955). The Merriam-Webster definition of scientific method (https://www.merriam-webster.com/dictionary/scientific%20method) is:

Principles and procedures for the systematic pursuit of knowledge involving the recognition and formulation of a problem, the collection *of data though observation and experiment, and the formulation and testing of hypotheses.*

What you may find interesting is that different authors suggest different numbers of steps in the process called scientific method—between four and ten! Let's look at what these descriptions have in common:

Step 1. It all begins with an observation about something that arouses curiosity. For example, a social work investigator and her colleague observed an emerging population of people who rely on emergency food pantry support for relatively long periods of time—despite the programs' intended period of emergency support being relatively short-term (Kaiser & Cafer, 2016).

Step 2. The observation leads the curious social worker to develop a specific research question. In our example, the investigators asked what might be "the differences, particularly in food security status and use of federal support programs, between traditional, short-term emergency pantry users and an emerging population of long-term users" (Kaiser & Cafer, 2016, p. 46).

Step 3. Out of their questions, the social work scientist develops hypotheses. This may come from a review of existing knowledge about the topic—knowledge that is relevant but may not be sufficient to directly answer the question(s) developed in step 2. In our food security example, the investigators reviewed information from literature, community practice, and community sources to help identify key dimensions to explore. The hypothesis was, essentially, that exploring three factors would be informative: longevity, regularity, and trends of food pantry use.

Step 4. The social work scientist develops a systematic experiment or other means of systematically collecting data to answer the research question(s) or test the research hypotheses from steps 2 and 3. Kaiser and Cafer (2016) described their research methodology for randomly selecting their sample of study participants, their research variables and measurement tools, and their data collection procedures.

Step 5. The social worker will analyze the collected data and draw conclusions which answer the research question. This analysis not only includes the results of the study, but also an assessment of the study's limitations of the methods used for the knowledge developed and the implications of the study findings. Our working example from the food security study described the approaches to data analysis that were used. They found that almost 67% of 3,691 food pantry users were long-term users, most commonly female and white. Being a persistent user (versus short-term user) was predicted by having Social Security, retirement, pension, or SSI/disability payment income. Also, the probability of being a persistent pantry user increased with participant age. The authors also discussed other patterns in use as related to some of the other study variables and the implications related to food security and economic vulnerability (Kaiser & Cafer, 2016).

Step 6. Social work scientists share their resulting knowledge with colleagues and others through professional presentations, training sessions, and professional publications. As noted in the reference list to this module, the investigators in our study example published their findings in a professional journal (see Kaiser & Cafer, 2016).

Step 7. The social work scientist then may ask new, related questions, and progress through the stages anew. In our food security example, the authors concluded with a new question: would raising minimum wage or establishing a living wage be a better means of supporting families in the precarious position of long-term dependence on emergency food pantry support?

As you can see, the scientific process potentially incorporates multiple sources of knowledge that we have discussed in this module and allows for inclusion of various methods of investigation depending on the nature of the research questions being asked. The result of quality science, regardless of methodology, is called research or scientific evidence.

In an interview where he was asked if religion and science could ever be reconciled, Stephen Hawking responded:

> *"There is a fundamental difference between religion, which is based on authority, [and] science, which is based on observation and reason. Science will win because it works."*

It may not be necessary to continue pitting these two ways of knowing against each other. Engaging in science involves multiple ways of knowing. Scientific activity requires a certain degree of imagination about which might be true, rational reasoning about how to demonstrate what might be true, perception and observation of what is happening, memory about what has happened in the past, language to communicate about the process and results, an emotional investment in what is being explored, (sometimes) a certain degree of intuition, and faith in the scientific process and what has come before. Acknowledging the collaborative, constructive, dynamic nature of scientific knowledge development, consider a quote Sir Isaac Newton borrowed from an earlier popular saying:

> "*If I have seen further, it is by standing on the shoulders of giants.*"

This does not mean always having noncritical faith in the results of science, however. Scientific results are wrong sometimes. Sometimes scientific results are inconsistent. Sometimes scientific results are difficult to interpret. For example, are eating eggs bad for you? In the early 1970s, a link between cholesterol levels in the body and heart disease was detected. The advice coming from that observation was that people should limit their consumption of cholesterol-containing foods (like eggs) to prevent heart disease. However, this conclusion was based on faulty logic:*eating*cholesterol was not the cause of the elevated cholesterol levels; cholesterol was being produced by the body. Further research indicated no increased risk of heart disease with eating eggs; the elevated cholesterol problem stems from other unhealthy behavior patterns. This example demonstrates the serial, sequential, accumulating nature of scientific knowledge where previous assumptions and conclusions are questioned and tested in further research. We need to accept a degree of uncertainty as complex problems are unraveled, because science is imperfect. Again, quoting Carl Sagan (1987):

> "*Of course, scientists make mistakes in trying to understand the world, but there is a built-in error-correcting mechanism: The collective enterprise of creative thinking and skeptical thinking together keeps the field on track*" *(p. 12).*

Critically important is that science be conducted and shared with integrity. Research integrity concerns conducting all aspects of research with honesty, fairness, and accuracy. This includes objectively examining data, being guided by results rather than preconceived notions, and accurately reporting results and implications. ***Research integrity***, or "good" science means that a study's methods and results can be objectively evaluated by others, and the results can be replicated when a study is repeated by others. Science nerds enjoy the playfulness presented in the scientific humor magazine called *The Journal of Irreproducible Results* (www.jir.com).

Take a moment to complete the following activity. For each statement, decide if it is accurate (True) or not (False).

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=22

Imperfect Science versus Pseudoscience. Even "good" science has certain limitations, imperfections, and uncertainties. For example, conclusions drawn from statistical analyses are always based on probability. In later modules you will learn what it means when a statistical test has $p<.05$ and about the 95% confidence interval (CI) of a statistical value. In practical terms, this means that we cannot be 100% certain that conclusions drawn from any one study's statistical results are accurate. While accepting that we should not engage in unquestioning acceptance even of "good" science, it is even more important to avoid unquestioning acceptance of information offered through "poor" science. Especially, it is important to recognize what is being presented as scientific knowledge when it is not science. By way of introducing the importance of scientific data and statistics, Professor Michael Starbird (2006, p. 12) introduced a famous quote by the author Mark Twain:

> *"There are three kinds of lies: lies, damned lies, and statistics."*

Professor Starbird's response was:

> *"It is easy to lie with statistics, but it is easier to lie without them."*

Pseudoscience is about information that pretends to be based on science when it is not. This includes practices that are claimed to be based on scientific evidence when the "evidence" surrounding their use was gathered outside of appropriate scientific methods. You may have learned about "snake oil salesmen" from our country's early history—people who fraudulently sold products with (usually secret) ingredients having unverifiable or questionable benefits for curing health problems—these products often caused more harm than cure! The pseudoscientific "evidence" used in marketing these products often included (unverifiable) expert opinions or anecdotal evidence from "satisfied" customers. See the historical advertisement about Wizard Oil and all the ills that it will cure: rheumatism (joint pain), neuralgia (nerve pain), toothache, headache, diphtheria, sore throat, lame back, sprains, bruises, corns, cramps, colic, diarrhea, and all pain and inflammation. The product was 65% alcohol and other active ingredients included turpentine, ammonia, capsicum, camphor, gum camphor, fir oil, sassafras oil, cajeput oil, thyme oil, according to the Smithsonian National Museum of American History. Fortunately, Wizard Oil was a liniment, applied topically and not consumed orally.

What does this look like in present times? We still, to this day, have people relying on unsubstantiated products and treatment techniques, much of which can be identified as "quackery." For example, 9% of out-of-pocket health care spending in the United States pays for complementary health approaches (e.g., vitamins, dietary supplements, natural product supplements and alternative practitioners): over $30 billion (Nahin, Barnes, & Stussman, 2016). Many alternative and complementary health approaches lack a strong scientific evidence base—that does not necessarily mean that they offer no health benefits, but it may be that they do not help and sometimes they cause harm, especially with overuse. Proponents of these approaches adopt what Gambrill and Gibbs (2017) refer to as a "casual approach" to evidence, where weak evidence is accepted equally to strong evidence, and an open mind is equated with being non-critical.

Three reasons why this matters in social work (and other professions) are:

- There may be harmful "ingredients" in these interventions. For example, an untested treatment approach to helping people overcome the effects of traumatic experiences might have the unintended result of their being re-traumatized or further traumatized.
- Engaging in unproven interventions, even if not containing harmful "ingredients," can delay a person receiving critically needed help of a type that has evidence supporting its use. For example, failure with an untested approach to treating a person's opioid use disorder extends that person's vulnerability to the harmful effects of opioid misuse—including potentially fatal overdose. This is difficult to condone, particularly in light of emerging evidence about certain combination approaches, like medications combined with evidence-supported behavioral therapies (medication assisted treatment, or MAT), having significant effectiveness for a range of individuals under a range of circumstances (Banta-Green, 2015).

Treatment failure with untested, ineffectual interventions can discourage a person from ever seeking help again, including from programs that have a strong track record of success. Engaging in these ineffectual interventions may also deplete a person's resources to the point where evidence-supported interventions may not be affordable.

Despite limitations, pseudoscience does become applied. For example, the book about "psychomythology" called 50 *Great Myths of Popular Psychology* (Lilienfeld, Lynn, Ruscio, & Beyerstein, 2010) discusses implications of the myth that most humans use only 10% of their potential brain power. The authors suggest that this myth is pervasive, in part because it would be great if true, and because it is exploitable and lucratively marketable through programs, devices, and practices available for purchase that will help people tap into the unused 90% reserve.

Learning to Recognize Pseudoscience. In *Science and Pseudoscience in Clinical Psychology*, the authors identified ten signs that raise the probability that we are looking at pseudoscience rather than science (Lilienfeld, Lynn, & Lohr, 2015). In their book *Science and Pseudoscience in Social Work*, Thyer and Pignotti (2015) help explain these signs. The ten signs are summarized here.

1. *Excuses, excuses, excuses.*When a strongly held belief fails to be supported by rigorous empirical evidence, proponents of that belief may generate excuses to explain away the failure. These excuses are often unsupported by evidence. Hundreds of examples are present in recent political arguments aired on news talk shows. In science, when an expected result is not observed through application of scientific method, new hypotheses are generated to explain the unexpected results—then these new hypotheses are tested through scientific methods.

2. *Absence of self-correction makes the belief grow stronger*. In pseudoscience, erroneous conclusions go untested, so corrections of theory and conclusions do not happen. Proponents of an untested, erroneous belief have no impetus to make corrections and critical questioning may lead them to "dig in their heels" more strongly. Scientists admittedly do make mistakes, but errors are eventually eliminated through application of scientific methods.
3. *Conspiracy theories.*Proponents of a pseudoscientific belief may argue that scientific review of their "evidence" is never going to be favorable because of bias built into the scientific process and the science community. The approach is a "blame the system" strategy, as opposed to recognizing the flaws in their work. Pseudoscience is not likely to survive peer-review scrutiny because the methodologies, assumptions, and/or conclusions are flawed. Of concern in the information technology era are the many open-source journals appearing globally that claim peer-review and falsify editorial board membership. These "predatory journals" do not provide the review services present with legitimate journals and may be challenging for information consumers to spot (more on this topic in Chapter 4).
4. & 5. combined: *Shifting the burden of proof.*A sign of pseudoscientific beliefs involves demanding that critics prove them wrong (rather than assuming responsibility for proving the claim right). In science, the burden of proof is on the one making the claim. This goes along with the sign where promoters of a pseudoscientific claim seek only confirmatory evidence, not any evidence that might disprove the claim. Scientists, however, "bend over backward" to design studies that "test and potentially falsify our most cherished notions" (Thyer & Pignotti, p. 10).
6. *Unlike anything seen before.*A common claim with any pseudoscientific beliefs is that it represents a totally new

phenomenon, a dramatically different innovation, unlike anything seen before. In science, innovation is derived from existing models, theories, or interventions. Often the innovation involves a unique, novel, new way of thinking about what came before, but connectivity between past, present, and future can be demonstrated. With pseudoscience, claims of the extraordinary are involved, disconnecting the innovation from anything in the past. "Hence, the well-known saying that extraordinary claims require extraordinary evidence applies" (Thyer & Pignotti, 2015, p. 11).

7. *The power of testimonials.* Promoters of pseudoscientific innovations rely heavily on testimonial statements and anecdotal "evidence" provided by those who have been swayed into belief. While it may be true that an individual providing the testimonial was helped by the innovation, this alone does not constitute reliable scientific evidence. It does not attempt to rule out influences such as social desirability, placebo effects, and poor retrospective recall on the person's perceptions. The proponents of a pseudoscientific claim also do not share any complaints or reports of negative results in their advertising. This is at the heart of recent challenges to how social media product reviews are biased and potentially unreliable.
8. *Relying on jargon.*Pseudoscientific claims often introduce technical-sounding jargon that merely obscures the facts. "Technobabble" may be invented for a specific purpose, it may involve misapplication of scientific terminology, and/or it may involve "mashing" together concepts from different fields of science. Here are some examples presented in an online video demonstration of pseudoscientific language, Skeptic Presents: Get Your Guru Going (https://scienceornot.net/2012/05/25/technobabble-and-tenuous-terminology-the-use-of-pseudo-scientific-language/). These examples were presented in a "TED Talks" style and context.

> *"Ladies and Gentlemen, consciousness is merely the potentiality of quantum locality expressing itself beyond the constraints of four-dimensional space-time. And, that is the externalization of what we call love."*

> *"Mindfulness, on the other hand, differentiates itself into the multi-dimensional expanse of external reality. It really is that simple. Quantum perception embraces all potential space-time events exponentially. Cosmic balance is simply the way the universe transcodes the raw potential of quantum energy into unbridled happiness."*

9. *One size fits everything*. In the world of fashion, "one size fits all" is a false claim—there are always going to be individuals for whom that item is too large, too small, too short, or too long. The same is true in social work—the most effective possible interventions still fail to help some individuals. It is important for science to identify for whom an intervention might not work as intended; future scientific work might be able to unravel the reasons and create interventions targeted specifically to those individuals. Pseudoscientific claims tend to be overly inflated—the interventions are often advertised as working for everyone and/or for everything. A scientific approach requires evidence that it works for each condition claimed.
10. *The Combo Meal.* Just as fast-food eateries offer consumers a combination meal option, rather than ordering individual items, pseudoscientific approaches are often presented as part of a package—where the other parts of the package might have supporting evidence. What makes this pseudoscience is that the new add-ons have not been scientifically evaluated independently, and the combination may not have the support of evidence, either. Science does not stop at demonstrating that an intervention has an effect on outcomes; a great deal of effort is directed to understanding the mechanisms of change, how

> the changes are produced. Not only would each intervention alone be tested, the mechanisms by which they operate would be studied, as would their interactions when combined.

When broken down and analyzed this way, you might wonder why people believe in pseudoscience, especially in the face of challenges from more critical thinkers. At least two forces are in play at these times. One is that people sometimes desperately seek simple, effective solutions to difficult problems. Parents tormented by watching their child disappear into the altered state of schizophrenia, individuals haunted by the cravings induced by addiction, anyone who experiences wrenching grief over the death of someone they love: these people are particularly susceptible to pseudoscientific claims.

The second reason people often stick to pseudoscientific beliefs despite a lack of evidence relates to a psychological phenomenon known as cognitive dissonance. A significant internal conflict is created when a person's deeply held belief or attitude is challenged by contradictory beliefs, attitudes, or evidence. The discomfort of this conflicted experience is called ***cognitive dissonance***. The person is left with few choices: change the original belief/attitude to fit the new information, remain in limbo with the two in conflict, or refute the new information and stick to the original belief/attitude. Remaining in limbo is an extremely uncomfortable state, so the person is likely to restore balance and comfort by selecting one of the other two options. Thus, it is not unusual for a person to reject the new evidence and cling to the original belief.

Consider the example of how you might respond to a homeless person on the street who asks you for money. Why does this scenario make you feel uncomfortable? On one hand, you see yourself as a kindly, generous, giving person who cares about other people. On the other hand, you do not want to give your hard-earned money away to strangers. These two beliefs come into conflict when the person in need asks for money, a conflict you experience as the uncomfortable state of cognitive dissonance. (One solution is to resort to rationalization to resolve the conflict and not give the money: it is not good to reinforce this behavior, the money might do harm if the homeless person buys alcohol or drugs with it, someone else with more money than I have will donate, my taxes go to providing services to this type of person so I have already donated, letting on that I have any money will make me vulnerable to being robbed, and other reasons.)

Emerging Science. Does all of this mean that all unproven techniques are pointless and should be avoided? No, not really. After all, we should consider the observation attributed to Sir Isaac Newton:

> *"What we know is a drop, what we don't know is an ocean."*

Knowledge and science are always emerging and evolving. It is possible that what is unheard of today is going to be common knowledge in the future. In other words, what is today an alternative or complementary treatment approach may eventually become a standard practice based on evidence of its effectiveness and safety. For example, not long ago, meditation practices were considered to be unconfirmed alternative mental health intervention strategies. Evidence supports the inclusion of mindfulness-based interventions (MBIs) and meditation alongside other intervention strategies for addressing a range of mental health concerns (Edenfield & Saeed, 2012). This includes stress reduction, anxiety, depression, and combat veterans' post-traumatic stress disorder (PTSD) (Khusid & Vythilingam, 2016, Perestelo-Perez, et al., 2017).

A more current conundrum concerns the use of EMDR (Eye Movement Desensitization and Reprocessing) therapy as a treatment approach for addressing PTSD. As scientific study progresses, elements of this therapy that are similar to other evidence-supported approaches seem to have significant treatment benefits, but it is unclear that all of the eye-movement procedures are necessary elements (Lancaster, Teeters, Gros, & Back, 2016). Investigators are learning more about EMDR as the mechanisms of change in the treatment process are being more systematically and scientifically investigated. This EMDR therapy example fits into that part of the ocean about which we do not know enough. It serves as a reminder that we need to be mindful of the tension between what we know is true, what could be true, and what masquerades as truth.

Distinguishing between Facts and Opinions

We hear hundreds of statements of fact and opinion every day—from family, friends, the media, and experts in our field. It is important to decipher which are facts and which are opinions in making decisions about important matters. A fact is information that can be objectively proven or demonstrated. An opinion is a personal belief or point of view, is subjective, and proof is not relevant. Just because a person expresses an opinion does not mean that person has useful knowledge to impart.

> "*There are in fact two things, science and opinion; the former begets knowledge, the latter ignorance*" (*Hippocrates, circa* 395 BC).

In a 1987 essay,*The Fine Art of Baloney Detection: How Not to be Fooled*, Carl Sagan addressed the importance of critical thinking and a healthy degree of skepticism about the information and knowledge that we encounter. He stated that:

> "*Skeptical habits of thought are essential for nothing less than*

> *our survival—because baloney, bamboozles, bunk, careless thinking, flimflam and wishes disguised as facts are not restricted to parlor magic and ambiguous advice on matters of the heart. Unfortunately, they ripple through mainstream political, social, religious and economic issues in every nation*"(p. 13).

It is worth noting that they also ripple through the world of professional practice, as well. So, how do we tell the difference when we are provided with practice-related information? Here are some ways of assessing the information (Surbhi, 2016).

- Is the information based on observation or research, can it be validated or verified with pieces of evidence (fact)? Or, is it based on assumptions and personal views, representing a personal perception, judgment, or belief (opinion)?
- Does it hold true in general (fact) or does it differ individually (opinion)?
- Is it presented or described in objective, unbiased words (fact) or is it expressed with subjective, biased words (opinion)?
- Is it a debatable view (opinion) or non-debatable (fact)?
- Facts can change opinions but opinions cannot change facts.

The Problem with Expert Opinion. Over years of practice, social workers develop a wealth of experience in identifying and addressing certain problems encountered in routine practice. Under the heading of "practice makes perfect," experience contributes to developing a certain degree of practice wisdom and expert opinion. What is important in this label is the word "opinion." Experience does not change facts, it shapes opinion. Expert opinion is not, in itself, a bad thing to turn to. The problem lies in non-critical reliance on expert opinion. Let's consider an example that a social work intern experienced in practice.

At intake in a mental health center, Mr. R (aged 34) describes recent scary incidents where he has experienced a loss of memory for certain events and found himself in dangerous situations. Most recently, he "came to" standing on the edge of a river embankment with the potential of falling into rushing water at night. He has complete amnesia for how he came to be at the river. In supervision, the intern suggested two hypotheses: Mr. R has a dissociative disorder and these incidents might involve suicidality, or Mr. R has alcohol (or other substance-induced) blackouts. The intern's social work supervisor discounted the first hypothesis on the grounds that dissociative disorders do not occur in men, only women. Furthermore, the supervisor stated that dissociative disorders are rare, making it unlikely that the intern would encounter this in a first semester field placement, since most practitioners never see such a case in an entire career. The supervisor indicated that the proper assessment would be a substance use disorder.

Relatively few facts are present in this example. They include:

- Mr. R was seen by the intern at intake to the mental health center.
- The intern was in the first semester of field placement.
- The client provided a description of events.
- The intern offered two hypotheses.
- Dissociative disorders are rare.
- The supervisor did not address the potential suicidality risk of the client's reported incidents.

The following either fall short of being demonstrable facts or represent opinion:

- The intern has the client's description of what happened but

does not know for a fact that these things happened.
- Dissociative disorders do not occur in men, only in women (they are more common among women but do occur among men).
- Because the disorder is rare, this would not show up in a first-semester internship caseload (a classic logic fallacy).
- The proper assessment is substance use disorder.

Based on this analysis, the social work intern should consider the "expert opinion" of the supervisor but should also seek more information to confirm the hypothesis of a substance use disorder. The intern should also keep an open mind to the other possibilities, including dissociative disorder (often related to a trauma experience) and suicidality, and seek information that either supports or refutes these alternative conclusions.

The Problem with Groupthink. Sometimes an individual makes decisions or judgments in conformity with a group's thinking, decisions that run contrary to what that individual would have decided alone. The process of group think typically refers to faulty decision-making resulting in a group's systematic errors. Groupthink is characterized by individuals abandoning their own critical analysis, reality testing, and reasoning—possibly abandoning their own moral reasoning (Street, 1997). An example discussed by the originator of the groupthink concept was the space shuttle *Challenger* disaster: the commission appointed to determine the probable causes of this incident concluded, among other things, that "a highly flawed decision process was an important contributing cause of the disaster (Janis, 1991, p. 235). A more everyday example follows:

> *Four friends on vacation at Hawai'i Volcanoes National Park hiked out to view an active volcano. They arrived at a point where park rangers had posted signs indicating that no one*

should go further because of the danger of breaking through thin lava. The group members were disappointed that they could only see the active lava flow in the distance. The group's informal leader announced they had come this far and it looked safe enough; he encouraged everyone to follow him out onto the lava bed to get a closer look. Two other members started to willingly go along with him. One member decided that she was not going to cave in to group think, and told the others so. The two following the leader soon turned back, as well, conceding that this was ill-advised. The "leader" eventually agreed, and returned safely, too.

Critics of the original groupthink model proposed by Janis (1991) argue with the emphasis placed on certain aspects of group cohesion as a precursor to the observation that group decision-making can be faulty. However, researchers have observed that "the propensity for the group to display groupthink symptoms" is greatest when the group is characterized by a high degree of cohesion based on the socioemotional dimension and a low degree of task-oriented cohesion (Street, 2017, p. 78). This difference in groupthink occurrence is attributed to the importance placed on analytical (critical) thinking in task-oriented groups. These findings have important implications for social work practitioners working in organizations and communities, in particular.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=22

Module 1 Chapter 2: Research and the Professional Ethics Mandate

America's long, complicated, ambivalent, and confusing relationship with research, evidence, and factual information have profound implications for how social workers practice our profession. In 1964, a Pulitzer Prize was awarded to Richard Hofstadter's book (1963) analyzing the role played by anti-intellectual passions in a democratic society—intellect being about critical, creative, and contemplative thinking. The author's analysis is relevant today because it addresses the depreciation of knowledge as a societal counter-force acting against social change; counter-forces that are especially prevalent in periods characterized by "complex and bewildering global change" (Baer, 2017).

What you will learn from reading this chapter is:

- Why research matters to the social work profession and the nature of the relationship between research and ethical social work practice
- Principles and standards in the social work code of ethics that address the relationship between research and practice.

Before you begin reading, please answer the following question for yourself–make a note of your answer using the 5-point scale.

When I think about the importance of research and social work practice skills, I think that:

1. practice skills are far more important
2. practice skills are somewhat more important
3. research and practice skills are about equally important
4. research skills are somewhat more important
5. research skills are far more important

A Place for Research in The Social Work Profession

One feature of a profession is that members of that profession participate in the systematic study of its practices and engaging in knowledge-building processes that shape practice (Kirk & Reid, 2002; Zlotnik, 2008). Having a body of knowledge and theory, as opposed to simply having a set of practices or skills, is a distinguishing feature between an occupation and a profession (Kirk & Reid, 2002).

Research is an important source of knowledge for the profession of social work. Translated from an original Arabic analysis is the phrase "knowledge is power" (attributed to Imam Ali). Similar is a quote from Ralph Waldo Emerson (1862): "…knowledge comes by eyes always open, and working hands and there is no knowledge that is not power" (p. 136).

Consider this description of social work research:

> *"Social work research is the use of social research methods for developing, producing, and disseminating knowledge that is pertinent to policies and practices…It aims to describe and explain phenomena relevant to social work" (Tripodi & Lalayants, 2008, p. 512)*

And, in an essay about the science of the social work profession, the authors answer the question "Why social science?" with the conclusion, "Because social science—and social work research in particular—is an engine for social progress" (Coffey & Williams, 2017). This observation leads to an important point: social work is a long-standing, proud profession that draws on an expanding body of knowledge, theory, and research—both its own and that generated through multiple other disciplines. Because social work practice covers so many broad topics, problems, and populations, and because social work professionals work in collaboration with so many other professions, social work scholars are well positioned to contribute to and draw from knowledge in multiple professions and disciplines.

The NASW Code of Ethics

By now you may have been introduced to the 36-page National Association of Social Workers ***Code of Ethics***(NASW, 2017) that outlines patterns of professional behavior that define ethical social work practice. In this chapter, we examine several sections of the professional code of ethics that directly address the relationship between research/evidence and professional social work practice. This is not quite the same thing as what we explore in Chapter 4 where we focus more specifically on the ethical conduct of research.

You can access the NASW Code of Ethics at the interactive website https://www.socialworkers.org/About/Ethics/Code-of-Ethics/Code-of-Ethics-Englishor as a pdf link in the Carmen course site.

Beginning with The Sixth Principle. The first place where the code of ethics content directly relates to the relationship between ethical social work practice and research is in the last of the six broad ethical principles (NASW, 2017). This principle states:

> *Social workers practice within their areas of competence and develop and enhance their professional expertise* (p. 6).

This single statement is followed by two additional statements that elaborate the point:

> *Social workers continually strive to increase their professional knowledge and skills and to apply them in practice. Social workers should aspire to contribute to the knowledge base of the profession* (p. 6).

How can you as a social worker contribute to the knowledge base of the profession? Engaging in research is a major mechanism for doing so.

Standard 4.01: Competence. Next, let's turn to the Standards identified in the professional code of ethics for social workers (NASW, 2017). The first place where evidence and research are directly mentioned is in Standard 4.01, which is about competence. Two statements are relevant here, the first of which is:

> *(b) Social workers should strive to become and remain proficient in professional practice and the performance of professional functions. Social workers should critically examine and keep current with emerging knowledge relevant to social work. Social workers should routinely review the professional literature and participate in continuing education relevant to social work practice and social work ethics (p. 25).*

What does this mean for you as a social worker?

- First, you need to learn how to identify current relevant knowledge.
- Second, you need to develop the skills necessary to critically examine such knowledge.
- Third, you need to develop a strong value and appreciation for putting forth the effort to maintain your knowledge of the professional literature–social work education is a lifelong commitment, it does not end the day you graduate!

The second statement elaborating on the Competence Standard 4.01 is:

> *(c) Social workers should base practice on recognized knowledge, including empirically based knowledge, relevant to social work and social work ethics (p. 25).*

For you as a social worker, this means knowing:

- how and where to identify sources of knowledge,
- how to evaluate the information provided through different

sources—including the evaluation of research evidence, and
- how research evidence relates to social work practice.

Standard 5.01: Integrity of the Profession. The next place where the code of ethics directly refers to research and evidence is in Standard 5.01 addressing ***Integrity of the Profession***. Here it is stated:

> *(b) Social workers should uphold and advance the values, ethics, knowledge, and mission of the profession. Social workers should protect, enhance, and improve the integrity of the profession through appropriate study and research, active discussion, and responsible criticism of the profession (p. 27).*

What this means is that social workers have a responsibility to engage in critical thinking about what we know and what we do as professionals.

Research is mentioned again in the next statement as one of the key activities that promote the profession's integrity:

> *(c) Social workers should contribute time and professional expertise to activities that promote respect for the value, integrity, and competence of the social work profession. These activities may include teaching, research, consultation, service, legislative testimony, presentations in the community, and participation in their professional organizations (p. 27).*

The topic appears again in the last of the Integrity of the Profession statements:

> *(d) Social workers should contribute to the knowledge base of social work and share with colleagues their knowledge related to practice, research, and ethics. Social workers should seek to contribute to the profession's literature and to share their knowledge at professional meetings and conferences (p. 27).*

Standard 5.02: Evaluation and Research. This standard is the most obvious and lengthy place where the relationship between research and social work practice is presented in the NASW ethical code (2017). Here are the first three points stated in this Evaluation and Research standard—we will look at the remaining points from this standard in Chapter 4 where we take a close look at the ethical conduct of research:

- (a)Social workers should monitor and evaluate policies, the implementation of programs, and practice interventions (p. 27).
- (b)Social workers should promote and facilitate evaluation and research to contribute to the development of knowledge (p. 27).
- (c)Social workers should critically examine and keep current with emerging knowledge relevant to social work and fully use evaluation and research evidence in their professional practice (p. 27).

What do these points mean to you as a social work professional? Taken together, they reinforce some of what was emphasized in earlier standards—social workers:

- engage in efforts to evaluate their practices and programs (which includes policy as an intervention strategy),
- promote research that helps build the profession's knowledge base,
- keep up with emerging knowledge (especially in our areas of practice),
- continue to think critically about the emerging knowledge, and
- use evidence in our professional practice.

This last point is the focus of the entire second course in our sequence, SWK 3402.

Before you began reading, you answered the following question for yourself using the 5-point scale. Now compare your answers from before and after reading the chapter. If your answer changed after reading the chapter, what do you think was the most important information to stimulate this change? If your answer remained the same, what do you think was the most important information that you read to reinforce your earlier answer?

When I think about the importance of research and social work practice skills, I think that:

1. practice skills are far more important
2. practice skills are somewhat more important
3. research and practice skills are about equally important
4. research skills are somewhat more important
5. research skills are far more important

Module 1 Chapter 3: Ethical Conduct of Research

The NASW professional code of ethics has quite a lot to say about ensuring that research be conducted in an ethical manner, so here we revisit the code of ethics to review these research-specific standards. We begin with an overview of historical events that bring us to where we are today in the research ethics arena. This content is covered in greater depth and detail in your CITI (Collaborative Institutional Training Initiative) program training modules. Then, we return to the section of NASW's (2017) Code of Ethics that we initially visited in Chapter 1.

In reading this chapter, you will learn:

- Key facts about research ethics;
- Statements from the NASW Code of Ethics specific to the ethical conduct of research

History and Research Ethics

Following the conclusion of World War II and subsequent "Doctors' Trial" portion of the Nuremberg Trials, the 10-point Nuremberg Code was developed as a response to numerous examples of unethical, inhumane "medical" experiments conducted in concentration camps. The Nuremberg Code was relatively ignored for many years, at least in the United States. Subsequently, a number of grievously unethical experiments were conducted on prisoners, institutionalized patients, and children—disproportionately on poor and racial minority populations—including exposure to toxins, diseases, radiation, or torture. The Tuskegee syphilis experiment represents a critical turning point in the nation's tolerance of unethical human research. In 1932, the U.S. Public Health Service partnered with the Tuskegee Institute to study the natural course of syphilis among 399 black men. Over 40 years, the men believed they were being treated for syphilis, but were in fact receiving no treatment for the disease, even when penicillin proved to be an effective form of treatment by 1947. Beginning in 1968, concerns were being raised and condemning news reports were widely circulated in 1972, leading to the study being ended. By 1974, the Tuskegee Health Benefit Program was established by the U.S. government to ensure health and burial benefits to the study's remaining survivors, wives, widows, and children.

In response to concerns about unethical research practices, a national group was formed (in 1974, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research), and created what is now known as The Belmont Report (1979). The Belmont Report, in turn, influenced development of federal policy (1981, revised 2009) concerning protections for human research subjects—the Common Rule. The Belmont Report presented a summary of ethical guidelines for engaging in research that involves human subjects. The guidelines are based on three core principles:

- **Respect for persons.**This principle is founded on a conviction that "individuals should be treated as autonomous agents" (Belmont Report, 1979, p. 4). The report explains this autonomy in terms of respecting a person's right to making considered, informed choices, and places a responsibility on researchers to ensure that a person has all of the information necessary for self-determined choices and is free from constraints on making self-determined choices. The discussion also addresses situations where a person might not be capable of self-determination and what protections might be necessary in these instances.
- **Beneficence.**This principle is about an obligation for protecting

participants from harm and "making efforts to secure their well-being" (Belmont Report, 1979, p. 5). This translates into ensuring that benefits to participants are maximized and possible harms or risks are minimized.
- **Justice.** This principle is about fairness. Justice in this context is about ensuring a fair and just distribution of both the potential burdens and the potential benefits of participating in research. "Another way of conceiving the principle of justice is that equals ought to be treated equally" (Belmont report, 1979, p. 5).

These three core principles are closely aligned with the core values of our social work profession. Our code of ethics is based on principles that include respect for the dignity and worth of the individual, which includes fostering client self-determination, as well as practicing with integrity.

The Belmont Report (1979) also contains language about distinguishing between research and professional practice. The report defines practice in this way:

> *"…interventions that are designed solely to enhance the well-being of an individual patient or client and that have a reasonable expectation of success"* (p. 3).

Research is defined in the Belmont Report as:

> *"…an activity designed to test an hypothesis, permit conclusions to be drawn, and thereby to develop or contribute to generalizable knowledge"* (p. 3).

These two types of professional activity are clearly contrasted—their goals are markedly different. Practice goals relate to benefits for the individuals being served; research goals are served by the individuals who participate.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=199

Revisiting the NASW Code of Ethics

We previously examined what the NASW Code of Ethics had to say about the role of research and evidence in professional practice and the role of social work professionals in engaging with research and evidence. At this point, we turn again to the Code of Ethics to see what is specified in terms of the ethical conduct of research. These details are in addition to what was discussed earlier about research integrity.

Standard 5.02: Evaluation and Research continued. Let's pick up where we left off in reviewing the NASW Code of Ethics content related to engaging in evaluation and research as a social work professional. The Code of Ethics (p. 27-28) replicates many elements found in the current federal policy concerning the protection of human subjects in research and in the 1996 Health Insurance Portability and Accountability Act (HIPAA). The Code of Ethics states that:

- *(d) Social workers engaged in evaluation or research should carefully consider possible consequences and should follow guidelines developed for the protection of evaluation and research participants. Appropriate institutional review boards should be consulted*

- *(e) Social workers engaged in evaluation or research should obtain voluntary and written informed consent from participants, when appropriate, without any implied or actual deprivation or penalty for refusal to participate; without undue inducement to participate; and with due regard for participants' well-being, privacy, and dignity. Informed consent should include information about the nature, extent, and duration of the participation requested and disclosure of the risks and benefits of participation in the research.*

- *(f) When using electronic technology to facilitate evaluation or research, social workers should ensure that participants provide informed consent for the use of such technology. Social workers should assess whether participants are able to use the technology and, when appropriate, offer reasonable alternatives to participate in the evaluation or research.*

- *(g) When evaluation or research participants are incapable of giving informed consent, social workers should provide an appropriate explanation to the participants, obtain the participants' assent to the extent they are able, and obtain written consent from an appropriate proxy.*

- *(h) Social workers should never design or conduct evaluation or research that does not use consent procedures, such as certain forms of naturalistic observation and archival research, unless rigorous and responsible review of the research has found it to be* *justified because of its prospective scientific, educational, or applied value and unless equally effective alternative procedures that do not involve waiver of consent are not feasible.*

- *(i) Social workers should inform participants of their right to withdraw from evaluation and research at any time without penalty.*
- *(j) Social workers should take appropriate steps to ensure that participants in evaluation and research have access to appropriate supportive services.*
- *(k) Social workers engaged in evaluation or research should protect participants from unwarranted physical or mental distress, harm, danger, or deprivation.*
- *(l) Social workers engaged in the evaluation of services should discuss collected information only for professional purposes and only with people professionally concerned with this information.*
- *(m)Social workers engaged in evaluation or research should ensure the anonymity or confidentiality of participants and of the data obtained from them. Social workers should inform participants of any limits of confidentiality, the measures that will be taken to ensure confidentiality, and when any records containing research data will be destroyed.*
- *(n) Social workers who report evaluation and research results should protect participants' confidentiality by omitting identifying information unless proper consent has been obtained authorizing disclosure.*
- *(o) Social workers should report evaluation and research findings accurately. They should not fabricate or falsify results and should*

take steps to correct any errors later found in published data using standard publication methods.

- *(p) Social workers engaged in evaluation or research should be alert to and avoid conflicts of interest and dual relationships with participants, should inform participants when a real or potential conflict of interest arises, and should take steps to resolve the issue in a manner that makes participants' interests primary.*

- *(q) Social workers should educate themselves, their students, and their colleagues about responsible research practices.*

Together, these 14 statements reflect a responsibility for social workers to :

- participate in appropriate Institutional Review Board (IRB) procedures,
- ensure the safety and protection of participants in social work research and evaluation studies,
- engage in effective informed consent procedures,
- responsibly utilize information technology in research,
- collect and discuss only information relevant to the study,
- preserve participant privacy and confidentiality (or anonymity),
- engage in research with integrity,
- prevent or responsibly resolve conflict of interest instances, and
- remain informed about responsible research practices.

Module 1 Summary

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=83

In Module 1, we covered a range of topics related to the sources of knowledge social workers might utilize in practice. One of the major themes of this module was that there are many different sources of knowledge available to us as human beings, and additional sources to consider in relation to ourselves as professional social workers. The contents of the first chapter reflected different ways of knowing that might also help us to better understand our clients and colleagues, as well as ourselves. The Chapter 1 contents set the stage for Chapter 2, creating a context for understanding one of the major ways of knowing related to professional social work practice: evidence derived through research. Finally, Chapter 3 addressed issues in professional responsibility to engage in research that is both ethical and characterized by integrity. Using this first module as a springboard, we are now ready to begin looking more closely at the nature of the evidence that informs social work practice.

Module 1 Key Terms and Definitions

abductive reasoning: beginning with a set of observations, then seeking the most likely explanation for what was observed.

Code of Ethics: standards for professional practice, such as the NASW Code of Ethics for professional social work practice.

cognition: mental processes (thinking, experiencing, sensing) involved in knowledge and understanding.

cognitive dissonance: uncomfortable experience of inconsistencies in thoughts, beliefs, or attitudes; often motivates changing one or more of these cognitions.

critical thinking: forming judgments or conclusions on the basis of objective analysis and evaluation of information.

deductive reasoning: deriving a specific conclusion based on application of logical processes, assembling a series of truthful general premises.

empiricism: belief that knowledge is primarily derived from sensory experiences and observation.

epistemology: theory of knowledge, its nature, and how it is acquired.

ethnoepistemology: theory of knowledge based on acceptance of all human experience and epistemology as being equally valid; strongly based in anthropology.

fallacious argument: logic based on faulty reasoning, resulting in faulty conclusions.

inductive reasoning: deriving a general conclusion or hypothesis based on a set of specific observations.

linguistic determinism: the limits on knowledge imposed by the structure and contents of a language.

metacognition: the act of thinking about or being aware of thought processes and contents (thinking about cognition)

positivism: belief in the possibility of rational/logical or scientific proof of all justifiable knowledge.

pseudoscience: premises or assertions that have the appearance of being based on science without their being based on the application of strong scientific method.

qualitative research: based on a belief that knowledge is socially constructed, explores individuals' perceptions, beliefs, attitudes, experiences, and meaning-making to derive general conclusions or develop hypotheses.

quantitative research: based on objective measurement for variables of interest to answer research questions/test hypotheses about variables or their inter-relationships.

research integrity: a commitment to honesty and assuming responsibility for research conduct, resulting in replicable, reliable research results.

Module 1 References

Adams, C. (2017). The gender identity terms you need to know. CBS News (March 24). Retrieved from https://www.cbsnews.com/news/transgender-gender-identity-terms-glossary/

Bacon, F. (1955). *Selected writings of Francis Bacon*(edited by H.G. Dick). NY: Modern Library.

Baer, N. (2017). American idiot: Rethinking anti-intellectualism in the age of Trump. *Transformation*(August 27). Retrieved from https://www.opendemocracy.net/transformation/nicholas-baer/american-idiot-rethinking-anti-intellectualism-in-age-of-trump

Banta-Green, C. (2015). *Medication assisted treatment for opioid use disorders: Overview of the evidence.*University of Washington Alcohol & Drug Abuse Institute brief (June). Retrieved from http://adai.uw.edu/pubs/infobriefs/MAT.pdf

Baskin, E.J., & Hess, R.D. (1980). Does affective education work? A review of seven programs. *Journal of School Psychology*, *18*(1), 40-50.

Belenky, M.F., Clinchy, B.M., Goldberger, N.R., & Tarule, J.M. (1986). *Women's ways of knowing: The development of self, voice, and mind.* NY: Basic Books.

Belmont Report, The. (1979). *The Belmont Report: Ethical principles and guidelines for the protection o human subjects research.*Prepared by The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (April 18). Washington, DC: Department of Health, Education, and Welfare. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

Birkenmaier, J., & Berg-Wegner, M. (2017). *The practice of generalist social work, 4thed.* NY: Routledge.

Coffey, D.S., & Williams, J.H. (2017).*Why social science? Because it is an engine for social progress.*Consortium of Social Science Associations. Retrieved from www.whysocialscience.com

DeLecce, T. (n.d.). What is critical thinking? Definition, skills & meaning. Lesson transcript. Retrieved from https://study.com/academy/lesson/what-is-critical-thinking-definition-skills-meaning.html

Drisko, J.W., & Grady, M.D. (2012). *Evidence-based practice in clinical social work*. NY: Springer.

Edenfield, T.M., & Saeed, S. (2012). An update on mindfulness meditation as a self-help treatment for anxiety and depression. *Psychology Research and Behavior Management*, 5, 131-141.

Emerson, R.W. (1862). Old age. *The Atlantic Monthly*, 9(51), 134-140 (January). Retrieved from https://www.theatlantic.com/past/docs/issues/1862jan/emerson.htm

Gambrill, E.D.(1997). *Social work practice*: A *critical thinker's guide.*NY: Oxford University Press.

Gibbs, L.E., & Gambrill, E.D. (1999). *Critical thinking for social workers: Exercises for the helping professions.*Thousand Oaks, CA: Pine Forge Press.

Gibbs, L.E., Gambrill, E.D., Blakemore, J., Begun, A., Keniston, A., Peden, B., & Lefcowitz, J. (1995). A measure of critical thinking about practice. *Research on Social Work Practice*, 5(2), 193-204.

Goleman, D. (1995). *Emotional intelligence.*NY: Bantam Books.

Heussner, K.M. (2010). Stephen Hawking on religion: 'Science will win.' ABC News (June 7). Retrieved from https://abcnews.go.com/WN/Technology/stephen-hawking-religion-science-win/story?id=10830164

Hofstader, R. (1963). *Anti-intellectualism in American life.*NY: Knopf.

Holton, G. (1993). *Science and anti-science.*Cambridge, MA: Harvard University Press.

International Baccalaureate Organization (IB). (n.d.). Theory of knowledge guide: Ways of knowing. Retrieved from https://ibpublishing.ibo.org/exist/rest/app/tsm.xql?doc=d_0_tok_gui_1304_1_e&part=2&chapter=3

James, W. (1902). The varieties of religious experience: A study in human nature (Lectures VI and VII). Retrieved from http://ebooks.adelaide.edu.au/j/james/william/varieties/complete

Janis, I.L. (1991). Groupthink. Reprinted in E. Griffin, (Eds), A *first look at communication theory*,(pp. 235-246). NY: McGraw Hill.

Jones, M.A. (2002). Deafness as culture: A psychosocial perspective. *Disability Studies Quarterly*, 22(2), 51-60.

Kaiser, M.L., & Cafer, A.M. (2016). Exploring long-term food pantry use: Differences between persistent and prolonged typologies of use. J*ournal of Hunger & Environmental Nutrition*, 12, 46-63.

Khusid, M.A., & Vythilingam, M. (2016). The emerging role of mindfulness meditation as effective self-management strategy, Part 1: Clinical implications for depression, post-traumatic stress disorder, and anxiety. *Military medicine*, *181*(9), 961-968.

Kirk, S.A., & Reid, W.J. (2002). *Science and social work: A critical appraisal*. NY: Columbia University Press.

Lancaster, C.L., Teeters, J.B., Gros, D.F., & Back, S.E. (2016). Posttraumatic stress disorder: Overview of evidence-based assessment and treatment. *Journal of Clinical Medicine*, (5),1-13. doi: 10.3390/jcm5110105

Lilienfeld, S.O., Lynn, S.J., & Lohr, J.M. (2015). Science and pseudoscience in clinical psychology: Initial thoughts, reflections, and considerations. In S.O.Lilienfeld, S.J. Lynn, & J.M. Lohr (Eds.), *Science and pseudoscience in clinical psychology*, (pp. 1-16). NY: Guilford.

Lilienfeld, S.O., Lynn, S.J., Ruscio, J., & Beyerstein, B.L. (2010). 50 *great myths of popular psychology: Shattering widespread misconceptions about human behavior*. West Sussex, UK: Wiley-Blackwell.

Maffie, J. (1990). Recent work on naturalized epistemology. *American Philosophical Quarterly*, 27(4), 281-294.

Maffie, J. (1995). Towards an anthropology of epistemology. *The Philosophical Forum*, 26(3),218-241.

Mantell, A., Simpson, G.K., Vungkhanching, M., Jones, K.F., Strandberg, T., & Simonson, P. (2017). Social work-generated evidence in traumatic brain injury from 1975 to 2014: A systematic scoping review. *Health and Social Care in the Community*, 00(1), 1-16. doi: 10.1111/hsc.12476

Mathias, J. (2015). Thinking like a social worker: Examining the meaning of critical thinking in social work. *Journal of Social Work Education*, 51(3), 457-474.

Milner, M., & Wolfer, T. (2014). The use of decision cases to foster critical thinking in social work students. *Journal of Teaching in Social Work*, 34, 269-284.

Nahin, R.L., Barnes, P.M., & Stussman, B.A. (2016). Expenditures on complementary health approaches: United States, 2012. National Health Statistics Report, No. 95. Hyattsville, MD: National Center for Health Statistics. Retrieved from https://www.cdc.gov/nchs/data/nhsr/nhsr095.pdf

National Association of Social Workers (NASW). (2017). *Code of ethics of the National Association of Social Workers.*Washington, DC: NASW Press. Retrieved from https://www.socialworkers.org/LinkClick.aspx?fileticket=ms_ArtLqzeI%3D&portalid=0

Perestelo-Perez, L., Barraca, J., Penate, W., Rivero-Santana, A., & Alvarez-Perez, Y. (2017). Mindfulness-based interventions for the treatment of depressive rumination: Systematic review and meta-analysis. *International Journal of Clinical and Health Psychology*, 17,282-295.

Rimland, B. (1964).*Infantile autism: The syndrome and its implications for a neural theory of behavior*. Note: 50th anniversary edition (2015), edited by W.M. Edelson. Philadelphia, PA: Jessica Kingsley Publishers.

Russell, J.B. (1991). *Inventing the Flat Earth: Columbus and modern historians*. NY: Praeger.

Sagan, C. (1987). The fine art of baloney detection: How not to be fooled. Parade Magazine, Science on Parade, (February 1), p. 12-13. Retrieved from https://www.csicop.org/uploads/files/ParadeFeb11987.pdf

Schick, T., & Vaughn, L. (2010). *How to think about weird things: Critical thinking for a new age*, *6thedition.*NY: McGraw-Hill.

Soydan, H., & Palinkas, L.A. (2014).*Evidence-based practice in social work: Development of a new professional culture.*NY: Routledge/Taylor & Francis Group.

Surbhi, S. (2016). Difference between fact and opinion. Retrieved from Key Differences athttps://keydifferences.com/difference-between-fact-and-opinion.html

Street, M.D. (1997). Groupthink: An examination of theoretical issues, implications, and future research suggestions. *Small Group Research*, 28(1), 72-93.

Theory of Knowledge Network (TOK). (n.d.). Ways of knowing. Retrieved from http://www.theoryofknowledge.net/ways-of-knowing/

Thyer, B.A., & Pignotti, M.G. (2015). *Science and pseudoscience in social work*.NY: Springer Publishing Co.

Wagner, M.F., & Skowronski, J.J. (2017). Social influence and mental routes to the production of authentic false memories and inauthentic false memories. Consciousness and Cognition, 51, 34-52.

Williams, J.H. (2016). [Editorial] Grand challenges for social work: Research, practice, and education. *Social Work Research*, 40(2), 67-70.

MODULE 2

Module 2 Introduction

Before we can design social work interventions to address specific problems and help specific populations impacted by the problem, we first need to understand the nature of these problems for which populations might be seeking help. In this module, we explore how social work professionals identify, access, assess, and utilize evidence in understanding diverse populations, social work problems, and social phenomena. (The SWK 3402 course involves exploring research and evidence related to understanding interventions.) In this Module 2 coursebook, we learn about the implications of different types of research questions, the role of theory in addressing research questions, and strategies for identifying and analyzing existing knowledge related to a research question.

READING OBJECTIVES

After engaging with these reading materials and learning resources, you should be able to:

- Recognize different types of research, research questions, and specific aims of a study to address those questions;
- Begin linking theory to research questions about diverse populations, social work problems, and social phenomena;
- Begin identifying and assessing empirical evidence about diverse populations, social work problems, and social phenomena.
- Define key terms related to identifying and assessing evidence.

Module 2 Chapter 1: The Nature of Social Work Research Questions

The search for empirical evidence typically begins with a question or hypothesis. The nature of the questions asked determine many features of the studies that lead to answers: the study approach, design, measurement, participant selection, data collection, data analysis, and reporting of results. Not just any type of question will do, however:

> "*When the question is poorly formulated, the design, analysis, sample size calculations, and presentation of results may not be optimal. The gap between research and clinical practice could be bridged by a clear, complete, and informative research question*" (*Mayo, Asano, & Barbic*, 2013, 513).

The topic concerning the nature of social work research questions has two parts: what constitutes a *research*question, and what makes it a *social work*question. We begin this chapter by examining a general model for understanding where different types of questions fit into the larger picture of knowledge building explored in Module 1. We then look at research questions and social work questions separately. Finally, we reassemble them to identify strong social work research questions.

In this chapter, you will learn:

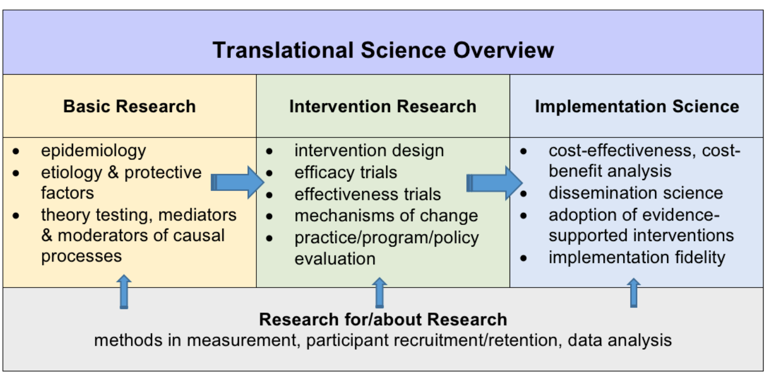
- 4 types of social work research for knowledge building,
- characteristics of research questions,

- characteristics of social work research questions.

Translational Science

The concept of ***translational science***addresses the application of basic science discoveries and knowledge to routine professional practice. In medicine, the concept is sometimes described as "bench to trench," meaning that it takes what is learned at the laboratory "bench" to practitioners' work in the real-world, or "in the trenches." This way of thinking is about applied science–research aimed at eventual applications to create or support change. Figure 1-1 assembles the various pieces of the translational science knowledge building enterprise:

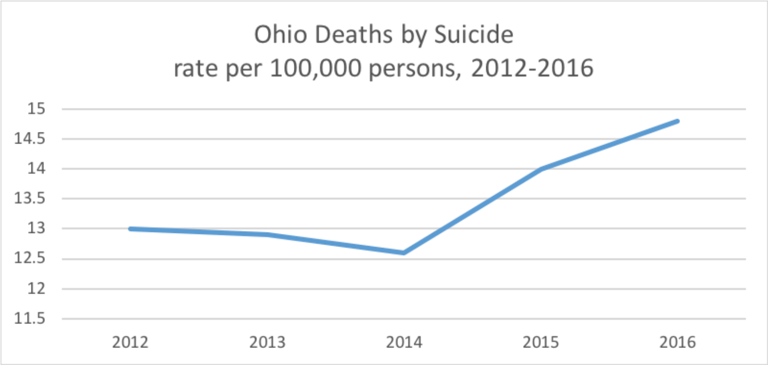
Figure 1-1. Overview of translational science elements



Basic Research. Federal policy defines ***basic research*** as systematic study that is directed toward understanding the fundamental aspects of phenomena without specific applications in mind (adapted from 32 CFR 272.3). Basic research efforts are those designed to describe something or answer questions about its nature. Basic research in social and behavioral science addresses questions of at least two major types: ***epidemiology*** and ***etiology*** questions.

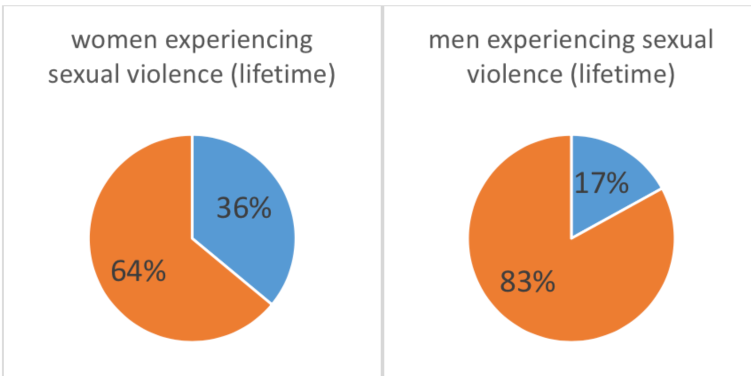
*Epidemiology questions.*Questions about the nature of a population, problem, or social phenomenon are often answered through epidemiological methods. Epidemiology is the branch of science (common in public health) for understanding how a problem or phenomenon is distributed in a population. Epidemiologists also ask and address questions related to the nature of relationships between problems or phenomena—such as the relationship between opioid misuse and infectious disease epidemics (NAS, 2018). One feature offered by epidemiological research is a picture of trends over time. Consider, for example, epidemiology data from the Centers for Disease Control and Prevention (the CDC) regarding trends in suicide rates in the state of Ohio over a four-year period (see Figure 1-2, created from data presented by CDC WONDER database).

Figure 1-2. Graph reflecting Ohio trend in suicide rate, 2012-2016



Since the upward trend is of concern, social workers might pursue additional questions to examine possible causes of the observed increases, as well as what the increase might mean to the expanded need for supportive services to families and friends of these individuals. The epidemiological data can help tease out some of these more nuanced answers. For example, epidemiology also tells us that firearms were the recorded cause in 46.9% of known suicide deaths among individuals aged 15-24 years across the nation during 2016 (CDC, WONDER database). Not only do we now know the numbers of suicide deaths in this age group, we know something about a relevant factor that might be addressed through preventive intervention and policy responses.

Epidemiology also addresses questions about the size and characteristics of a population being impacted by a problem or the scope of a problem. For example, a social worker might have a question about the "shape" of a problem defined as sexual violence victimization. Data from the United States' 2010-2012 National Intimate Partner and Sexual Violence Survey (NISVS) indicated that over 36% of woman (1 in 3) and 17% of men (1 in 6) have experienced sexual violence involving physical contact at some point in their lives; the numbers vary by state, from 29.5% to 47.5% for women and 10.4% to 29.3% for men (Smith et al., 2017).



In developing informed responses to a problem, it helps to know for whom it is a problem. Practitioners, program administrators, and policy decision makers may not be aware that the problem of sexual violence is so prevalent, or that men are victimized at worrisome rates, as well as women. It is also helpful to know how the problem of interest might interface with other problems. For example, the interface between perpetrating sexual assault and alcohol use was examined in a study of college men (Testa & Cleveland, 2017). The study investigators determined that frequently attending parties and bars was associated with a greater probability of perpetrating sexual assault. Thus, epidemiological research helps answer questions about the scope and magnitude of a problem, as well as how it relates to other issues or factors, which can then inform next steps in research to address the problem.

Etiology questions. Etiology research tests theories and hypotheses about the origins and natural course of a problem or phenomenon. This includes answering questions about factors that influence the appearance or course of a problem—these may be factors that mediate or moderate the phenomenon's development or progression (e.g., demographic characteristics, co-occurring problems, or other environmental processes). To continue with our intimate partner violence example, multiple theories are presented in the literature concerning the etiology of intimate partner violence perpetration—theories also exist concerning the etiology of being the target of intimate partner violence (Begun, 2003). Perpetration theories include:

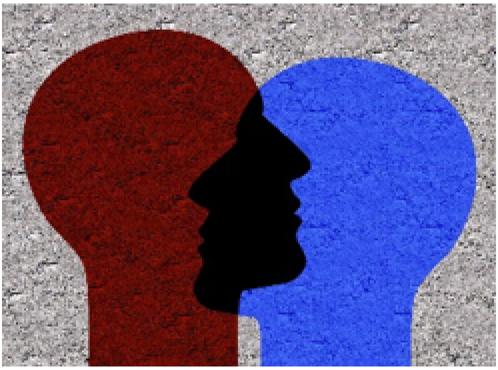
- personality/character traits
- biological/hereditary/genetic predisposition
- social learning/behavior modeling
- social skills
- self-esteem
- cultural norms (Begun, 2003, p. 642).

Evidence supporting each of these theories exists, to some degree; each theory leads to the development of a different type of prevention or intervention response. The "best" interventions will be informed by theories with the strongest evidence or will integrate elements from multiple evidence-supported theories.

Etiology research is often about understanding the mechanisms underlying the phenomena of interest. The questions are "how" questions—how does this happen (or not)? For example, scientists asked the question: how do opioid medications (used to manage pain) act on neurons compared to opioids that naturally occur in the brain (Stoeber et al., 2018)? They discovered that opioid medications used to treat pain bind to receptors *inside*n erve cells, which is a quite different mechanism than the conventional wisdom that they behave the same way that naturally occurring (endogenous) opioids do—binding only on the *surface* of nerve cells. Understanding this mechanism opens new options for developing pain relievers that are less- or non-addicting than current opioid medicines like morphine and oxycodone. Once these mechanisms of change are understood, interventions can be developed, then tested through intervention research approaches.

Intervention Research. Interventions are designed around identified needs: epidemiology research helps to support intervention design by identify the needs. Epidemiology research also helps identify theories concerning the causes and factors affecting social work problems. Intervention development is further supported by later theory-testing and etiology research. However, developing an intervention is not sufficient: interventions need to be tested and evaluated to ensure that they are (1) safe, (2) effective, and (3) cost-efficient to deliver. This is where ***intervention research*** comes into play. Consider the example of Motivational Interviewing (MI) approaches to addressing client ambivalence about engaging in a behavior change effort. Early research concerning MI addressed questions about its effectiveness. For example, a meta-analytic review reported that "MI should be considered as a treatment for adolescent substance abuse" because the evidence demonstrated small, but significant effect sizes, and that the treatment gains were retained over time (Jensen et al., 2011). Subsequently, when its safety and effectiveness were consistently demonstrated through this kind of evidence, investigators assessed MI as cost-efficient or cost-effective. For example, MI combined with providing feedback was demonstrated to be cost-effective in reducing drinking among college students who engaged in heavy drinking behavior (Cowell et al., 2012).

Intervention research not only is concerned with the outcomes of delivering an intervention, but may also address the ***mechanisms of change*** through which an intervention has its effects–not only *what changes* happen, but *how* they happen. For example, investigators are exploring *how* psychotherapy works, moving beyond demonstrating *that* it works (Ardito & Rabellino, 2011; Kazdin, 2007; Wampold, 2015). One mechanism that has garnered attention is the role of therapeutic alliance–the relationships, bonds, and interactions that occur in the context of treatment–on treatment outcomes.



Therapeutic alliance is one common factor identified across numerous types of effective psychotherapeutic approaches (Wampold, 2015). Authors summarizing a number of studies about therapeutic alliance and its positive relationship to treatment outcomes concluded that the quality of therapeutic alliance may be a more powerful predictor of positive outcome than is the nature or type of intervention delivered (Ardito & Rabellino, 2011). However, it is important to determine the extent to which (a) therapeutic alliance enhances clients' symptom improvement, (b) gradual improvements in symptoms lead to enhanced therapeutic alliance, or (c) the relationship between therapeutic alliance and symptom improvement are iterative—they go back and forth, influencing each other over time (Kazdin, 2007).

Implementation Science. Social work and other disciplines have produced a great deal of evidence about "what works" for intervening around a great number of social work problems. Unfortunately, many best practices with this kind of evidence support are slow to become common practices. ***Implementation science*** is about understanding facilitators and barriers to these evidence-supported interventions becoming adopted into routine practice: characteristics of the interventions themselves, conditions and processes operating in the organizations where interventions are implemented, and factors external to these organizations all influence practitioners' adoption of evidence supported interventions.

Even under optimal internal organizational conditions, implementation can be undermined by changes in organizations' external environments, such as fluctuations in funding, adjustments in contracting practices, new technology, new legislation, changes in clinical practice guidelines and recommendations, or other environmental shifts" (Birken, et al, 2017).

Research for/about Research. In addition, social work investigators engage in research that is specifically about scientific methodology. This is where advances in measurement, participant recruitment and retention, and data analysis emerge. The results of these kinds of research studies are used to improve the research in basic, intervention, and implementation research. Later in the course you will see some of these products in action as we learn about best practices in research and evaluation methodology. Here are a few examples related to measurement methods:

- Concept mapping to assess community needs of sexual minority youth (Davis, Saltzburg, & Locke, 2010)
- Field methodologies for measuring college student drinking in natural environments (Clapp et al., 2007)
- Intergenerational contact measurement (Jarrott, Weaver, Bowen, & Wang, 2018)
- Perceived Social Competence Scale-II (Anderson-Butcher et al., 2016)
- Safe-At-Home Instrument to measure readiness to change intimate partner violence behavior (Begun et al., 2003; 2008; Sielski, Begun, & Hamel, 2015)
- Teamwork Scale for Youth (Lower, Newman, & Anderson-Butcher, 2016)

And, here are a few examples related to involving participants in research studies:

- Conducting safe research with at risk populations (Kyriakakis, Waller, Kagotho, & Edmond, 2015)
- Recruitment strategies for non-treatment samples in addiction studies (Subbaraman et al., 2015)
- Variations in recruitment results across Internet platforms (Shao et al., 2015)

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=109

Research Questions

In this section, we take a closer look at research questions and their relationship to the types of research conducted by investigators. It may be easier to understand research questions by first ruling out what are *not*research questions. In that spirit, let's begin with examples of questions where applying research methods will *not* help to find answers:

- *Trauma informed education.*The first issue with this example is obvious: it is not worded as a question. The second is critically important: this is a general topic, it is not a research question. This topic is too vague and broad making it impossible to determine what answers would look like or how to approach finding answers.
- *How is my client feeling about what just happened?*This type of question about an individual is best answered by asking clinical questions of that individual, within the context of the

therapeutic relationship, not by consulting research literature or conducting a systematic research study.

- *Will my community come together in protest of a police-involved shooting incident*?This type of question may best be answered by waiting to see what the future brings. Research might offer a guess based on data from how other communities behaved in the past but cannot predict how groups in individual situations will behave. A better research question might be: What factors predict community protest in response to police-involved shooting incidents?
- *Should I order salad or soup to go with my sandwich*?This type of question is not of general interest, making it a poor choice as a research question. The question might be reframed as a general interest question: Is it healthier to provide salad or soup along with a sandwich? The answer to that researchable question might inform a personal decision.
- *Why divorce is bad for children.*There are two problems with this example. First, it is a statement, not a question, despite starting with the word "why." Second, this question starts out with a biased assumption—that divorce *is*bad for children. Research questions should support unbiased investigation, leading to evidence and answers representative of what exists rather than what someone sets out wanting to prove is the case. A better research question might be: How does divorce affect children?

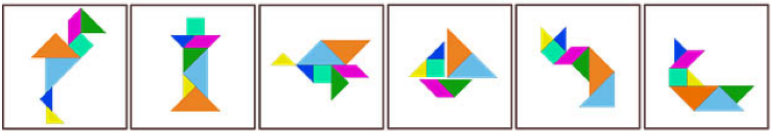
Tuning back to our first example of what is not a research question, consider several possible school social work research questions related to that general topic:

- To what extent do elementary school personnel feel prepared to engage in trauma informed education with their students?
- What are the barriers and facilitators of integrating trauma informed education in middle school?
- Does integrating trauma informed education result in lower rates of suicidal ideation among high school students?
- Is there a relationship between parent satisfaction and the implementation of trauma informed education in their children's schools?
- Does implementing trauma informed education in middle schools affect the rate of student discipline referrals?

What is the difference between these research questions and the earlier "not research" questions? First, research questions are specific. This is an important distinction between identifying a topic of interest (e.g., trauma informed education) and asking a researchable question. For example, the question "How does divorce affect children?" is not a good research question because it remains too broad. Instead, investigators might focus their research questions on one or two specific effects of interest, such as emotional or mental health, academic performance, sibling relationships, aggression, gender role, or dating relationship outcomes.

Related to a question being "researchable" is its feasibility for study. Being able to research a question requires that appropriate data can be collected with integrity. For example, it may not be feasible to study what would happen if every child was raised by two parents, because (a) it is impossible to study every child and (2) this reality cannot ethically be manipulated to systematically explore it. No one can ethically conduct a study whereby children are randomly assigned by study investigators to the compared conditions of being raised by two parents versus being raised by one or no parents. Instead, we settle for observing what has occurred naturally in different families.

Second, "good" research questions are relevant to knowledge building. For this reason, the question about what to eat was not a good research question—it is not relevant to others' knowledge development. Relevance is in the "eye of the beholder," however. A social work researcher may not see the relevance of using a 4-item stimulus array versus a 6-item stimulus array in testing children's memory, but this may be an important research question for a cognitive psychology researcher. It may, eventually, have implications for assessment measures used in social work practice.



Third, is the issue of bias built into research questions. Remembering that investigators are a product of their own developmental and social contexts, what they choose to study and how they choose to study it are socially constructed. An important aspect at the heart of social work research relates to a question's cultural appropriateness and acceptability. To demonstrate this point, consider an era (during the 1950s to early 1970s) when research questions were asked about the negative effects on child development of single-parent, black family households compared to two-parent, white family households in America. This "majority comparison" frame of reference is not culturally appropriate or culturally competent. Today, in social work, we adopt a strengths perspective, and avoid making comparisons of groups against a majority model. For example, we might ask questions like: What are the facilitators and barriers of children's positive development as identified by single parents of diverse racial/ethnic backgrounds? What strengths do African American parents bring to the experience of single-parenting and how does it shape their children's development? What are the similar and different experiences of single-parenting experienced by families of different racial/ethnic composition?

Research Questions versus Research Hypotheses. You have now seen examples of "good" research questions. Take, for example, the last one we listed about trauma informed education:

> *Does implementing trauma informed education in middle schools affect the rate of student discipline referrals*?

Based on a review of literature, practice experience, previous research efforts, and the school's interests, an investigator may be prepared to be even more specific about the research question (see Figure 1-3). Assume that these sources led the investigator to believe that implementing the trauma informed education approach will have the effect of reducing the rate of disciplinary referrals. The investigator may then propose to test the following hypothesis:

> *Implementing trauma informed education in middle schools*

will result in a reduction in the number of student discipline referrals.

The ***research hypothesis*** is a clear statement that can be tested with quantitative data and will either be rejected or not, depending on the evidence. Research hypotheses are predictions about study results—what the investigator expects the results will show. The prediction, or hypothesis, is based on theory and/or other evidence. A study hypothesis is, by definition, quantifiable—the answer lies in numerical data, which is why we do not generally see hypotheses in qualitative, descriptive research reports.

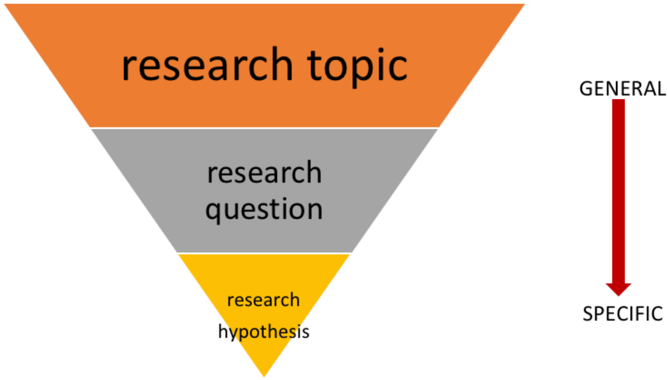
Hypotheses are also specific to one question at a time. Thus, an investigator would need to state and test a second hypothesis to answer the question:

Is there a relationship between parent satisfaction and the implementation of trauma informed education in their children's schools?

The stated hypothesis might be:

Parent satisfaction is higher in middle schools where trauma informed education is implemented.

Figure 1-3. Increasing specificity from research topic to question to hypothesis



Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=109

Social Work Questions

It is difficult to find a simple way to characterize social work research. The National Institutes of Health (NIH) described social work research in the following way:

> *Historically, social work research has focused on studies of the individual, family, group, community, policy and/or organizational level, focusing across the lifespan on prevention, intervention, treatment, aftercare and rehabilitation of acute and chronic conditions, including the effects of policy on social work practice* (OBSSR, 2003, *p.* 5).

For all the breadth expressed in this statement, it reflects only how social work research relates to the health arena–it does not indicate many other domains and service delivery systems of social work influence:

- physical, mental, and behavioral health
- substance misuse/addiction and other addictive behaviors
- education
- workplace
- income/poverty
- criminal justice
- child and family welfare
- housing and food security/insecurity
- environmental social work
- intimate partner, family, and community violence
- and others.

In addition to breadth of topic, social work research is characterized by its biopsychosocial nature. This means that social work researchers not only pursue questions relating to biological, psychological, and social context factors, but also questions relating to their intersections and interactions. Related to this observation is that social work not only addresses questions related to the multiple social system levels, social work also addresses the ways multiple levels intersect and interact (i.e., those levels represented in the NIH statement about individuals, families, groups, communities, organizations, and policy).

It is worth noting that research need not be conducted by social workers to be relevant to social work–many disciplines and professions contribute to the knowledge base which informs social work practice (medicine, nursing, education, occupational therapy, psychology, sociology, criminal justice, political science, economics, and more). Authors of one social work research textbook summarize the relevance issue in the following statement:

> *"To social workers, a relevant research question is one whose answers will have an impact on policies, theories, or practices related to the social work profession" (Grinnell & Unrau, 2014, p. 46).*

Social Work Research Questions and Specific Aims

The kinds of questions that help inform social work practice and policy are relevant to understanding social work problems, diverse populations, social phenomena, or interventions. Most social work research questions can be divided into two general categories: ***background questions*** and ***foreground questions***. The major distinction between these two categories relates to the specific aims that emerge in relation to the research questions.

Background Questions. This type of question is answerable with a fact or set of facts. Background questions are generally simple in structure, and they direct a straightforward search for evidence. This type of question can usually be formulated using the classic 5 question words: who, what, when, where, or why. Here are a few examples of social work background questions related to the topic of fetal alcohol exposure:

- *Who*is at greatest risk of fetal alcohol exposure?
- *What*are the developmental consequences of fetal alcohol exposure?

- *When*in gestation is the risk of fetal alcohol exposure greatest?
- *Where*do women get information about the hazards of drinking during pregnancy?
- *Why*is fetal alcohol exposure (FAE) presented as a spectrum disorder, different from fetal alcohol syndrome (FAS)?

These kinds of questions direct a social worker to review literature about human development, human behavior, the distribution of the problem across populations, and factors that determine the nature of a specific social work problem like fetal exposure to alcohol. Where the necessary knowledge is lacking, investigators aim to explore or describe the phenomenon of interest. Many background questions can be answered by epidemiology or etiology evidence.

Foreground Questions. This type of question is more complex than the typical background question. Foreground questions typically are concerned with making specific choices by comparing or evaluating options. These types of questions required more specialized evidence and may lead to searching different types of resources than would be helpful for answering background questions. Foreground questions are dealt with in greater detail in our second course, SWK 3402 which is about understanding social work interventions. A quick foreground question example related to the fetal exposure to alcohol topic might be:

> *Which is the best tool for screening pregnant women for alcohol use with the aim of reducing fetal exposure, the* T-ACE, TWEAK, *or* AUDIT?

This type of question leads the social worker to search for evidence that compares different approaches. These kinds of evidence are usually found in comparative reviews, or require the practitioner to conduct a review of literature, locating individual efficacy and effectiveness studies. Where knowledge is found to be lacking, investigators aim to experiment with different approaches or interventions.

Three Question Types and Their Associated Research Aims

Important distinctions exist related to different types of background questions. Consider three general categories of questions that social workers might ask about populations, problems, and social phenomena: exploratory, descriptive, and explanatory. The different types of questions matter because the nature of the research questions determines the specific aims and most appropriate research approaches investigators apply in answering them.

Exploratory Research Questions. Social workers may find themselves facing a new, emerging problem where there is little previously developed knowledge available—so little, in fact, that it is premature to begin asking any more complex questions about causes or developing testable theories. ***Exploratory research questions***open the door to beginning understanding and are basic; answers would help build the foundation of knowledge for asking more complex descriptive and explanatory questions. For example, in the early days of recognition that HIV/AIDS was emerging as a significant public health problem, it was premature to jump to questions about how to treat or prevent the problem. Not enough was known about the nature and scope of the problem, for whom it was a problem, how the problem was transmitted, factors associated with risk for exposure, what factors influenced the transition from HIV exposure to AIDS as a disease state, and what issues or problems might co-occur along with either HIV exposure or AIDS. In terms of a knowledge evolution process, a certain degree of exploration had to occur before intervention strategies for prevention and treatment could be developed, tested, and implemented.

In 1981, medical providers, public health officials, and the Centers for Disease Control and Prevention (CDC) began to circulate and publish observations about a disproportionate, unexpectedly high incidence rate of an unusual pneumonia and Kaposi's sarcoma appearing in New York City and San Francisco/California among homosexual men (Curran, & Jaffe, 2011). As a result, a task force was formed and charged with conducting an epidemiologic investigation of this outbreak; "Within 6 months, it was clear that a new, highly concentrated epidemic of life threatening illness was occurring in the United States" (Curran & Jaffe, 2011, p. 65). The newly recognized disease was named for its symptoms: acquired immune deficiency syndrome, or AIDS. Exploratory research into the social networks of 90 living patients in 10 different cities indicated that 40 had a sexual contact link with another member of the 90-patient group (Auerbach, Darrow, Jaffe, & Curran, 1984). Additionally, cases were identified among persons who had received blood products related to their having hemophilia, persons engaged in needle sharing during substance use, women who had sexual contact with a patient, and infants born to exposed women. Combined, these pieces of information led to an understanding that the causal infectious factor (eventually named the human immunodeficiency virus, HIV) was transmitted by sexual contact, blood, and placental connection. This, in turn, led to knowledge building activities to develop both preventive and treatment strategies which could be implemented and studied. Social justice concerns relate to the slow rate at which sufficient resources were committed for evolving to the point of effective solutions for saving lives among those at risk or already affected by a heavily stigmatized problem.

The exploratory research approaches utilized in the early HIV/AIDS studies were both qualitative and quantitative in nature. Qualitative studies included in-depth interviews with identified patients—anthropological and public health interviews about many aspects of their living, work, and recreational environments, as well as many types of behavior. Quantitative studies included comparisons between homosexually active men with and without the diseases of concern. In addition, social network study methods combined qualitative and quantitative approaches. These examples of early exploratory research supported next steps in knowledge building to get us to where we are today. "Today, someone diagnosed with HIV and treated before the disease is far advanced can live nearly as long as someone who does not have HIV" (hiv.gov). While HIV infection cannot (yet) be "cured," it can be controlled and managed as a chronic condition.

Descriptive Research Questions. Social workers often ask for descriptions about specific populations, problems, processes, or phenomena. ***Descriptive research questions*** might be expressed in terms of searching to create a profile of a group or population, create categories or types (typology) to describe elements of a population, document facts that confirm or contradict existing beliefs about a topic or issue, describe a process, or identify steps/stages in a sequential process (Grinnell & Unrau, 2014). Investigators may elect to approach the descriptive question using qualitative methods that result in a rich, deep description of certain individuals' experiences or perceptions (Yegidis, Weinbach, & Meyers, 2018). Or, the descriptive question might lead investigators to apply quantitative methods, assigning numeric values, measuring variables that describe a population, process, or situation of interest. In descriptive research, investigators do not manipulate or experiment with the variables; investigators seek to describe what naturally occurs (Yegidis, Weinbach, & Meyers, 2018). As a result of studies answering descriptive questions, tentative theories and hypotheses may be generated.

Here are several examples of descriptive questions.

- How do incarcerated women feel about the option of medication-assisted treatment for substance use disorders?
- What barriers to engaging in substance misuse treatment do previously incarcerated persons experience during community reentry?
- How often do emerging adults engage in binge drinking in different drinking contexts (e.g., bars, parties, sporting events, at home)?
- What percent of incarcerated adults experience a substance use disorder?
- What is the magnitude of racial/ethnic disparities in access to treatment for substance use disorders?

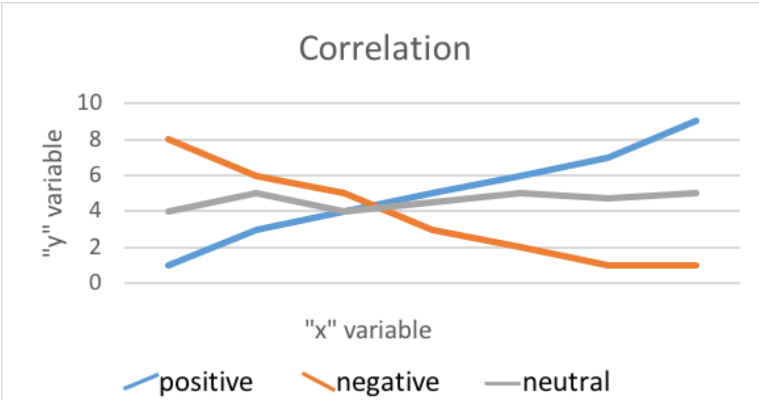
- Who provides supervision or coordination of services for aging adults with intellectual or other developmental disabilities?
- What is the nature of the debt load among students in doctoral social work programs?

An example of descriptive research, derived from a descriptive question, is represented in an article where investigators addressed the question: How is the topic of media violence and aggression reported in print media (Martins et al., 2013)? This question led the investigators to conduct a qualitative content analysis, resulting in a description showing a shift in tone where earlier articles (prior to 2000) emphasized the link as a point of concern and later articles (since 2000) assumed a more neutral stance.

Correlational Research Questions. One important type of descriptive question asks about relationships that might exist between variables—looking to see if *variable x* and *variable y* are associated or correlated with each other. This is an example of a correlational research question; it does not indicate whether “x” causes “y” or “y” causes “x”, only whether these two are related. Consider again the topic of exposure to violence in the media and its relationship to aggression. A descriptive question asked about the existence of a relationship between exposure to media violence (*variable x*) and children’s expression of aggression (*variable y*). Investigators reported one study of school-aged children, examining the relationship between exposure to three types of media violence (television, video games, and movies/videos) and three types of aggression (verbal, relational, and physical; Gentile, Coyne, & Walsh, 2011). The study investigators reported that media violence exposure was, indeed, correlated with all three types of aggressive behavior (and less prosocial behavior, too).

For a positive correlation (the blue line), as the value of the “x” variable increases, so does the value of the “y” variable (see Figure 1-4 for a general demonstration). An example might be as age or grade in school increases (“x”), so does the number of preadolescent, adolescent, and emerging adults who have used alcohol (“y”). For a negative correlation (the orange line), as the value of the “x” variable increases, the value of the “y” variable decreases. An example might be as the number of weeks individuals are in treatment for depression symptoms (“x”), the reported depression symptoms decreases (“y”). The neutral of non-correlation line (grey) means that the two variables, “x” and “y” do not have an association with each other. For example, number of years of teachers’ education (“x”) might be unrelated to the number of students dropping out of high school (“y”).

Figure 1-4. Depicting positive, negative, and neutral correlation lines

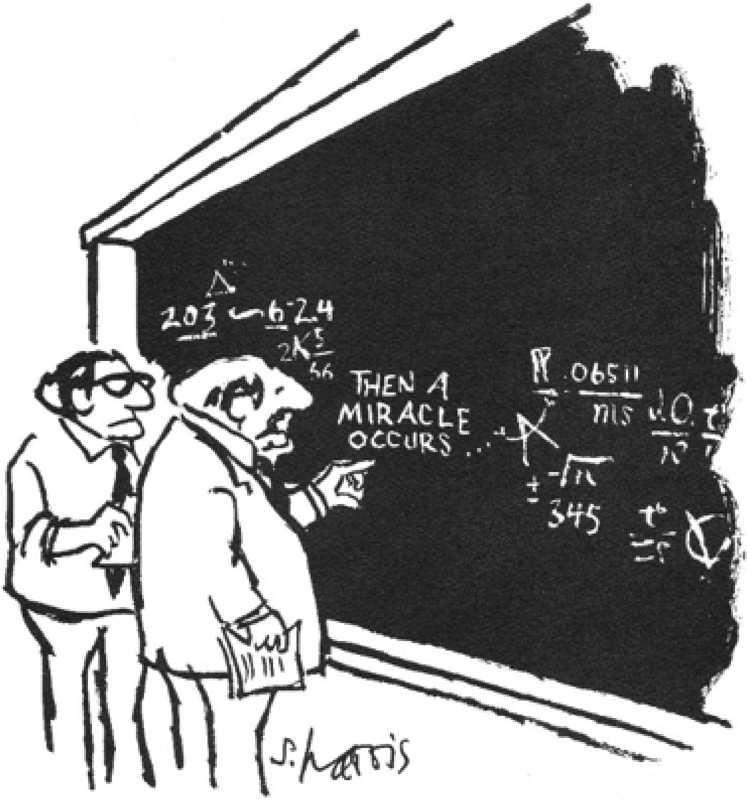


Descriptive correlational studies are sometimes called comparison studies because the descriptive question is answered by comparing groups that differ on one of the variables (low versus high media violence exposure) to see how they might differ on the other variable (aggressive behavior).

Explanatory Research Questions. To inform the design of evidence-informed interventions, social workers need answers to questions about the nature of the relationships between potentially influential factors or variables. An ***explanatory research question*** might be mapped as: Does variable *x* cause, lead to or prevent changes in variable *y* (Grinnell & Unrau, 2014)? These types of questions often test theory related to etiology.

Comparative research might provide information about a relationship between variables. For example, the difference in outcomes between persons experiencing a substance use disorder and have been incarcerated compared to others with the same problem but have not been incarcerated may be related to their employability and ability to generate a living-wage income for themselves and their families. However, to develop evidence-informed interventions, social workers need to know that variables are not only related, but that one variable actually plays a causal role in relation to the other. Imagine, for example, that evidence demonstrated a significant relationship between adolescent self-esteem and school performance. Social workers might spend a great deal of effort developing interventions to boost self-esteem in hopes of having a positive impact on school performance. However, what if self-esteem comes from strong school performance? The self-esteem intervention efforts will not likely have the desired effect on school performance. Just because research demonstrates a significant relationship between two variables does not mean that the research has demonstrated a *causal* relationship between those variables. Investigators need to be cautious about the extent to which their study designs can support drawing conclusions about causality; anyone reviewing research reports also needs to be alert to where causal conclusions are properly and improperly drawn.

The questions that drive intervention and evaluation research studies are explanatory in nature: does the intervention (x) have a significant impact on outcomes of interest (y)? Another type of explanatory question related to intervention research concerns the mechanisms of change. In other words, not only might social workers be interested to find out *what* outcomes or changes can be attributed to an intervention, they may also be interested to learn *how* the intervention causes those changes or outcomes.



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

Cartoon used with permission, ©Sidney Harris

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=109

Chapter Summary

In this chapter, you learned about different aspects of the knowledge building process and where different types of research questions might fit into the big picture. No single research study covers the entire spectrum; each study contributes a piece of the puzzle as a whole. Research questions come in many different forms and several different types. What is important to recall as we move through the remainder of the course is that the decisions investigators make about research approaches, designs, and procedures all start with the nature of the question being asked. And, the questions being asked are influenced by multiple factors, including what is previously known and remains unknown, the culture and context of the questioners, and what theories they have about what is to be studied. That leads us to the next chapter.

Module 2 Chapter 2: The Link Between Theory, Research, and Social Justice

Theory has been mentioned several times in Chapter 1 discussions. In this chapter, we explore the relationship between theory and research, paying particular attention to how theory and research relate to promoting social justice.

In this chapter you will read about:

- Why theory matters to social work
- How theory and research relate to social justice

The Significance of Theory

It is helpful to begin with thinking about what ***theory*** is. Theory is defined as a belief, idea, or set of principles that explain something–the set of principles may be organized in a complex manner or may be quite simple in nature. Theories are not facts; they are conjectures and predictions that need to be tested for their goodness-of-fit with reality. A scientific theory is an explanation supported by empirical evidence. Therefore, scientific theory is based on careful, reasonable examination of facts, and theories are tested and confirmed for their ability to explain and predict the phenomena of interest.

Theory is central to the development of social work interventions, as it determines the nature of our solutions to identified problems. Consider an example whereby programmatic and social policy responses might be influenced by the way that a problem like teen pregnancy is defined and the theories about the problem. In Table 2-1 you can see different definitions or theories of the problem on the left, and the logical responses on the right. In many cases, the boxes on the right can also be supplemented with content from other boxes addressing other problem definitions or theories.

Table 2-1. Analysis of teen pregnancy: How defining a problem determines responses

Problem definition/ theory	Policy response option
Teen pregnancy is not a problem	No need to respond
Teen girls are having sex	Keep them otherwise occupied allowing no opportunity; keep them separated from potential sex partners (including sexual abuse); teach/ preach "Just Say No!"
Teen girls are getting pregnant	Make birth control available, accessible, affordable, and palatable; keep them separated from boys and men capable of impregnating them (including sexual abuse)
Teen girls are having babies	Terminate teen pregnancies; prevent pregnancy—see above
Babies born to teen mothers are less healthy	Provide resources for healthy prenatal development; prevent pregnancy/having babies—see above
Teen mothers have poorer health than other teens, other mothers	Provide health-promoting resources to teen mothers; prevent pregnancy/having babies—see above
Teen mothers are raising babies inadequately	Provide resources to promote positive parenting by teen mothers (or all parents); place babies with other parents (other family members or foster/ adoptive families); prevent pregnancy/having babies—see above
Being a teen mother interrupts her education and future employability/ income potential	Provide resources to promote positive educational and vocational/income producing outcomes for teen mothers; prevent teen mothers raising babies, prevent pregnancy/having babies—see above

Hopefully, through this example, you can see how the way we define a problem and our theories about its causes determines the types of solutions we develop. Solutions are also dependent on whether they are feasible, practical, reasonable, ethical, culturally appropriate, and politically acceptable.

Theory, Research, and Social Justice

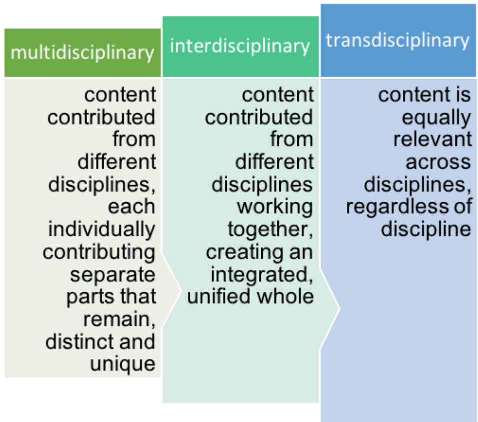
Theory is integral to research and research is integral to theory. Theory guides the development of many research questions and research helps generate new theories, as well as determining whether support for theories exists. What is important to remember is that theory is not fact: it is a belief about how things work; theory is a belief or "best guess" that awaits the support of empirical evidence.

> *"The best theories are those that have been substantiated or validated in the real world by research studies"* (Dudley, 2011, p. 6).

Countless theories exist, and some are more well-developed and well-defined than others. More mature theories have had more time and effort devoted to supporting research, newer theories may be more tentative as supporting evidence is being developed. Theories are sometimes disproven and need to be scrapped completely or heavily overhauled as new research evidence emerges. And, exploratory research leads to the birth of new theories as new phenomena and questions arise, or as practitioners discover ways that existing theory does not fit reality.

Examples of theories and theoretical models with which you may have become familiar in other coursework are developed, tested, and applied in research from multiple disciplines, including social work. You may be familiar with the concepts of ***multidisciplinary*** and ***interdisciplinary*** practice, research and theory, but you also might be interested to learn the concept of ***transdisciplinary*** research and theory. The social work profession engages in all three, as described in Figure 2-1.

Figure 2-1. Comparison of multidisciplinary, interdisciplinary, and transdisciplinary as concepts



An example of transdisciplinarity is demonstrated in motivational interviewing. The principles and practice of motivational interviewing skills transcends disciplines—it is relevant and effective regardless of a practitioner's discipline and regardless of which discipline is conducting research that provides supporting evidence. An example of interdisciplinarity is when social workers, nurses, pediatricians, occupational therapists, respiratory therapists, and physical therapists work together to create a system for reducing infants' and young children's environmental exposure to heavy metal contamination in their home, childcare, recreational, food and water (e.g., lead and mercury). An example of multidisciplinarity is when a pediatrician, nurse, occupational therapist, social worker, and physical therapist each deliver discipline-specific services to children with intellectual disabilities.

Here are examples of theories that you may have encountered or eventually will encounter in your career:

- behavioral theory

- cognitive theory
- conflict theory
- contact theory (groups)
- critical race theory
- developmental theory (families)
- developmental theory (individuals)
- feminist theory
- health beliefs model
- information processing theory
- learning theory
- lifecourse or lifespan theory
- neurobiology theories
- organizational theory
- psychoanalytic theory
- role theory
- social capital model
- social ecological theory
- social learning theory
- social network theory
- stress vulnerability model
- systems theory
- theory of reasoned action/planned behavior
- transtheoretical model of behavior change

Social work practitioners generate new theories in the field all the time, but these theories are rarely documented or systematicallyl tested. Applying systematic research methods to test these practice-generated theories can help expand the social work knowledge base. Well-developed theories must be testable through the application of research methods. Furthermore, they both supplement and complement other theories that have the support of strong evidence behind them. In addition to these points, for theories to be relevant to social work, they need to:

- have practice implications at one or more level of intervention–suggest principles of action;
- be responsive to human diversity;
- contribute to the promotion of social justice (Dudley, 2011).

Social Justice and The Grand Challenges for Social Work. In 2016, the American Academy of Social Work & Social Welfare (AASWSW) rolled out a set of 12 initiatives, challenging the social work profession to develop strategies for having a significant impact on a broad set of problems challenging the nation: "their introduction truly has the potential to be a defining moment in the history of our profession" (Williams, 2016, p. 67). The Grand Challenges directly relate to content presented in the Preamble to the Code of Ethics of the National Association of Social Workers (2017):

> *The primary mission of the social work profession is to enhance human well-being and help meet the basic human needs of all people, with particular attention to the needs and empowerment of people who are vulnerable, oppressed, and living in poverty. A historic and defining feature of social work is the profession's dual focus on individual well-being in a social context and the well-being of society. Fundamental to social work is attention to the environmental forces that create, contribute to, and address problems in living (p. 1).*

The ambitious Grand Challenges (http://aaswsw.org/grand-challenges-initiative/12-challenges/) call for a synthesis of research and evidence generating endeavors, social work education, and social work practice (at all levels), with promoting social justice and transforming society at the forefront of attention. None of the Grand Challenges can be achieved without collaboration between social work researchers, practitioners, key stakeholders/constituents, policy makers, and members of other professions and disciplines (see Table 2-2). Each of the papers published under the 12 umbrella challenges not only reviews evidence related to the challenge, but also identifies a research and action agenda for promoting change and achieving the stated challenge goals.

Table 2-2. 12 Grand Challenges for Social Work

Grand Challenge	Topics Under the Challenges
Ensure healthy development for all youths	• Prevention of schizophrenia and severe mental illness • Unleashing the power of prevention
Close the health gap	• Reducing and preventing alcohol misuse and its consequences • Health equity: Eradicating health inequalities for future generations • Strengthening health care systems: Better health across America
Stop family violence	• Ending gender-based violence • Safe children: Reducing severe and fatal maltreatment
Advance long and productive lives	• Increasing productive engagement in later life • Increasing productive engagement early in life: Civic and volunteer service as a pathway to development
Eradicate social isolation	• Social isolation presents a grand challenge to social work
End homelessness	• End homelessness
Create social responses to a changing environment	• Strengthening the social response to the human impacts of environmental change

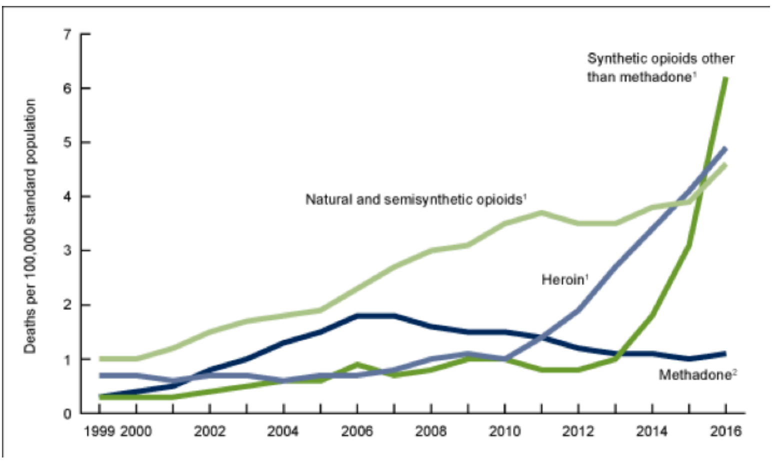
Grand Challenge	Topics Under the Challenges
Harness technology for social good	• Practice innovation through technology in the digital age • Harnessing big data for social good
Promote smart decarceration	• From mass incarceration to smart decarceration
Reduce extreme economic inequality	• Reversing extreme inequality
Build financial capability for all	• Financial capability and asset building for all
Achieve equal opportunity and justice	• Promoting equality by addressing social stigma • Achieving equal opportunity and justice: The integration of Latina/o immigrants into American society • Increasing success for African American children and youth • Fair housing and inclusive communities

Theory and its related research is presented in the literature of social work and other disciplines. Being able to locate the relevant literature is an important skill for social work professionals and researchers to master. The next chapter introduces some basic principles about identifying literature that informs us about theory and the research that leads to the development of new theories and the testing of existing theories.

Using Data to Support Social Justice Advocacy

It is one thing to collect data, statistics, and information about the dimensions of a social problem; it is another to apply those data, statistics, and information in action to promote social change around an identified social justice cause. Advocacy is one tool or role important to the social work profession since its earliest days. Data and empirical evidence should routinely support social workers' social justice advocacy efforts at the micro level—when working with specific client systems (case advocacy). Social justice advocacy also is a macro-level practice (cause advocacy), one that often involves the use of data to raise awareness about a cause, establish change goals, and evaluate the impact of change efforts. Consider, for example, the impact of data on social justice advocacy related to opioid misuse and addiction across the United States. Data regarding the sharp, upward trend in opioid-related deaths have had a powerful impact on public awareness, having captivated the attention of mass media outlets. These are data from the Centers for Disease Control and Prevention (CDC) for the period 1999 to 2016 (see Figure 2-2). The death rate from opioid overdose increased remarkably in the most recent years depicted, including heroin, natural and semisynthetic (morphine, codeine, hydrocodone, oxycodone), and synthetic opioids (fentanyl, fentanyl analogs, tramadol); the rate declined somewhat for deaths related to methadone, a highly controlled prescription medication used to treat opioid addiction (Hedegaard, Warner, & Miniño, 2017). In 2016 the overdose death rate was more than triple the 1999 rate and surpassed all other causes of death (including homicide and automobile crashes) for persons aged 50 or younger (Vestal, 2018). The good news is that the rates dropped significantly across 14 states during 2017 as new policy approaches were implemented—including more widespread use and access to naloxone (an opioid overdose reversal drug). The bad news is that the opioid overdose death rates increased significantly across 5 states and the District of Columbia as fentanyl and related drugs became integrated into the illicit drug supply: "Nationally, the death toll is still rising" (Vestal, 2018, p. 1). The Ohio Department of Health, with Ohio being one of the 5 states with more than a 30% increase in opioid overdose deaths, now recommends that naloxone be used more widely in overdose situations, whether or not opioids are known to be involved.

Figure 2-2. Opioid drug overdose death rate trends 1999-2016, by opioid category



Having data of this sort available to argue the urgency of a cause is of tremendous value. Imagine the potential impact of data on social issues such as child maltreatment, intimate partner violence, human trafficking, and death by suicide. Imagine also the potential impact of data concerning disparities in health, incarceration, mental and traumatic stress disorders, and educational or employment achievement across members of different racial/ethnic, age, economic, and national origin groups. This type of evidence has the potential to be a powerful element in the practice of social justice advocacy.

Follow links below to specific AASWSW Grand Challenges and consider the ways that research evidence is being used to advocate for at least one of the following social justice causes:

- Health Equity: Eradicating Health Inequalities for Future Generations
- From Mass Incarceration to Smart Decarceration
- Safe Children: Reducing Severe and Fatal Maltreatment
- Ending Gender Based Violence: A Grand Challenge for Social Work
- Prevention of Schizophrenia and Severe Mental Illness
- Increasing Productive Engagement in Later Life
- Strengthening the Social Responses to the Human Impacts of Environmental Change

Module 2 Chapter 3: What is Empirical Literature & Where can it be Found?

In Module 1, you read about the problem of pseudoscience. Here, we revisit the issue in addressing how to locate and assess scientific or ***empirical literature***. In this chapter you will read about:

- distinguishing between what IS and IS NOT empirical literature
- how and where to locate empirical literature for understanding diverse populations, social work problems, and social phenomena.

Probably the most important take-home lesson from this chapter is that one source is not sufficient to being well-informed on a topic. It is important to locate multiple sources of information and to critically appraise the points of convergence and divergence in the information acquired from different sources. This is especially true in emerging and poorly understood topics, as well as in answering complex questions.

What Is Empirical Literature

Social workers often need to locate valid, reliable information concerning the dimensions of a population group or subgroup, a social work problem, or social phenomenon. They might also seek information about the way specific problems or resources are distributed among the populations encountered in professional practice. Or, social workers might be interested in finding out about the way that certain people experience an event or phenomenon. Empirical literature resources may provide answers to many of these types of social work questions. In addition, resources containing data regarding ***social indicators*** may also prove helpful. Social indicators are the “facts and figures” statistics that describe the social, economic, and psychological factors that have an impact on the well-being of a community or other population group.The United Nations (UN) and the World Health Organization (WHO) are examples of organizations that monitor social indicators at a global level: dimensions of population trends (size, composition, growth/loss), health status (physical, mental, behavioral, life expectancy, maternal and infant mortality, fertility/child-bearing, and diseases like HIV/AIDS), housing and quality of sanitation (water supply, waste disposal), education and literacy, and work/income/unemployment/economics, for example.

Three characteristics stand out in empirical literature compared to other types of information available on a topic of interest: systematic observation and methodology, objectivity, and transparency/replicability/reproducibility. Let’s look a little more closely at these three features.

Systematic Observation and Methodology. The hallmark of empiricism is "repeated or reinforced observation of the facts or phenomena" (Holosko, 2006, p. 6). In empirical literature, established research methodologies and procedures are systematically applied to answer the questions of interest.

Objectivity. Gathering "facts," whatever they may be, drives the search for empirical evidence (Holosko, 2006). Authors of empirical literature are expected to report the facts as observed, whether or not these facts support the investigators' original hypotheses. Research integrity demands that the information be provided in an objective manner, reducing sources of investigator bias to the greatest possible extent.

Transparency and Replicability/Reproducibility. Empirical literature is reported in such a manner that other investigators understand precisely what was done and what was found in a particular research study–to the extent that they could replicate the study to determine whether the findings are reproduced when repeated. The outcomes of an original and replication study may differ, but a reader could easily interpret the methods and procedures leading to each study's findings.

What is NOT Empirical Literature

By now, it is probably obvious to you that literature based on "evidence" that is not developed in a systematic, objective, transparent manner is not empirical literature. On one hand, non-empirical types of professional literature may have great significance to social workers. For example, social work scholars may produce articles that are clearly identified as describing a new intervention or program without evaluative evidence, critiquing a policy or practice, or offering a tentative, untested theory about a phenomenon. These resources are useful in educating ourselves about possible issues or concerns. But, even if they are informed by evidence, they are not empirical literature. Here is a list of several sources of information that do not meet the standard of being called empirical literature:

- your course instructor's lectures
- political statements
- advertisements
- newspapers & magazines (journalism)
- television news reports & analyses (journalism)
- many websites, Facebook postings, Twitter tweets, and blog postings
- the introductory literature review in an empirical article
- textbooks

You may be surprised to see the last two included in this list. Like the other sources of information listed, these sources also might lead you to look for evidence. But, they are not themselves sources of evidence. They may summarize existing evidence, but in the process of summarizing (like your instructor's lectures), information is transformed, modified, reduced, condensed, and otherwise manipulated in such a manner that you may not see the entire, objective story. These are called secondary sources, as opposed to the original, primary source of evidence. In relying solely on secondary sources, you sacrifice your own critical appraisal and thinking about the original work—you are "buying" someone else's interpretation and opinion about the original work, rather than developing your own interpretation and opinion. What if they got it wrong? How would you know if you did not examine the primary source for yourself? Consider the following as an example of "getting it wrong" being perpetuated.

Example: Bullying and School Shootings. One result of the heavily publicized April 1999 school shooting incident at Columbine High School (Colorado), was a heavy emphasis placed on bullying as a causal factor in these incidents (Mears, Moon, & Thielo, 2017), "creating a powerful master narrative about school shootings" (Raitanen, Sandberg, & Oksanen, 2017, p. 3). Naturally, with an identified cause, a great deal of effort was devoted to anti-bullying campaigns and interventions for enhancing resilience among youth who experience bullying. However important these strategies might be for promoting positive mental health, preventing poor mental health, and possibly preventing suicide among school-aged children and youth, it is a mistaken belief that this can prevent school shootings (Mears, Moon, & Thielo, 2017). Many times the accounts of the perpetrators having been bullied come from potentially inaccurate third-party accounts, rather than the perpetrators themselves; bullying was not involved in all instances of school shooting; a perpetrator's perception of being bullied/persecuted are not necessarily accurate; many who experience severe bullying do not perpetrate these incidents; bullies are the least targeted shooting victims; perpetrators of the shooting incidents were often bullying others; and, bullying is only one of many important factors associated with perpetrating such an incident (Ioannou, Hammond, & Simpson, 2015; Mears, Moon, & Thielo, 2017; Newman &Fox, 2009; Raitanen, Sandberg, & Oksanen, 2017). While mass media reports deliver bullying as a means of explaining the inexplicable, the reality is not so simple: "The connection between bullying and school shootings is elusive" (Langman, 2014), and "the relationship between bullying and school shooting is, at best, tenuous" (Mears, Moon, & Thielo, 2017, p. 940). The point is, when a narrative becomes this publicly accepted, it is difficult to sort out truth and reality without going back to original sources of information and evidence.

What May or May Not Be Empirical Literature: Literature Reviews

Investigators typically engage in a review of existing literature as they develop their own research studies. The review informs them about where knowledge gaps exist, methods previously employed by other scholars, limitations of prior work, and previous scholars' recommendations for directing future research. These reviews may appear as a published article, without new study data being reported (see Fields, Anderson, & Dabelko-Schoeny, 2014 for example). Or, the literature review may appear in the introduction to their own empirical study report. These literature reviews are not considered to be empirical evidence sources themselves, although they may be based on empirical evidence sources. One reason is that the authors of a literature review may or may not have engaged in a *systematic*search process, identifying a full, rich, multi-sided pool of evidence reports.

There is, however, a type of review that applies systematic methods and is, therefore, considered to be more strongly rooted in evidence: the ***systematic review***.

Systematic review of literature. A systematic reviewis a type of literature report where established methods have been systematically applied, objectively, in locating and synthesizing a body of literature. The systematic review report is characterized by a great deal of transparency about the methods used and the decisions made in the review process, and are replicable. Thus, it meets the criteria for empirical literature: systematic observation and methodology, objectivity, and transparency/reproducibility. We will work a great deal more with systematic reviews in the second course, SWK 3402, since they are important tools for understanding interventions. They are somewhat less common, but not unheard of, in helping us understand diverse populations, social work problems, and social phenomena.

Locating Empirical Evidence

Social workers have available a wide array of tools and resources for locating empirical evidence in the literature. These can be organized into four general categories.

Journal Articles. A number of professional journals publish articles where investigators report on the results of their empirical studies. However, it is important to know how to distinguish between empirical and non-empirical manuscripts in these journals. A key indicator, though not the only one, involves a ***peer review process***. Many professional journals require that manuscripts undergo a process of peer review before they are accepted for publication. This means that the authors' work is shared with scholars who provide feedback to the journal editor as to the quality of the submitted manuscript. The editor then makes a decision based on the reviewers' feedback:

- Accept as is
- Accept with minor revisions
- Request that a revision be resubmitted (no assurance of acceptance)
- Rejection

When a “revise and resubmit” decision is made, the piece will go back through the review process to determine if it is now acceptable for publication and that all of the reviewers’ concerns have been adequately addressed. Editors may also reject a manuscript because it is a poor fit for the journal, based on its mission and audience, rather than sending it for review consideration.

*Indicators of journal relevance.*Various journals are not equally relevant to every type of question being asked of the literature. Journals may overlap to a great extent in terms of the topics they might cover; in other words, a topic might appear in multiple different journals, depending on how the topic was being addressed. For example, articles that might help answer a question about the relationship between community poverty and violence exposure might appear in several different journals, some with a focus on poverty, others with a focus on violence, and still others on community development or public health. Journal titles are sometimes a good starting point but may not give a broad enough picture of what they cover in their contents.

In focusing a literature search, it also helps to review a journal's mission and target audience. For example, at least four different journals focus specifically on poverty:

- *Journal of Children & Poverty*
- *Journal of Poverty*
- *Journal of Poverty and Social Justice*
- *Poverty & Public Policy*

Let's look at an example using the *Journal of Poverty and Social Justice*. Information about this journal is located on the journal's webpage: http://policy.bristoluniversitypress.co.uk/journals/journal-of-poverty-and-social-justice. In the section headed "About the Journal" you can see that it is an internationally focused research journal, and that it addresses social justice issues in addition to poverty alone. The research articles are peer-reviewed (there appear to be non-empirical discussions published, as well). These descriptions about a journal are almost always available, sometimes listed as "scope" or "mission." These descriptions also indicate the sponsorship of the journal—sponsorship may be institutional (a particular university or agency, such as *Smith College Studies in Social Work*), a professional organization, such as the Council on Social Work Education (CSWE) or the National Association of Social Work (NASW), or a publishing company (e.g., Taylor & Frances, Wiley, or Sage).

Indicators of journal caliber. Despite engaging in a peer review process, not all journals are equally rigorous. Some journals have very high rejection rates, meaning that many submitted manuscripts are rejected; others have fairly high acceptance rates, meaning that relatively few manuscripts are rejected. This is not necessarily the best indicator of quality, however, since newer journals may not be sufficiently familiar to authors with high quality manuscripts and some journals are very specific in terms of what they publish. Another index that is sometimes used is the journal's ***impact factor***. Impact factor is a quantitative number indicative of how often articles published in the journal are cited in the reference list of other journal articles—the statistic is calculated as the number of times on average each article published in a particular year were cited divided by the number of articles published (the number that could be cited). For example, the impact factor for the *Journal of Poverty and Social Justice*in our list above was 0.70 in 2017, and for the *Journal of Poverty*was 0.30. These are relatively low figures compared to a journal like the *New England Journal of Medicine*with an impact factor of 59.56! This means that articles published in that journal were, on average, cited more than 59 times in the next year or two.

Impact factors are not necessarily the best indicator of caliber, however, since many strong journals are geared toward practitioners rather than scholars, so they are less likely to be cited by other scholars but may have a large impact on a large readership. This may be the case for a journal like the one titled Social Work, the official journal of the National Association of Social Workers. It is distributed free to all members: over 120,000 practitioners, educators, and students of social work world-wide. The journal has a recent impact factor of.790. The journals with social work relevant content have impact factors in the range of 1.0 to 3.0 according to Scimago Journal & Country Rank (SJR), particularly when they are interdisciplinary journals (for example, *Child Development*, *Journal of Marriage and*

Family, *Child Abuse and Neglect*, *Child Maltreatment*,*Social Service Review*, and *British Journal of Social Work*). Once upon a time, a reader could locate different indexes comparing the "quality" of social work-related journals. However, the concept of "quality" is difficult to systematically define. These indexes have mostly been replaced by impact ratings, which are not necessarily the best, most robust indicators on which to rely in assessing journal quality. For example, new journals addressing cutting edge topics have not been around long enough to have been evaluated using this particular tool, and it takes a few years for articles to begin to be cited in other, later publications.

Beware of pseudo-, illegitimate, misleading, deceptive, and suspicious journals. Another side effect of living in the Age of Information is that almost anyone can circulate almost anything and call it whatever they wish. This goes for "journal" publications, as well. With the advent of open-access publishing in recent years (electronic resources available without subscription), we have seen an explosion of what are called ***predatory or junk journals***. These are publications calling themselves journals, often with titles very similar to legitimate publications and often with fake editorial boards. These "publications" lack the integrity of legitimate journals. This caution is reminiscent of the discussions earlier in the course about pseudoscience and "snake oil" sales. The predatory nature of many apparent information dissemination outlets has to do with how scientists and scholars may be fooled into submitting their work, often paying to have their work peer-reviewed and published. There exists a "thriving black-market economy of publishing scams," and at least two "journal blacklists" exist to help identify and avoid these scam journals (Anderson, 2017).

This issue is important to information consumers, because it creates a challenge in terms of identifying legitimate sources and publications. The challenge is particularly important to address when information from on-line, ***open-access journals*** is being considered. Open-access is not necessarily a poor choice—legitimate scientists may pay sizeable fees to legitimate publishers to make their work freely available and accessible as open-access resources. On-line access is also not necessarily a poor choice—legitimate publishers often make articles available on-line to provide timely access to the content, especially when publishing the article in hard copy will be delayed by months or even a year or more. On the other hand, stating that a journal engages in a peer-review process is no guarantee of quality—this claim may or may not be truthful. Pseudo- and junk journals may engage in some quality control practices, but may lack attention to important quality control processes, such as managing conflict of interest, reviewing content for objectivity or quality of the research conducted, or otherwise failing to adhere to industry standards (Laine & Winker, 2017).

One resource designed to assist with the process of deciphering legitimacy is the Directory of Open Access Journals (DOAJ). The DOAJ is not a comprehensive listing of all possible legitimate open-access journals, and does not guarantee quality, but it does help identify legitimate sources of information that are openly accessible and meet basic legitimacy criteria. It also is about open-access journals, not the many journals published in hard copy.

An additional caution: Search for article corrections. Despite all of the careful manuscript review and editing, sometimes an error appears in a published article. Most journals have a practice of publishing corrections in future issues. When you locate an article, it is helpful to also search for updates. Here is an example where data presented in an article's original tables were erroneous, and a correction appeared in a later issue.

- Marchant, A., Hawton, K., Stewart A., Montgomery, P., Singaravelu, V., Lloyd, K., Purdy, N., Daine, K., & John, A. (2017). A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. PLoS One, 12(8): e0181722. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5558917/
- Marchant, A., Hawton, K., Stewart A., Montgomery, P., Singaravelu, V., Lloyd, K., Purdy, N., Daine, K., & John, A. (2018).Correction—A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. PLoS One, 13(3): e0193937. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0193937

Search Tools. In this age of information, it is all too easy to find items—the problem lies in sifting, sorting, and managing the vast numbers of items that can be found. For example, a simple Google® search for the topic "community poverty and violence" resulted in about 15,600,000 results! As a means of simplifying the process of searching for journal articles on a specific topic, a variety of helpful tools have emerged. One type of search tool has previously applied a filtering process for you: ***abstracting and indexing databases***. These resources provide the user with the results of a search to which records have already passed through one or more filters. For example, PsycINFO is managed by the American Psychological Association and is devoted to peer-reviewed literature in behavioral science. It contains almost 4.5 million records and is growing every month. However, it may not be available to users who are not affiliated with a university library. Conducting a basic search for our topic of "community poverty and violence" in PsychINFO returned

1,119 articles. Still a large number, but far more manageable. Additional filters can be applied, such as limiting the range in publication dates, selecting only peer reviewed items, limiting the language of the published piece (English only, for example), and specified types of documents (either chapters, dissertations, or journal articles only, for example). Adding the filters for English, peer-reviewed journal articles published between 2010 and 2017 resulted in 346 documents being identified.

Just as was the case with journals, not all abstracting and indexing databases are equivalent. There may be overlap between them, but none is guaranteed to identify all relevant pieces of literature. Here are some examples to consider, depending on the nature of the questions asked of the literature:

- Academic Search Complete—multidisciplinary index of 9,300 peer-reviewed journals
- AgeLine—multidisciplinary index of aging-related content for over 600 journals
- Campbell Collaboration—systematic reviews in education, crime and justice, social welfare, international development
- Google Scholar—broad search tool for scholarly literature across many disciplines
- MEDLINE/ PubMed—National Library of medicine, access to over 15 million citations
- Oxford Bibliographies—annotated bibliographies, each is discipline specific (e.g., psychology, childhood studies, criminology, social work, sociology)
- PsycINFO/PsycLIT—international literature on material relevant to psychology and related disciplines
- SocINDEX—publications in sociology
- Social Sciences Abstracts—multiple disciplines
- Social Work Abstracts—many areas of social work are covered
- Web of Science—a “meta” search tool that searches other search

tools, multiple disciplines

Placing our search for information about “community violence and poverty” into the Social Work Abstracts tool with no additional filters resulted in a manageable 54-item list. Finally, abstracting and indexing databases are another way to determine journal legitimacy: if a journal is indexed in a one of these systems, it is likely a legitimate journal. However, the converse is not necessarily true: if a journal is not indexed does not mean it is an illegitimate or pseudo-journal.

Government Sources.A great deal of information is gathered, analyzed, and disseminated by various governmental branches at the international, national, state, regional, county, and city level. Searching websites that end in.gov is one way to identify this type of information, often presented in articles, news briefs, and statistical reports. These government sources gather information in two ways: they fund external investigations through grants and contracts and they conduct research internally, through their own investigators. Here are some examples to consider, depending on the nature of the topic for which information is sought:

- Agency for Healthcare Research and Quality (AHRQ) at https://www.ahrq.gov/
- Bureau of Justice Statistics (BJS) at https://www.bjs.gov/
- Census Bureau at https://www.census.gov
- Centers for Disease Control and Prevention (CDC) at https://www.cdc.gov
    - Morbidity and Mortality Weekly Report of the CDC (MMWR-CDC) at https://www.cdc.gov/mmwr/index.html
- Child Welfare Information Gateway at https://www.childwelfare.gov
- Children’s Bureau/Administration for Children & Families at https://www.acf.hhs.gov

- Forum on Child and Family Statistics at https://www.childstats.gov
- National Institutes of Health (NIH) at https://www.nih.gov, including (not limited to):
- National Institute on Aging (NIA at https://www.nia.nih.gov
- National Institute on Alcohol Abuse and Alcoholism (NIAAA) at https://www.niaaa.nih.gov
- National Institute of Child Health and Human Development (NICHD) at https://www.nichd.nih.gov
- National Institute on Drug Abuse (NIDA) at https://www.nida.nih.gov
- National Institute of Environmental Health Sciences at https://www.niehs.nih.gov
- National Institute of Mental Health (NIMH) at https://www.nimh.nih.gov
- National Institute on Minority Health and Health Disparities at https://www.nimhd.nih.gov
- National Institute of Justice (NIJ) at https://www.nij.gov
- Substance Abuse and Mental Health Services Administration (SAMHSA) at https://www.samhsa.gov/
- United States Agency for International Development at https://usaid.gov

Each state and many counties or cities have similar data sources and analysis reports available, such as Ohio Department of Health at https://www.odh.ohio.gov/healthstats/dataandstats.aspxand Franklin County at https://statisticalatlas.com/county/Ohio/Franklin-County/Overview. Data are available from international/global resources (e.g., United Nations and World Health Organization), as well.

Other Sources. The Health and Medicine Division (HMD) of the

National Academies—previously the Institute of Medicine (IOM)—is a nonprofit institution that aims to provide government and private sector policy and other decision makers with objective analysis and advice for making informed health decisions. For example, in 2018 they produced reports on topics in substance use and mental health concerning the intersection of opioid use disorder and infectious disease, the legal implications of emerging neurotechnologies, and a global agenda concerning the identification and prevention of violence (see http://www.nationalacademies.org/hmd/Global/Topics/Substance-Abuse-Mental-Health.aspx). The exciting aspect of this resource is that it addresses many topics that are current concerns because they are hoping to help inform emerging policy. The caution to consider with this resource is the evidence is often still emerging, as well.

Numerous "think tank" organizations exist, each with a specific mission. For example, the Rand Corporation is a nonprofit organization offering research and analysis to address global issues since 1948. The institution's mission is to help improve policy and decision making "to help individuals, families, and communities throughout the world be safer and more secure, healthier and more prosperous," addressing issues of energy, education, health care, justice, the environment, international affairs, and national security (https://www.rand.org/about/history.html). And, for example, the Robert Woods Johnson Foundation is a philanthropic organization supporting research and research dissemination concerning health issues facing the United States. The foundation works to build a culture of health across systems of care (not only medical care) and communities (https://www.rwjf.org).

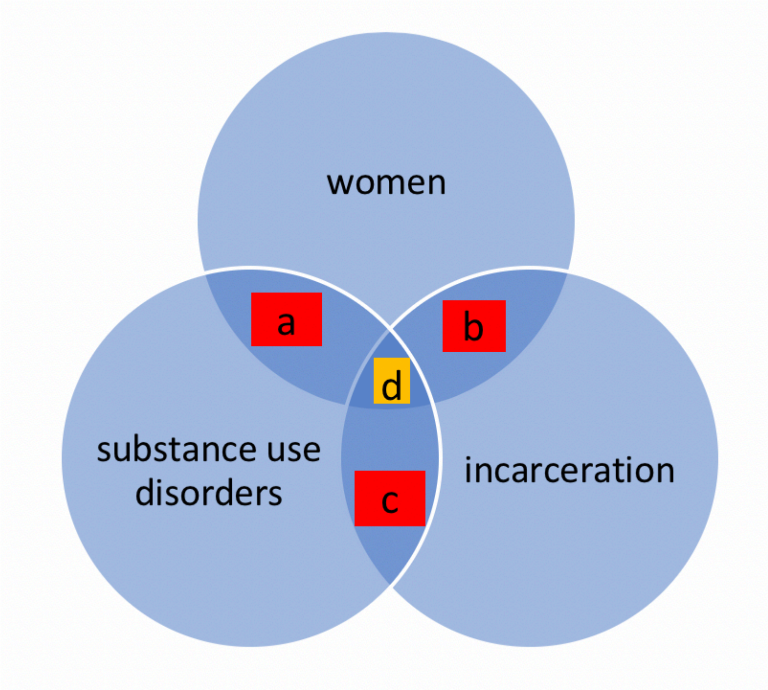
While many of these have a great deal of helpful evidence to share, they also may have a strong political bias. Objectivity is often lacking in what information these organizations provide: they provide evidence to support certain points of view. That is their purpose—to provide ideas on specific problems, many of which have a political component. Think tanks "are constantly researching solutions to a variety of the world's problems, and arguing, advocating, and lobbying for policy changes at local, state, and federal levels" (quoted from https://thebestschools.org/features/most-influential-think-tanks/). Helpful information about what this one source identified as the 50 most influential U.S. think tanks includes identifying each think tank's political orientation. For example, The Heritage Foundation is identified as conservative, whereas Human Rights Watch is identified as liberal.

While not the same as think tanks, many mission-driven organizations also sponsor or report on research, as well. For example, the National Association for Children of Alcoholics (NACOA) in the United States is a registered nonprofit organization. Its mission, along with other partnering organizations, private-sector groups, and federal agencies, is to promote policy and program development in research, prevention and treatment to provide information to, for, and about children of alcoholics (of all ages). Based on this mission, the organization supports knowledge development and information gathering on the topic and disseminates information that serves the needs of this population. While this is a worthwhile mission, there is no guarantee that the information meets the criteria for evidence with which we have been working. Evidence reported by think tank and mission-driven sources must be utilized with a great deal of caution and critical analysis!

In many instances an empirical report has not appeared in the published literature, but in the form of a technical or final report to the agency or program providing the funding for the research that was conducted. One such example is presented by a team of investigators funded by the National Institute of Justice to evaluate a program for training professionals to collect strong forensic evidence in instances of sexual assault (Patterson, Resko, Pierce-Weeks, & Campbell, 2014): https://www.ncjrs.gov/pdffiles1/nij/grants/247081.pdf. Investigators may serve in the capacity of consultant to agencies, programs, or institutions, and provide empirical evidence to inform activities and planning. One such example is presented by Maguire-Jack (2014) as a report to a state's child maltreatment prevention board: https://preventionboard.wi.gov/Documents/InvestmentInPreventionPrograming_Final.pdf.

When Direct Answers to Questions Cannot Be Found. Sometimes social workers are interested in finding answers to complex questions or questions related to an emerging, not-yet-understood topic. This does not mean giving up on empirical literature. Instead, it requires a bit of creativity in approaching the literature. A Venn diagram might help explain this process. Consider a scenario where a social worker wishes to locate literature to answer a question concerning issues of intersectionality. Intersectionality is a social justice term applied to situations where multiple categorizations or classifications come together to create overlapping, interconnected, or multiplied disadvantage. For example, women with a substance use disorder and who have been incarcerated face a triple threat in terms of successful treatment for a substance use disorder: intersectionality exists between being a woman, having a substance use disorder, and having been in jail or prison. After searching the literature, little or no empirical evidence might have been located on this specific triple-threat topic. Instead, the social worker will need to seek literature on each of the threats individually, and possibly will find literature on pairs of topics (see Figure 3-1). There exists some literature about women's outcomes for treatment of a substance use disorder (a), some literature about women during and following incarceration (b), and some literature about substance use disorders and incarceration (c). Despite not having a direct line on the center of the intersecting spheres of literature (d), the social worker can develop at least a partial picture based on the overlapping literatures.

Figure 3-1. Venn diagram of intersecting literature sets.



Take a moment to complete the following activity. For each statement about empirical literature, decide if it is true or false.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=144

Module 2 Chapter 4: Reviewing Empirical Articles

After carefully reviewing the source you have located, it is time to critically review the piece itself. In this chapter, you will read about:

- Steps in reviewing different sections of an empirical article
- The importance of maintaining a critical perspective on what you are reading–being an "active reader"

The seven steps considered in this chapter relate to the structure of typical journal articles published in social work and allied discipline journals. The structure is familiar to anyone who works with the American Psychological Association (APA) guide to how journal articles are written and structured, the*Publication Manual of the American Psychological Association*(APA, 2009 for the sixth edition).

Step 1. What is in a Name? Reviewing the title

Authors vary tremendously in their approach to titling their work, much as parents differ markedly in their approach to naming their babies. Ideally, a title is sufficiently precise and specific to tell a reader what the article is about. Sometimes titles have attention catching phrases added at the front or back end. Ideally, a title is also not overly elaborate and lengthy.

For example, the following article titles clearly and succinctly communicates what each article is about.

- "Race and ethnic differences in early childhood maltreatment in

the United States" (Lanier, Maguire-Jack, Walsh, & Hubel, 2014).
- "Tracking the when, where, and with whom of alcohol use: Integrating ecological momentary assessment and geospatial data to examine risk for alcohol-related problems" (Freisthler, Lipperman-Kreda, Bersamin, & Gruenewald, 2014).
- "Meeting them where they are: An exploration of technology use and help seeking behaviors among adolescents and young adults" (Cash & Bridge, 2012).
- "A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown" (Marchant et al, 2017). [Note the word "behavior" is spelled "behaviour" in the United Kingdom, but without the "u" in the United States.]

Examples of (hypothetical) article titles that are non-communicative or miscommunicate include:

- "The problem of drugs in America." This title is not sufficiently specific for an empirical article. First, a reader does not know if this is about how the pharmacology industry manufactures drugs (quality control), an epidemiology report about the scope of the problem, an etiology report about theories related to the causes of the problem, a test of a sociological, psychological, or biological theory. It is more likely a title about an opinion piece, a general book chapter, or even an entire book.
- "The effectiveness of PFI and MI in AOD treatment." This title is swimming in jargon, making it difficult to interpret without looking into what each acronym means, and also makes it difficult to locate in the literature. (PFI is Personal Feedback Intervention, MI is motivational interviewing, and AOD is about alcohol and other drugs.) In addition, there exists a problem of ambiguity with acronyms: BPD could refer to either bronchopulmonary dysplasia, a lung complication common

among prematurely born infants, borderline personality disorder, or bipolar disorder (both are mental disorders).
- "The effects of self-esteem on high school student retention and drop-out." This title is only good if the study design and methods actually allow for a causal inference. If the study design only allows for conclusions about the existence of a relationship between these two variables, the title is a poor choice—the word "effects" implies causality and misrepresents the study.

Step 2. What is it about? Reviewing the abstract

Authors present a summary of their manuscript in a brief abstract that appears at the start of a published article. Abstracts are also published in a number of indexing and abstracting resources, making them relatively easy to access. Journals limit the length of abstracts, usually to somewhere between 150-250 words depending on the journal. This makes it challenging to explain the important aspects of the manuscript, with enough detail to be clear, but without the luxury of unlimited space to present nuances. An abstract should address the following points:

- Study's purpose, research aims, questions, and/or hypotheses
- Study approach
- Study design and methods (including study participants and measures)
- Data analysis and key results
- Key implications of the study results.

As a reader, the abstract should provide enough information for you to determine whether it is relevant for you to pursue the full article. An ***article abstract***is not sufficient information for you to evaluate the evidence, even the evidence that appears in the description of results! To evaluate the evidence, you need to acquire and review the full article.

This point is so important that it warrants repeating:

> *An article's abstract is not sufficient information for you to evaluate the evidence, even the evidence that appears in the description of results! To evaluate the evidence, you need to acquire and review the full article.*

A Note about research abstracts. An excellent search of literature will often turn up published ***research abstracts***. In this case, there will not be a full article to locate. The abstract describes a conference presentation, and these are the precursors to publishing an article about a research study. Many professional organizations publish these abstracts in a special journal issue, sometimes a "supplement" to the journal. The next step for your search, if the title and abstract seem relevant, will be to determine whether a paper was ever published based on the study described in the abstract. This is a place where you would search by author name(s) rather than by subject or topic alone.

Step 3. What is the rationale and background knowledge? Reviewing the introduction

Once you have acquired an article that seems interesting and relevant (based on its title and abstract), you will next encounter its introduction. The purpose of an introduction is to provide the reader with a background orientation to the study that was conducted. This might include background information regarding the scope of the problem being addressed by the research. It should certainly provide the reader with a review of literature related to the topic and research questions. This might include an overview of the theory or theories related to the research that was conducted. In the end, the reader should understand the following:

1. Why was the study was undertaken, why was it important, why does it matter?
2. What was known from the literature that informed the study's development?
3. What knowledge gap or gaps did the study aim to fill, or what did the study aim to contribute to the body of knowledge?
4. What research questions did the investigators address in their study?

The introduction often also informs readers about the study's approach (e.g., qualitative, quantitative, mixed methods) and type of study that was implemented (e.g., exploratory, descriptive, experimental). After reviewing the introduction, you should have an even better idea of whether the article is relevant for your purposes.

Step 4. What happened? Reviewing the methods

If it was not made clear in the introduction, the study approach and type of study should be made explicit in the methods section of the article. The methods section, at a minimum, needs to explain who participated in the study and how data were collected. In a quantitative study, the study design is also described in the methods section; in a qualitative study, the type of study is described. There are three basic sub-sections in the methods section of an empirical article: study participants, study measures, and study procedures.

Study participants. A methods sub-section describes who actually participated in the study, including numbers and characteristics of the study participants, as well as the pool from which these participants were drawn. The purpose of this sub-section in describing a quantitative study is to inform readers about generalizability of the study's results and inform other investigators about what they would need to do to replicate the study to determine if they achieve similar results. Authors may present some of the description material in the form of tables with information about numbers and proportions reflecting categorical variables (like gender or race/ethnicity) and distribution on scale/continuous variables (like age). The method of selecting these participants should be clear along with any inclusion or exclusion criteria that were applied. The participant response rate might also be calculated as the number of participants enrolled in the study divided by the number of persons eligible to be enrolled (the "pool"), multiplied by 100%. Very low response rates make a study vulnerable to selection bias—the few persons who elected to participate might not represent the general population. In a qualitative study, the study participants section again describes how the individuals were selected for participation, and details describing these individuals are provided. Generalizability is not a goal in qualitative studies, but information about study participants should provide an indication to a reader of how robust the results might be. Robust descriptions come from participants who exhibit a range of defining and/or experiential characteristics. Finally, regardless of study approach, authors typically make evident that the study was reviewed by an Institutional Review Board for the inclusion of human participants.

Study Measures. Another methods sub-section explains how the data were collected. For quantitative studies, the data collection instruments used to measure each study variable are described. If the tools used for data collection were previously published, authors cite the sources of those tools and published literature about them—their reliability and validity, for example. The authors may also summarize literature concerning how the measures are known to perform with the specific type of study participants involved in the study—for example, different ages, diagnoses, races/ethnicities, or other characteristics. For qualitative studies the interview protocols or questions asked of participants are described in detail. In observational data collection studies, the approach to recording and scoring/coding observed behavior are described. In any case, the approach to data collection or measurement is described in sufficient detail for a reader to critically appraise the adequacy of the data collection approach and conclusions that can be drawn from the data collection process, and for other investigators to be able to replicate the study should they wish to confirm the results.

Study procedures. Sometimes study procedures is a separate methods sub-section and sometimes this content is incorporated into the participants and measures sub-sections. This sub-section includes information about activities in which the study participants engaged during the study. In a quantitative, experimental study, the methods utilized to assign study participants to different experimental conditions might be described here (i.e., the randomization approach used). Additionally, procedures used in handling data are usually described. In a quantitative study, investigators may report how they scored certain measures and what evidence from the literature informs their scoring approach. In a qualitative study, details about how data were coded are reported here. Procedures for ensuring inter-observer or inter-rater reliability and agreements will also be reported for either type of study. Regardless of the study's research approach, a reader should come away with a detailed understanding of how the study was executed. As a result, the reader should be sufficiently informed about the study's execution to be able to critically analyze the strength of the evidence developed from the methods that were applied.

Step 5. What was found? Reviewing the results

The results section is where investigators describe the data they collected, how it was analyzed, and what was observed in the data.The structure and format of the results section varies markedly for different research approaches.

Qualitative methods results description. The nature of qualitative research questions and methods leads to data that are richly descriptive. The results derived from the data are, therefore, generally descriptive in nature. There may be a great deal of direct quotes, presenting information in participants' own words. Descriptions may include thematic or concept maps constructed by the study investigators as a means of "sense making" from the data. If statistics are included, they tend to be of a descriptive nature—perhaps demonstrating the frequency with which certain results were observed in the data. The results may include tables or figures representing results. Ideally, a reader can identify the way that reported results relate to the research aims or questions originally asked by the study.

Quantitative methods results description. The nature of quantitative research questions and methods leads to numeric data that can be summarized using various forms of statistical analyses. The results section of a quantitative study report will describe which statistical analyses were utilized, and should indicate the rationale for selecting those analyses, as well as discussing how well the data were suited to those analyses. Analyses that involve hypothesis testing will indicate the statistical support for conclusions drawn from the data (i.e., descriptive statistics, test statistics, significance levels, and confidence intervals). The results may be presented in a combination of text descriptions, tables, and figures. An informed reader should be able to determine the appropriateness of the statistical approaches used in the analyses and the conclusions drawn from those analyses. Ideally, study results are presented in association with each study question or hypothesis as it is answered. Problems encountered with any specific analyses are also reported here, such as when data were not suitably distributed, sample sizes were inadequate, or assumptions underlying specific types of analytic approaches were violated.

Mixed methods results description. As mixed methods approaches continue evolving, so are creative ways of presenting results of studies that integrate qualitative and quantitative approaches. Results for a study employing mixed methods are often presented question-by-question. Where qualitative questions were addressed, the descriptive results will be presented as outlined above. Where quantitative questions were addressed, the numeric and statistical results will be presented as outlined above.

Regardless of study approach, a reader should have a clear understanding of the way data were analyzed and results of those analyses. This is not a place where authors have drawn conclusions about the implications of those results—that belongs in the article's discussion section.

Step 6. What was concluded? Reviewing the Discussion

In the end, the authors will offer their interpretation of the evidence described in their Results section. This discussion should include several elements:

- A brief overview summary of the key results.
- Discussion of how each key result relates to the study aims,

questions, and/or hypotheses.

- Discussion of how the observed results relate to the previous existing literature (are they mutually confirming or contradicting), if the study results were completely new contributions, or if they were ambiguous and no conclusions could be drawn.
- Discussion of the study's methodological or analysis/results limitations.
- Practical implications of the study results for practice and for future research.

It is important to remember that the discussion is the authors' own interpretation of the results. This, again, is a place where readers must apply their own critical analysis to the study implications. For example, sometimes authors get a bit carried away with their interpretation and make suggestions that are not supported by the evidence in their studies. Or, they may not have gone far enough, and you see potential implications that they did not.

Step 7. Where are other relevant pieces? Reviewing the reference list

As you search for relevant literature, you might want to review the reference list to an article that you found to be relevant. Sometimes your own search methods and search terms might have missed some important items that the article's authors were able to identify. This review will provide you with titles to consider, and possibly you will recognize the names of key scholars in the topic area. You can then pursue these background resources as part of your own search.

Interactive Excel Workbook Activities

Complete the following Workbook Activities:

- Workbook Introduction and Steps to Install Microsoft Office 365
- SWK 3401.2-4.1 Getting Started: Excel and Data Analysis ToolPak Access

Module 2 Module Summary

The chapters in this module covered a wide range of topics. You read about the nature of social work research questions and how these relate to different types of research. The different types included those involved in a translational science process, as well as those addressing epidemiology, etiology, or mechanisms of change questions. The framework of exploratory, descriptive, and explanatory research was introduced, and this framework will be revisited in future modules as we examine different research methodologies. The role of theory in relation to research and the role of theory and research in promoting social justice were also examined in this module. The importance of recognizing how social work interfaces with other disciplines was emphasized, as well. Finally, you were presented with a great deal of information about empirical literature—what it is, where to find it, and how to review it. All of this information is designed to assist you as a social work professional as you strive to develop your understanding of diverse populations, social work problems, and social phenomena, and engage in efforts to develop culturally responsive intervention strategies. This information serves as a foundation, relevant in our next course about understanding interventions.

Module 2 Key Terms and Definitions

abstracting and indexing databases: provide article abstracts and information about accessing full articles of interest; may not be in the public domain; access depends on the library use.

article abstract: a summary of an article; helps to determine if the full article will be of interest/relevant.

background question: answerable with facts, direct a straightforward search for evidence; typically get at the epidemiology or etiology of a social work problem or social phenomenon (as compared to foreground question).

basic research: fundamental or foundational research, usually to build or test theory or improve prediction; tends to impose rigid control over variables (compared to applied research).

biopsychosocial: a perspective that integrates biological, psychological, and social system factors and forces in explaining a phenomenon; also, may consider the interactions among these factors and forces.

correlational research question: seeking descriptions about relationships that might exist between variables; questions of association.

descriptive research question: directs research around developing descriptions about populations, problems, processes, or phenomena.

empirical literature: literature providing information based on systematic observation and applying systematic methodology, objectively developed, characterized by transparency, and reproducible (meeting definition of research).

epidemiology: science of understanding the distribution of problems or phenomena in the population; also, may include understanding relationships between problems/phenomena, and the root causes; common approach used in public health.

etiology: science testing theories and hypotheses about the origins and natural course of a problem or phenomenon

explanatory research question: directs research around testing theories and hypotheses about causality and causal relationships between variables; usually leads to experimental study approaches.

exploratory research question: directs research around developing a basic, beginning understanding of a newly recognized phenomenon of interest where not much is yet known or understood, to the point where knowing what questions about theory and causes are premature.

foreground question: answers allow comparison of options, usually practice decisions; often complex (as compared to background questions).

impact factor: a statistic used to evaluate how much articles published in a particular journal have been cited by others, as a gauge of the journal's impact on the field.

implementation science: the study of how interventions are disseminated and adopted into practice.

interdisciplinary: where different disciplines work together, creating an integrated, unified whole.

intervention research: science of the impact of specific interventions on outcomes of interest; in addition to what outcomes are observed, may also include studies of mechanisms by which interventions have their impact on outcomes (mechanisms of change).

mechanisms of change: the specific factors, variables, and processes by which changes in outcome are produced; how an intervention has its effects.

multidisciplinary: where different disciplines contribute separate parts, individually, that maintain their distinct, unique characteristics in the whole.

open-access journals: electronic publications available to anyone without a subscription, designed to make information as widely accessible to public, global audiences as possible; note cautions regarding discerning between legitimate and predatory/junk open-access resources.

peer review process: publication (and grant award) process by which submissions are evaluated for adequacy and accuracy by other scholars familiar with the topic/field/area.

predatory or junk journals: publications lacking integrity of legitimate journals, proliferating in the open environment of the internet/web access.

research abstracts: published summaries of a research report; may or may not be accompanied by a full manuscript.

research hypothesis: a clear, testable statement predicting an outcome or results, and that will either be rejected or not, depending on the evidence; the prediction or hypothesis is based on theory and/or other evidence.

social indicators: statistics that describe the social, economic, and psychological factors that have an impact on the well-being of a community or other population group.

theory: explanatory belief(s), idea(s), or principles; may be simple or from complex system of principles.

think tank: formal organizations that provide ideas on specific problems, many of which have a political component and advocate/lobby for policy change based on their "research."

transdisciplinary: where disciplinary differences become insignificant; the "whole" is equally relevant across disciplines.

translational science: a system whereby basic research informs intervention research, and implementation science addresses adoption of evidence-supported interventions.

Module 2 References

American Psychological Association (APA). (2009). *Publication Manual of the American Psychological Association*, sixth edition. Washington, DC: American Psychological Association.

Anderson, R. (2017). *Cabell's new predatory journal blacklist: A review*. Blog posting of Society for Scholarly Publishing, The Scholarly Kitchen (July 25).

Anderson-Butcher, D., Amorose, A.J., Lowers, L., Gibson, A., Riley, A., & Ruch, D. (2016). The case for the Perceived Social Competence Scale-II. *Research on Social Work Practice*. 26(4), 419-428.

Ardito, R.B., & Rabellino, D. (2011). Therapeutic alliance and outcome of psychotherapy: Historical excursus, measurements and prospects for research. *Frontiers in Psychology*, 2, 270-281.

Begun, A.L., Brondino, M.J., Bolt, D., Weinstein, B., Strodthoff, T., & Shelley, G. (2008). The Revised Safe At Home Instrument for assessing readiness to change intimate partner violence. *Violence and Victims*, 23(4), 508-524.

Begun, A.L., Murphy, C., Bolt, D., Weinstein, B., Strodthoff, T., Short, L., & Shelley, G. (2003). Characteristics of the Safe At Home Instrument for assessing readiness to change intimate partner violence. *Research on Social Work Practice*, 13(1), 80-107.

Birken, S.A., Bunger, A.C., Powell, B.J., Turner, K., Clary, A.S., Klaman, S.L., Yu, Y., Whitaker, D.J., Self, Sh.R., Rostad, W.L., Chatham, J.R.S., Kirk, M.A., Shea, C.M., Haines, E., & Weiner, B.J. (2017). Organizational theory for dissemination and implementation research. *Implementation Science*, 12(62).

Centers for Disease Control and Prevention (CDC). *Multiple causes of death* 1999-2016, CDC WONDER database, National Center for Health Statistics (https://wonder.cdc.gov).

Clapp, J.D., Holmes, M.R., Reed, M.B., Shillington, A.M., Freisthler, B. & Lange, J.E. (2007). Measuring college students' alcohol consumption in natural drinking environments: Field methodologies for bars and parties. *Evaluation Review*, 31, 469-489.

Cowell, A.J., Brown, J.M., Mills, M.J., Bender, R.H., & Wedehase, B.J. (2012). Cost-effectiveness analysis of motivational interviewing with feedback to reduce drinking among a sample of college students. *Journal of Studies in Alcohol and Drugs*, 73(2), 226-237.

Davis, T. S., Saltzburg, S., & Locke, C. R. (2010). Assessing community needs of sexual minority youth: Modeling Concept Mapping for service planning. *Journal of Gay & Lesbian Social Services*, 22(3), 226- 249.

Dudley, J.R. (2011). *Research methods in social work: Being producers and consumers of research, 2ndedition/updated edition.* Boston: Allyn & Bacon.

Fields, N.L., Anderson, K.A. & Dabelko-Schoeny, H. (2014). The effectiveness of adult day services for older adults: A review of the literature from 2000 to 2011. *Journal of Applied Gerontology*, 33(2), 130-163.

Freisthler, B., Lipperman-Kreda, S., Bersamin, M., & Gruenewald, P.J. (2014). Tracking the when, where, and with whom of alcohol use: Integrating ecological momentary assessment and geospatial data to examine risk for alcohol-related problems. *Alcohol Research: Current Reviews*, 36, 29-38.

Grinnell, R.M., & Unrau, Y.A. (2014). *Social work research and* *evaluation: Foundations of evidence-based practice, 10thedition*. NY: Oxford University Press.

Hedegaard, H., Warner, M., & Miniño, A.M. (2017). *Drug overdose deaths in the United States, 1999-2016*. NCHS Data Brief, no 294. Hyattsville, MD: National Center for Health Statistics.

Ioannou, M., Hammond, L., & Simpson, O. (2015). A model for differentiating school shooters' characteristics. *Journal of Criminal Psychology*, 5(3), 188-200.

Jarrott, S.E., Weaver, R.H., Bowen, N.K., & Wang, N. (2018). Measuring dimensions of intergenerational contact: Factor analysis of the Queen's University Scale. *Aging & Mental Health*, 22(4), 568-573.

Jensen, C.D., Cushing, C.C., Aylward, B.S., Craig, J.T., Sorell, D.M., & Steele, R.G. (2011). Effectiveness of motivational interviewing interventions for adolescent substance use behavior change: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 79(4), 433-440.

Kazdin, A.E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, 3, 1-27.

Kyriakakis, S., Waller, B., Kagotho, N., & Edmond, T. (2015). Conducting safe research with at risk populations: Design strategies from a study with unauthorized immigrant women experiencing intimate abuse. *Qualitative Social Work*, 14(2) 259–274.

Laine, C., & Winker, M.A. (2017). Identifying predatory or pseudo-journals. *Biochemia Medica*, 27(2), 285-291.

Lanier, P., Maguire-Jack, K., Walsh, T., & Hubel, G. (2014). "Race and ethnic differences in early childhood maltreatment in the United

States." *Journal of Behavioral and Developmental Pediatrics*, 35(7), 419-426.

Langman, P. (2014). Statistics on bullying and school shootings. Retrieved from www.schoolshooters.info

Lower, L.M., Anderson-Butcher, D., & Newman, T.J. (2017). Validity and reliability of the Teamwork Scale for Youth, *Research on Social Work Practice*, 27(6), 716-725.

Maguire-Jack, K. (2104). Child maltreatment prevention programs—A worthwhile investment. What it Will Take: Investing in Wisconsin's future by keeping kids safe today, Madison, WI: Wisconsin Children's Trust Fund. Retrieved from https://preventionboard.wi.gov/Documents/InvestmentInPreventionPrograming_Final.pdf

Marchant, A., Hawton, K., Stewart A., Montgomery, P., Singaravelu, V., Lloyd, K., Purdy, N., Daine, K., & John, A. (2017). A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. PLoS One, 12(8): e0181722.

Marchant, A., Hawton, K., Stewart A., Montgomery, P., Singaravelu, V., Lloyd, K., Purdy, N., Daine, K., & John, A. (2018). Correction—A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. PLoS One, 13(3): e0193937.

Mears, D.P., Moon, M.M., & Thielo, A.J. (2017). Columbine revisited: Myths and realities about the bullying-school shootings connection. *Victims & Offenders*, 12(6), 939-955.

Mayo, N.E., Asano, M., & Barbic, S.P. (2013). When is a research question not a research question? *Journal of Rehabilitation Medicine*, 45(6), 513-518.

National Academies of Sciences, Engineering, and Medicine (NAS). (2018). *Integrating responses at the intersection of opioid use disorder and infectious disease epidemics: Proceedings of a workshop.*Washington, DC: The National Academies Press.

Newman, K.S., & Fox, C. (2009). Repeat tragedy: Rampage shootings in American high school and college settings, 2002-2008. *American Behavioral Scientist*, 52(9), 1286-1308.

Office of Behavioral and Social Sciences Research (OBSSR). (2003). *NIH Plan for social work research*. Retrieved from

https://obssr.od.nih.gov/wp-content/uploads/2016/07/SWR_Report.pdf.

Patterson, D., Resko, S., Pierce-Weeks, J., & Campbell, R. (2014). Delivery and evaluation of Sexual Assault Forensic (SAFE) training programs. Final Report. Submitted to National Institute for Justice. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/247081.pdf.

Raitanen, J., Sandberg, S., & Oksanen, A. (2017). The bullying-school shooting nexus: Bridging master narratives of mass violence with personal narratives of social exclusion. *Deviant Behavior, online*,1-14.

Shao, W., Guan, W., Clark, M.A., Liu, T., Santelices, C.C., Cortes, D.E., & Merchant, R.C. (2015). Variations in recruitment yield, costs, speed and participant diversity across internet platforms in a global study examining the efficacy of an HIV/AIDS and HIV testing animated and live-action video among English- or Spanish-speaking Internet or social media users. *Digital Culture and Education*, 7(1), 40-86.

Sielski, C.L., Begun, A.L., & Hamel, J. (2015). Expanding knowledge concerning the Safe at Home Instrument for assessing readiness to change among individuals in batterer treatment. *Partner Abuse*, 6(3), 255-272.

Smith, S.G., Chen, J., Basile, K.C., Gilbert, L.K., Merrick, M.T., Patel, N., Walling, M., & Jain, A. (2017). *The National Intimate Partner and Sexual Violence Survey* (NISVS): 2010-2012 *State Report.*Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Stoeber, M., Jullié, D., Lobingier, B.T., Laeremans, T., Steyaert, J., Schiller, P.W., Manglik, A., & von Zastrow, M. (2018). A genetically encoded biosensor reveals location bias of opioid drug action.*Neuron*, 98, 1-14.

Subbaraman, M.S., Ludet, A.B., Ritter, L.A., Stunz, A., & Kaskutas, L.A. (2015). Multisource recruitment strategies for advancing addiction recovery research beyond treated samples. *Journal of Community Psychology*, 43,560-575.

Testa, M., & Cleveland, M. (2017). Does alcohol contribute to college men's sexual assault perpetration? Between- and within-person effects over five semesters. *Journal of Studies on Alcohol and Drugs*, 78, 5-13.

Vestal, C. (2018). Overdose deaths fall in 14 states. PEW Trusts, retrieved from http://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2018/02/22/overdose-deaths-fall-in-14-states.

Wampold, B.E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14(3), 270-277.

Williams, J.H. (2016). [Editorial] Grand challenges for social work: Research, practice, and education. *Social Work Research*, 40(2), 67-70.

MODULE 3

Module 3 Introduction

In this module, we build on what we learned in earlier modules about the nature of our research questions concerning diverse populations and the problems that social workers address. This module affords us the opportunity to become familiar with several different approaches for answering questions about populations and social work problems. We explore the philosophies and assumptions underlying quantitative, qualitative, and mixed methods approaches, and we analyze various advantages/disadvantages associated with their use. We describe and analyze different types of study designs used to answer population and social work problem questions (survey, group comparison/experimental, grounded theory, narrative/ethnography, and others). We explore:

- Qualitative, quantitative, and mixed methods approaches to understanding diverse populations,social work problems, and social phenomena;
- The types of approaches used for drawing samples for different study design;
- Measurement strategies for different types of variables (independent, dependent, categorical, continuous);
- Varied approaches to data collection (e.g., observation, interview, self-administered survey, online surveys, focus groups, and others), and cultural competence in measurement

READING OBJECTIVES

After engaging with these reading materials and learning resources, you should be able to:

- Identify the similarities and differences underlying qualitative, quantitative, and mixed methods study approaches, and how they relate to different types of research questions;

- Recognize different qualitative research traditions;
- Describe longitudinal and cross-sectional approaches to designing research studies and how they relate to different types of research questions;
- Recognize different types of quantitative variables and how they relate to our research questions;
- Identify advantages and disadvantages of various measurement and data collection strategies;
- Critique measurement and data collection approaches on the dimension of cultural competence;
- Explain basic principles associated with designing a plan for who might be engaged as participants in research studies about diverse populations and social work problems;
- Define key terms related to social work research approaches and methods.

Module 3 Chapter 1: From Research Questions to Research Approaches

The approaches that social work investigators adopt in their research studies are directly related to the nature of the research questions being addressed.In Module 2 you learned about exploratory, descriptive, and explanatory research questions. Let's consider different approaches to finding answers to each type of question.

In this chapter we build on what was learned in Module 2 about research questions, examining how investigators' approaches to research are determined by the nature of those questions. The approaches we explore are all systematic, scientific approaches, and when properly conducted and reported, they all contribute empirical evidence to build knowledge. In this chapter you will read about:

- qualitative research approaches for understanding diverse populations, social problems, and social phenomena,
- quantitative research approaches for understanding diverse populations, social problems, and social phenomena,
- mixed methods research approaches for understanding diverse populations, social problems, and social phenomena.

Overview of Qualitative Approaches

Questions of a descriptive or exploratory nature are often asked and addressed through ***qualitative research***. The specific aim in these studies is to understand diverse populations, social work problems, or social phenomena as they naturally occur, situated in their natural environments, providing rich, in-depth, participant-centered descriptions of the phenomena being studied. Qualitative research approaches have been described as "humanistic" in aiming to study the world from the perspective of those who are experiencing it themselves; this also contributes to a social justice commitment in that the approaches give "voice" to the individuals who are experiencing the phenomena of interest (Denzen & Lincoln, 2011). As such, qualitative research approaches are also credited with being sensitive and responsive to diversity—embracing feminist, ethnic, class, critical race, queer, and ability/disability theory and lenses.

In qualitative research, the investigator is engaged as an observer and interpreter, being acutely aware of the subjectivity of the resulting observations and interpretations.

> "At *this level, qualitative research involves an interpretive, naturalistic approach to the world*" (*Denzin & Lincoln*, 2011, *p.* 3)

Because the data are rich and deep, a lot of information is collected by involving relatively few participants; otherwise, the investigator would be overwhelmed by a tremendous volume of information to collect, sift through, process, interpret, and analyze. Thus, a single qualitative study has a relatively low level of ***generalizability*** to the population as a whole because of its methodology, but that is not the aim or goal of this approach.

In addition, because the aim is to develop understanding of the participating individuals' lived experiences, the investigator in a qualitative study seldom imposes structure with standardized measurement tools. The investigator may not even start with preconceived theory and hypotheses. Instead, the methodologies involve a great deal of open-ended triggers, questions, or stimuli to be interpreted by the persons providing insight:

> “*Qualitative research’s express purpose is to produce descriptive data in an individual’s own written or spoken words and/or observable behavior*” (Holosko, 2006, p. 12).

Furthermore, investigators often become a part of the qualitative research process: they maintain awareness of their own influences on the data being collected and on the impact of their own experiences and processes in interpreting the data provided by participants. In some qualitative methodologies, the investigator actually enters into/becomes immersed in the events or phenomena being studied, to both live and observe the experiences first-hand.

Qualitative data and interpretations are recognized as being subjective in nature—that is the purpose—rather than assuming objectivity. Qualitative research is based on experientially derived data and is interpretive, meaning it is “concerned with understanding the meaning of human experience from the subject’s own frame of reference” (Holosko, 2006, p. 13). In this approach, conclusions about the nature of reality are specific to each individual study participant, following his or her own interpretation of that reality. These approaches are considered to flow from an inductive reasoning process where specific themes or patterns are derived from general data (Creswell & Poth, 2018).

Several purposes of qualitative approaches in social work include:

- describing and exploring the nature of phenomena, events, or relationships at any system level (individual to global)
- generating theory
- initially test ideas or assumptions (in theory or about practices)
- evaluate participants’ lived experiences with practices, programs, policies, or participation in a research study, particularly with diverse participants
- explore “fit” of quantitative research conclusions with

participants' lived experiences, particularly with diverse participants
- inform the development of clinical or research assessment/measurement tools, particularly with diverse participants.

Overview of Quantitative Approaches

Questions of the exploratory, descriptive, or explanatory type are often asked and addressed through ***quantitative research*** approaches, particularly questions that have a numeric component. Exploratory and descriptive quantitative studies rely on objective measures for data collection which is a major difference from qualitative studies which are aimed at understanding subjective perspectives and experiences. Explanatory quantitative studies often begin with theory and hypotheses, and proceed to empirically test the hypotheses that investigators generated. By their quantitative (numeric) nature, statistical hypothesis testing is possible in many types of quantitative studies.

Quantitative research studies utilize methodologies that enhance generalizability of results to the greatest extent possible—individual differences are de-emphasized, similarities across individuals are emphasized. These studies can be quite large in terms of participant numbers, and the study samples need to be developed in such a manner as to support generalization to the larger populations of interest.

The process is generally described as following a deductive logical system where specific data points are combined to lead to developing a generalizable conclusion. The philosophical roots (epistemology) underlying quantitative approaches is positivism, involving the seeking of empirical "facts or causes of social phenomena based on experimentallyderived evidence and/or valid observations" (Holosko, 2006, p. 13). The empirical orientation is objective in that investigators attempt to be detached from the collection and interpretation of data in order to minimize their own influences and biases. Furthermore, investigators utilize objective measurement tools to the greatest extent possible in the process of collecting quantitative study data.

Several purposes of quantitative approaches in social work include:

- describing and exploring the dimensions of diverse populations, phenomena, events, or relationships at any system level (individual to global)—how much, how many, how large, how often, etc. (including epidemiology questions and methods)
- testing theory (including etiology questions)
- experimentally determining the existence of relationships between factors that might influence phenomena or relationships at any system level (including epidemiology and etiology questions)
- testing causal pathways between factors that might influence phenomena or relationships at any system level (including etiology questions)
- evaluate quantifiable outcomes of practices, programs, or policies
- assess the reliability and validity of clinical or research assessment/measurement tools.

Overview of Mixed-Method Approaches

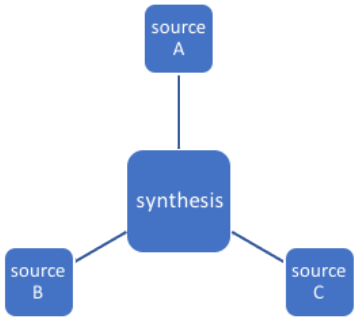
Important dimensions distinguish between qualitative and quantitative approaches. First, qualitative approaches rely on "insider" perspectives, whereas quantitative approaches are directed by "outsiders" in the role of investigator (Padgett, 2008). Second, qualitative results are presented holistically, whereas quantitative approaches present results in terms of specific variables dissected from the whole for close examination; qualitative studies emphasize the context of individuals' experiences, whereas quantitative studies tend to decontextualize the phenomena under study (Padgett, 2008). Third, quantitative research approaches tend to follow a positivist philosophy, seeking objectivity and representation of what actually exists; qualitative research approaches follow from a post-positivist philosophy, recognizing that observation is always shaped by the observer, therefore is always subjective in nature and this should be acknowledged and embraced. In post-positivist qualitative research traditions, realities are perceived as being socially constructed, whereas in positivist quantitative research, a single reality exists, waiting to be discovered or understood. The quantitative perspective on reality has a long tradition in the physical and natural sciences (physics, chemistry, anatomy, physiology, astronomy, and others). The social construction perspective has a strong hold in social science and understanding social phenomena. But what if an investigator's questions are relevant to both qualitative and quantitative approaches?

Given the fundamental philosophical and practical differences, some scholars argue that there can be no mixing of the approaches, that the underlying paradigms are too different. However, ***mixed-methods research*** has also been described as a new paradigm (since the 1980s) for social science:

> "*Like the mythology of the phoenix, mixed methods research has arisen out of the ashes of the paradigm wars to become the third methodological movement. The fields of applied social science and evaluation are among those which have shown the greatest popularity and uptake of mixed methods research designs*" (*Cameron & Miller*, 2007, *p.* 3).

Mixed-methods research approaches are used to address in a single study the acknowledged limitations of both quantitative and qualitative approaches. Mixed methods research combines elements of both qualitative and quantitative approaches for the purpose of achieving both depth and breadth of understanding, along with corroboration of results (Johnson, Onwuegbuzie, & Turner, 2007, p. 123). One mixed-methods strategy is related to the concept of ***triangulation***: understanding an event or phenomenon from the use of varied data sources and methods all applied to understanding the same phenomenon (Denzin & Lincoln, 2011; see Figure 1-1).

Figure 1-1. Depiction of triangulation as synthesis of different data sources



For example, in a survey research study of student debt load experienced by social work doctoral students, the investigators gathered quantitative data concerning demographics, dollar amounts of debt and resources, and other numeric data from students and programs (Begun & Carter, 2017). In addition, they collected qualitative data about the experience of incurring and managing debt load, how debt shaped students' career path decisions, practices around mentoring doctoral students about student debt load, and ideas for addressing the problem. Triangulation came into play in two ways: first, collecting data from students and programs about the topics, and second, a sub-sample of the original surveyed participants engaged in qualitative interviews concerning the "fit" or validity of conclusions drawn from the prior qualitative and quantitative data.

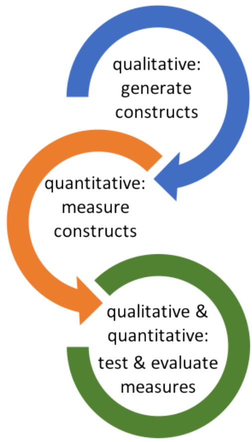
Three different types of mixed methods approaches are used:

- Convergent designs involve the simultaneous collection of both qualitative and quantitative data, followed by analysis of both

data sets, and merging the two sets of results in a comparative manner.
- Explanatory sequential designs use quantitative methods first, and then apply qualitative methods to help explain and further interpret the quantitative results.
- Exploratory sequential designs first explore a problem or phenomenon through qualitative methods, especially if the topic is previously unknown or the population is understudied and unfamiliar. These qualitative findings are then used to build the quantitative phase of a project (Creswell, 2014, p. 6).

Mixed methods approaches are useful in developing and testing new research or clinical measurement tools. For example, this is done in an exploratory sequential process whereby detail-rich qualitative data inform the creation of a quantitative instrument. The quantitative instrument is then tested in both quantitative and qualitative ways to confirm that it is adequate for its intended use. This iterative process is depicted in Figure 1-2.

Figure 1-2. Iterative qualitative and quantitative process of instrument development



One example of how this mixed-methods approach was utilized was in development of the Safe-At-Home instrument for assessing individuals' subjective readiness to change their intimate partner violence behavior (Begun et al., 2003; 2008). The transtheoretical model of behavior change (TMBC) underlies the instrument's development: identifying stages in readiness to change one's behavior and matching these stages to the most appropriate type of intervention strategy (Begun et al., 2001). The first step in developing the intimate partner violence Safe-At-Home instrument for assessing readiness to change was to qualitatively generate a list of statements that could be used in a quantitative rating scale. Providers of treatment services to individuals arrested for domestic or relationship violence were engaged in mutual teaching/learning with the investigators concerning the TMBC as it might relate to the perpetration of intimate partner violence. They independently generated lists of the kinds of statements they heard from individuals in their treatment programs, statements they believed were demonstrative of what they understood as the different stages in the change process. The investigators then worked with them to reduce the amassed list of statements into stage-representative categories, eliminating duplicates and ambiguous statements, and retaining the original words and phrases they heard to the greatest extent possible. The second phase was both quantitative and qualitative in nature: testing the instrument with a small sample of men engaged in batters' treatment programs and interviewing the men about the experience of using the instrument. Based on the results and their feedback, the instrument was revised. This process was followed through several iterations. The next phases were quantitative: determining the psychometric characteristics of the instrument and using it to quantitatively evaluate batterer treatment programs—the extent to which individuals were helped to move forward in stages of the change cycle.

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.3-1.1 Getting Started: Understanding an Excel Data File

Chapter Summary

In this chapter you were introduced to three general approaches for moving from research question to research method. You were provided with a brief overview of the philosophical underpinnings and uses of qualitative, quantitative, and mixed-methods approaches.

Next, you are provided with more detailed descriptions of qualitative and quantitative traditions and their associated methodologies.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=211

Module 3 Chapter 2: Overview of Qualitative Traditions

In Chapter 1 you read about three different approaches to social work research: qualitative, quantitative, and mixed-methods approaches. In this chapter we examine different traditions and methods commonly utilized in qualitative approaches to research. The qualitative research literature presents a wide array of approaches: ethnography, life histories, symbolic interactionism, grounded theory, phenomenology, hermeneutics/heuristic research, interpretivism, collaborative social research, social or cognitive anthropology, case studies, narrative inquiry, critical inquiry, participatory action research, content analysis, study of artifacts, and participant observation (Creswell & Poth, 2018). Here, we examine four of the five approaches identified by Creswell and Poth (2018), as well as two other approaches. You will read about:

- Narrative research
- Phenomenological research
- Grounded theory research
- Ethnographic research
- Participatory action research
- Content/artifact analysis research

Qualitative Research Approaches

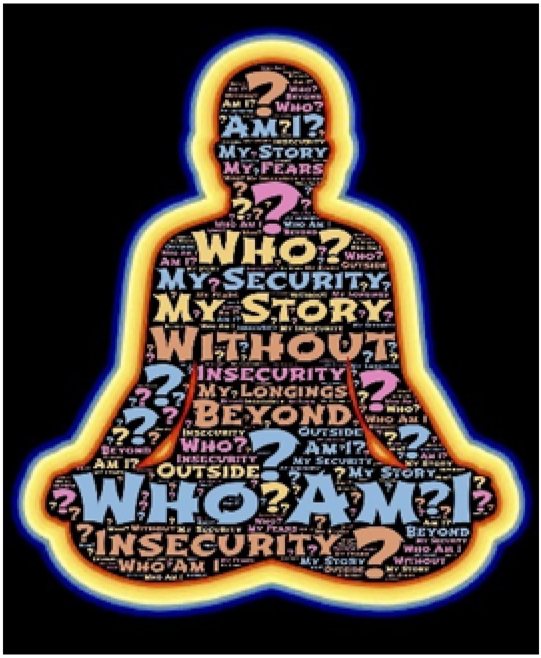
Qualitative research approaches are most appropriately used when the nature of the research questions and study goals meet certain criteria (Padgett, 2008). These include:

- Relatively little is known or understood about the topic of

interest
- The topic involves stigma and/or taboos, thereby lending itself to analysis with emotional depth and sensitivity (rather than quantifiable facts)
- The goal is to study and create meaning from the perspectives of individuals who have “lived the experiences” of study interest
- The goal is to understand processes, particularly complex processes, in social work problems and practices/interventions
- The goal is to interpret quantitative research results, particularly unexpected results
- Engaging in action or participatory research, involving social action or advocacy in combination with conducting the research.

Common across qualitative approaches is employing the “researcher-as-instrument” (Padgett, 2008, p. 17). Qualitative research results are heavily influenced by the experiences, perceptions, and personal qualities of the investigator. Thus, a critical component involves the investigator being reflective about the data, participants, analysis process, and results.

Narrative Research. Narrative research “begins with the experiences as expressed in lived and told stories of individuals” (Creswell & Poth, 2018). It is about both the nature of the expressed narratives and exploration of the contexts in which the narratives and narrators are situated. The method of gathering and interpreting these stories is called ***narrative inquiry***. This is typically performed with one or a very few individuals being involved with any one study. The purpose is to learn the individual stories and their meaning to individuals.



Of great interest in narrative inquiry is how a story is presented by the individual: it is seldom presented in a simple, chronological manner. Instead, stories are often conveyed in terms of critical transitions or nodal events where "turning points" occur in a person's lived experience. Narrative examples are present in biographies, autobiographies (autoethnography), and oral histories. For example, the movie documentary *Whitney*presents singer Whitney Houston's stories and lived experiences as a performer, celebrity, and family member, including her relationship with alcohol and other drugs. Parts of the documentary are presented as others' descriptions, parts are presented in her own words. The point is to better understand this person's life.

A social work example of narrative inquiry is reported as one woman's journey of recovery from mental illness (Thomas & Rickwood, 2016). Case interviews were conducted with a woman on three different occasions when she was (re)admitted to a residential mental health program. The woman's narratives demonstrated changes over time in hope, self-identity, a meaningful life, and assuming responsibility for one's own recovery. Furthermore, the investigators learned from the woman's perspective what elements of the services provided were helpful in her road to recovery and that repeated admissions were less indicative of repeated failures and more about progress toward recovery.

Phenomenological Research. Like narrative research, ***phenomenological research*** involves recording the stories of individuals who experience a particular phenomenon or event of interest. Unlike narrative inquiry, however, phenomenology is about deriving or constructing a shared meaning about those events among a group of participants: the goal is "meaning making" (Oktay, 2012). The emphasis in this approach is placed on identifying what the participants and their shared stories have in common. The emerging phenomenological description is a composite description of the phenomenon.

Ideally, the investigator engages with a heterogeneous group of individuals, all of whom have experienced the phenomenon of interest. For example, if the phenomenon of interest is about the experiences of siblings to persons with intellectual disabilities, the investigator would want to engage with families of different ages, social or economic standing, religious and cultural background, sibling composition (numbers, genders, and spacing of siblings), and individuals with different types or severity of disability.

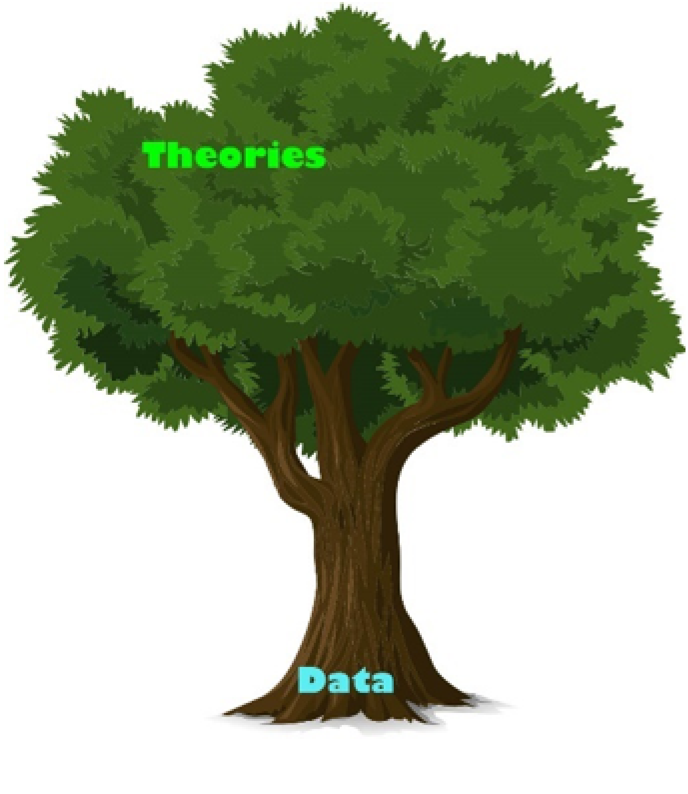
One example of this type of research is a qualitative study concerned with how trafficking affects the health and health care access among 15 women engaged in sex work in the red-light district (Kamathipura) of Mumbai, India (Karandikar, Gezinski, & Kaloga, 2016). The investigators worked with transcripts of interviews with the women, preserving the voice and language of the women who participated. They identified themes and relationships between themes in their coding of the data. Three themes were discussed in their manuscript:

how participants first entered into sex work (all but two reported deception by family members, friends, or acquaintances as the pathway to prostitution); health problems experienced by the women (physical violence, sexually transmitted infection, alcohol addiction, tuberculosis, HIV infection, miscarriage, abortion, cough/cold, and malaria were most frequently reported as notable events); and, disparities in access to health care resulting from being captively held in isolation or seclusion to prevent escape from trafficking and stigma attached to sex work.

Grounded Theory Research. ***Grounded theory research*** develops theoretical explanations about a phenomenon or process being studied—in other words, moving beyond description to generating or discovering a theory (Creswell & Poth, 2018). The theory that emerges can then be tested using other approaches and methods to determine its generalizability. The name of this approach comes from the theory being "grounded" in the data, rather than the data being used to test a theory that has been presented "off the shelf" (Creswell & Poth, 2018, p. 82). This approach, first described as a unified approach by Glaser and Strauss (1967), is particularly relevant to social work:

> *Because grounded theory creates theories that are derived directly from real-world settings, it has the potential to produce theories that can be used by social workers to guide practice. This theory can also be used to develop theoretically based interventions that can be tested in practice settings. In fact, when grounded theory was developed, the authors envisioned a collaborative venture between researchers and practitioners to test and adapt the theory in practice settings*" (Oktay, 2012, *p.* 5).

Grounded theory as an approach utilizes abductive reasoning, moving back and forth in an iterative manner between data gathering and analysis processes, including inductive and deductive logic in generating and cross-checking, testing the emerging theory (Oktay, 2012). Thus, theories grow from the grounded base of real-world data. The approach is heavily influenced by the discipline of sociology, but is used to inform other professions like social work and nursing. Study participants in grounded theory research are purposively selected because of their ability to provide relevant data.



An example of the grounded theory approach in use was presented in a study of professional decision-making processes undertaken by social workers in child protection roles (Kettle, 2018). Based on 22 interviews with social workers in Scotland, the investigator explored social workers' own perceptions of negotiating delicate balances between the interests of children and adults; the past, present, and future; functioning as an investigator versus building a relationship; getting too close versus not close enough; and, having power over versus power together. Also studied was the nature of their transactions and exchanges with other professionals. Emerging theory emphasized social workers' internal mental processes over rational-technical solutions for practice decisions, the role of perceived responsibility in degree of cooperation between professionals, and social workers' perceptions of different family strategies for managing closeness of and power in their relationships with the social worker.

Ethnographic Research. Like the qualitative approaches that we have explored so far, ethnographic research aims to understand phenomena from the perspectives of those experiencing it. However, there exists a major difference: the unit of analysis in ***ethnographic research*** is a collective group where the individuals interact with one another and have shared, common experiences. As a result, the individuals become a "culture-sharing group" by virtue of the "shared patterns of behavior, beliefs and language" that they develop (Creswell & Poth, 2018, p. 90). Ethnographic research focuses on cultural systems, "operating on several levels simultaneously to infer the tacit rules of the culture or subculture from the myriad of actions and interactions being witnessed" (Padgett, 2008, p. 32). The approach is strongly influenced by the discipline of anthropology, but is critically important as an option for social work to understand diverse populations, social work problems, and social phenomena.

Helm (2016) described an ethnographic study of how practice decisions are made within and across a team of social workers when the judgments are made under conditions of subjectivity and uncertainty, and when the available information is incomplete, inconclusive, or possibly contested. The aim of the study conducted in Scotland was to learn about "sense-making" under these circumstances, with attention directed to how this becomes a shared activity, rather than an individual activity. The study participants included a single team manager, four team leaders acting as supervisors, and six social workers from the larger team of professionals working at the study location. Data were collected through the investigator's direct,non-participatory observation of team member behavior, including what was said in the normal course of practice (***naturalistic observation***). The observation notes were coded for themes, the two strongest being "making sense through 'framing' the situation" and "constructing responsibility" (Helm, 2016, p. 29). Social work practitioners were observed to discuss cases with others, usually by opening the discussion with a framing statement, which then guided the discussion; for example, framing it in terms of the worker's own feelings about a case situation. Marked differences were noted between how social workers framed their discussions with each other, in supervision, and with clients. In addition, the social workers' sense-making occurred within the context of reference to their perceived professional roles and responsibilities. These perceptions were observed to influence decisions and actions by the social workers. The ethnographic observations help the profession better understand how practice decisions are made in complex, under-defined situations: in sense-making, individual practitioners rely on multiple sources of information, not the least of which is provided by their understanding of the professional context (roles and responsibilities) and input from co-workers and supervisors.

Ethnography might be an important approach for social workers to better understanding diverse populations in the ability/disability arena. Consider, for example, the meaning of a deaf culture. For hearing persons, not being able to hear and communicate through spoken language might seem like an overwhelming disability. For persons who have never been part of the dominant hearing culture, living and thriving in a deaf community is important and meaningful. Deaf culture is experienced as communicating and experiencing the world differently compared to the hearing community. Members of deaf culture do not view themselves as disabled and needing to be "fixed." Similar discussions have emerged among individuals who experience the world differently than the mainstream by virtue of their neurodiversity experience. These are individuals, many diagnosed with autism spectrum or other neurological disorders, who advocate for recognizing and appreciating variations in functioning as just that, not as mental disorders or disabilities needing to be cured. Ethnographic research helps social workers better understand their lived experiences from a strengths perspective.

Participatory Action Research. Sometimes social workers are members of a group attempting to solve a problem or influence change at the mezzo or macro level; sometimes members of such a group engage social workers to help them promote the problem solving and change-making processes. Either way, the key stakeholders in problem solving or advocating for change may decide to integrate research methodologies into their process. The result might be a ***participatory action research* (PAR)** scenario. This is a type of research activity where investigators are engaged in dual roles, both as researchers and change agents in the group or process under study.

PAR is conducted in the context of a non-hierarchical, democratic, engaged partnership of decision-making at every point in the process. Together, stakeholders and social workers develop the aims, questions, methods, analyses, results, and reporting of the change process. The outcomes of PAR are locally relevant solutions, developed by key stakeholders empowered to act as local experts and co-researchers supported by relevant theory and research knowledge shared by social work practitioners in a collaborative, problem-solving process. The results often include enhanced research skills among the key stakeholder partners as a form of further empowerment. What is learned through the PAR collaboration may have generalizable knowledge applicable to other communities, but the generalizability is carefully interpreted since it is, essentially, a single case study at the community level. Lawson (2015) described the relevance of PAR in community development efforts:

> *"Homogenization is rooted in part in mainstream researchers' claims about the wholesale generalizability of the knowledge gained from their investigations, and it is facilitated by the worldwide movement toward evidence-based policy and practice. Granting the importance, indeed the strengths, associated with this international development, there are manifest risks and dangers when local voices, choices, and knowledge are neglected, ignored, and discounted. Under these circumstances, rigorous research knowledge has the potential to serve as an instrument for domination, marginalization, and oppression. And when these conditions prevail, research-based knowledge has the potential to cause harm in the name of doing good work.* PAR's *expressly local knowledge for locally tailored solutions thus provides an important safeguard and a quality assurance mechanism for policy and practice"* (Lawson, 2015, *p. ix-x*).

***Community-based participatory research* (CBPR)** is a type of participatory research involving a collaborative inquiry process. As above, local experts are "people other than formally trained researchers" engaged to share their good ideas about every aspect of the research being conducted: the questions asked, study design features, how results are interpreted, and how findings should be used (Lawson, 2015, p. xv). University-community collaborations often provide the context through which CBPR occurs.

An example of CBPR was a needs assessment concerning parenting support programs for fathers in Detroit, Michigan (Lee, Hoffman, & Harris, 2016). The key stakeholders engaged in the effort with trained social work investigators from a university's school of social work included formal service delivery system providers, a community advisory board, other key informants, and fathers. The team proceeded through a five-stage community-engaged research process, and discussed the importance of operating as a "learning community" where team members function as both teachers and learners throughout the collaborative process.

Stage 1. Identification and assessment was centered around locating and interviewing formal service providers in the local community. As part of the interview process, additional providers were identified by the participants.

Stage 2. The team engaged in mapping nontraditional settings, as well. This was directed to identifying places where fathers might seek or receive parenting support outside of formal service delivery systems: barber shops, for example. They developed a map of the community served by formal and informal/nontraditional systems.

Stage 3. This phase was labeled “engagement & relationship building” (Lee, Hoffman, & Harris, 2016, p. 78). It involved creating a project community advisory board made up of key stakeholders, and developing working connections and relationships with researchers involved with this particular community.

Stage 4. At this point, key stakeholder and university partners collaborated to develop strategies and events to build the service providers’ capacity to provide parenting support as needed and desired by fathers in the community.

Stage 5. The final, evaluation stage involved trained investigators coding data, then stakeholders involved in checking/confirming the themes identified in the data (called “member checking” of results). This final phase also included dissemination of the study findings.

In the end, this CBPR project identified multiple themes and at least 15 recommendations. Service providers identified service needs of the community’s fathers and strategies for engaging fathers. The fathers identified themes around engaging with children and interacting with service providers. The recommendations are specific with regards to what can change to provide support for the parenting efforts of fathers in this community, and many are relevant to consider adopting in other communities.

Content or Artifact Analysis Research. ***Content analysis*** is not so much a qualitative approach as it is a means of analyzing qualitative data. The important distinction lies in the sources of data compared to studies where interviews are conducted. In content or artifact analysis, the investigator's source of information is records or artifacts previously produced during or about the phenomenon of interest. Content analysis is typically conducted on artifacts that were created for naturally occurring reasons, rather than for purposes of the study being conducted. Historically, content analysis was frequently conducted with personal diaries and letters written to family members or close associates. Examples include studies of first-hand report artifacts concerning how military personnel experienced military combat or their experiences of being separated from family members.

Given the important role of media on socialization, contents of media to which individuals are exposed might be analyzed for common themes and messages that audiences might be receiving. For example, a social worker might be concerned with learning about the contents of books commonly read to young children in families expecting a new baby. Parents read these books as a form of "bibliotherapy" to inform their children about what to expect and socialize them about how to behave in relation to the expected or new family member. A content analysis of these books divulged a common structure and theme: the new baby's arrival is initially met with a great deal of negative affect (jealousy, feeling neglected by parents, being annoyed by the baby's behavior), then in the end the child has an epiphany whereby the baby is fun and becomes loved (Begun, unpublished). Unfortunately, this plot structure and theme approach is adult-centric–young children learn concepts not by the logical sequencing of events but by what is most salient, affect-laden, or recent. Therefore, the message to these children could be "don't like the baby" as much as it is "like the baby" when exposed to these books.

Recent literature includes studies relying on content presented in social media postings as the studied artifacts (Facebook, twitter, Snapchat, MySpace, and others). For example, investigators reviewed the contents of more than 1,000 MySpace comments posted by adolescents and emerging adults (aged 13-24 years) for the presence of potential suicide statements (Cash et al, 2013). The content analysis resulted in identifying 64 messages where a theme of potential suicidality was present; these statements were associated with subthemes about relationships, mental health concerns, substance misuse, and suicide methods. The lesson learned from this study is that young people are expressing themselves about suicide-related thoughts and behaviors to their social networks using social networking sites. One implication is that individuals using social network sites might need to be educated about how to recognize, interpret, and respond to these types of messages to become part of a suicide prevention network.

Chapter Summary

In this chapter you read about different traditions of qualitative research. You were introduced to narrative, phenomenological, grounded theory, ethnographic, participatory action, and content/artifact analysis research traditions. In future chapters you will learn about the specific methods involved in data collection and engaging participants in qualitative research studies.

Module 3 Chapter 3: Overview of Quantitative Traditions

Just as there were multiple traditions under the umbrella heading of qualitative research approaches, different types of quantitative research approaches are used in social work, social, and behavioral sciences. In this chapter you are introduced to:

- cross-sectional research
- longitudinal research
- descriptive research
- correlational research
- experimental research.

Cross-Sectional and Longitudinal Research

Research for understanding diverse populations, social work problems, and social phenomena might be conducted either cross-sectionally or longitudinally. As you will see, the decision about using a ***cross-sectional research*** or ***longitudinal research*** design has a lot to do with the nature of the research questions being addressed, the type of data available to investigators, and a weighing of the advantages and disadvantages of either approach.

Cross-Sectional Research Design. A cross-sectional study involves data collection at just one point in time. It lends itself to research questions that do not relate to change over time but are descriptive, exploratory, or explanatory about a specific point in time. In a study addressing an epidemiological or other descriptive question, for example, the cross-sectional design allows investigators to describe a phenomenon at a single time point. An example might be a study of the prevalence of individuals' exposure to gun violence in their communities and ratings of collective efficacy for the community (Riley et al., 2017). Survey data were collected from participants at one point in time (2014) in two communities with high rates of violent crime. Collective efficacy was conceptualized as each individual's perception about formal and informal cohesion in their neighborhood, the strength of social bonds among neighbors, and the ability of community members to achieve public order through informal mechanisms. The team hypothesized an inverse relationship between these two factors: that higher rates of gun violence exposure would correlate with lower scores for perceived collective efficacy. Responses on 153 surveys were analyzed, demonstrating that exposure to gun violence was very high (95% had heard gunshots, 21% heard gunshots at least weekly, over 1/3 had been present when someone was shot). The hypothesized relationship between individuals' higher exposure to gun violence and their lower perceptions of community's collective efficacy was observed.

Cross-sectional research designs are also used in comparison studies where a phenomenon's expression is compared for two or more groups—groups differing in race or ethnicity, sex or gender identity, geographical location, or other social demographic factors, for example. These studies might also compare groups of individuals who received a certain intervention and those who did not. Or, this design might compare groups of individuals who differ in severity or other characteristics of social work or behavioral health problem.

A cross-sectional design was used to answer a question concerning how bipolar disorder might differ from major depressive disorder outside of major mood episodes (Nilsson, Straarup, & Halvorsen, 2015). Investigators compared 49 individuals with bipolar disorder to 30 participants with major depressive disorder at a point in time when each was in remission. Participants were compared on 15 "early maladaptive schemas" which included their self-ratings of abandonment, mistrust, social isolation, defectiveness, failure to achieve, dependence, emotional inhibition, insufficient self-control, and others. The investigators observed statistically different scores between the two groups on 7 of the early maladaptive schemas, with the bipolar diagnosis being associated with less favorable ratings. This evidence contributes to a clinical understanding of differences between these two diagnosed mood disorders, differences with potential implications for intervention, especially with those diagnosed with bipolar disorder.

In other studies, the groups being compared using a cross-sectional research design might also represent different age groups. The conclusions drawn from such a study could be intended to help develop or test theory concerning the developmental course of a phenomenon or changes in it over time.

A cross-sectional study example explored the potential impact of young children having been exposed to intimate partner violence (Bowen, 2017). The investigator utilized one data point (when the children were four years old) to assess the presence of conduct disorder symptoms among children whose mothers reported their partner having been emotionally cruel and/or having physically hurt them during the child's first 33 months. Among 7,743 children, the mothers of 18.4% of girls and 17% of boys reported experiencing a measured form of intimate partner violence during that developmental period. Conduct problems were significantly more common among these children compared to children whose mothers reported no exposure to intimate partner violence: 2.12 times more common among exposed boys and 2.85 times more common among exposed girls. The source of data for this study was longitudinal: one cohort of families participated in data collection at multiple time points during the course of the child's development; however, the study results reported by Bowen (2017) were based on one time point alone, making it a cross-sectional study.

The cross-sectional study presents a single snapshot description of the group's aggregated ratings. A longitudinal study design is more sensitive to fluctuations over time and to variation patterns experienced by individual couples, presenting a more video-like description. When the effect of time (or age) is of interest, conclusions from cross-sectional studies must be drawn with caution since many alternative explanations for any observed group differences exist, most of which have little to do with the passage of time or developmental changes. That is where longitudinal designs are more helpful.

Longitudinal Research Design. A longitudinal study involves data collection from the same participants at two or more points in time. The beauty of this approach is that it allows investigators to directly address questions involving change over time. Longitudinal studies are a common means of studying or evaluating social work interventions, for example. This design strategy is also common in studies about human developmental processes. For example, a longitudinal study design was used to study behavior and cognitive functioning in 189 young adolescents who had experienced prenatal cocaine exposure and 183 who had not (Minnes et al., 2016). Data were collected when these individuals were 12 years old, and again when they were aged 15. At age 12, girls with prenatal cocaine exposure had more problems regulating their behavior and with metacognition than did girls without prenatal cocaine exposure; there was no observed difference between the groups for boys. At this point, the study conclusions could have been derived from a cross-sectional study comparing the two groups at age 12 alone. However, the investigators also reported that both behavioral and metacognition problems improved from age 12 to 15 for the prenatally cocaine exposed girls at a greater rate than the other groups. While some differences did exist at age 15, it is important to note that the conclusions that might have been drawn from the cross-sectional element of the study (girls at age 12) are different from conclusions drawn from the longitudinal report comparing ages 12 and 15.

The group of individuals being followed in a longitudinal study are called a study ***cohort***. In the above example, the cohort was a group of babies born 12 years before the analyzed data were first collected. This group continued to be followed at age 15. In addition to studying a birth cohort like this, investigators might study other than birth cohorts. For example, an incoming class of BSW students might be followed each semester until graduation from the program. Or, investigators might longitudinally follow a cohort of veterans and their families as they return from deployment at about the same time. Or, investigators might follow a group of survivors of a traumatic event such as the 9/11 collapse of the twin towers in New York City or the Washington, DC attack on the Pentagon. The concept of "cohort" indicates that the study participants share a common experience or event.

Comparing Cross-Sectional and Longitudinal Designs. These two study designs result in somewhat different perspectives even if the same phenomenon is being examined; much as an orange sliced across its equator looks different from one sliced from top to bottom.

© A. Begun

Each design approach is associated with advantages and disadvantages that come under consideration by investigators as they plan their studies. Let's take a closer look at what these considerations might include.

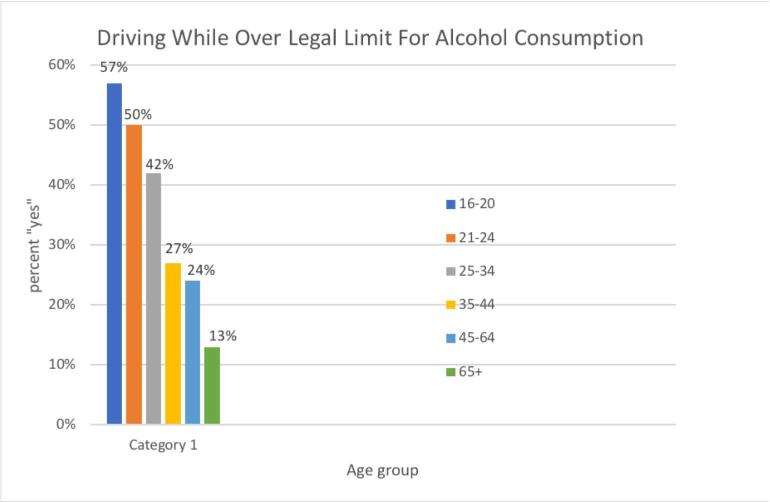
Individual variation. Longitudinal studies, compared to cross-sectional studies, are more sensitive to the unique, diverse patterns of behavior observed for individuals over time. In cross-sectional studies, individual differences are blended together for each group being compared. In longitudinal studies, behavior at one point in time is maintained within the context of behavior at prior and later points in time. This provides useful information about how individuals change.

Adaptability. One limitation of longitudinal studies is that investigators are "stuck" with continuing to use the same (or very similar) measurement approach throughout the life of a study. Measurement consistency ensures that differences at different time points are a function of time, rather than different measurement approaches. Unfortunately, the science of measurement may progress over the course of a longitudinal study so that by the end of the study, a better way to measure the variables of interest may have emerged. A cross-sectional study can capitalize on the best, latest measurement strategy available at the time the study is conducted. The longitudinal study must continue to use what was the best, latest measurement strategy at the start of the study, regardless of progress in measurement science that has occurred by the end of the study period.

*Concurrent versus retrospective measurement.*If investigators are addressing questions about change, data collected concurrently at each point in time may be more reliable than data collected at one point asking about previous points in time. This latter retrospective data may be contaminated by memory or other sources of bias, making inferences about change less accurate.

Time. Imagine that investigators wanted to know how driving after drinking excessive amount of alcohol might change as people age. A study conducted in 2008 demonstrated that the percent of persons who did so during the prior 12 months was highest among persons aged 16-20 years (57%), and declined with each subsequent age group to those aged 65 and older (see Figure 3-1). This cross-sectional study, covering an age span of more than 50 years was conducted in about four months. If the same question were to be studied longitudinally, following the 16-year-old participants until they were aged 65, and asking the question at intervals along the way, it would take every bit of those 50 years to complete. Sometimes longitudinal studies are simply impractical to undertake because of the time involved.

Figure 3-1. Percent driving during past 12 months when thought to be over the legal limit for alcohol consumption (N=1466 drivers who also consume alcohol, adapted from NHTSA, 2010).



Attrition. Investigators conducting cross-sectional studies, because they only involve one measurement point for each study participant, are not concerned with losing participants over time. Participant attrition (drop-out) from longitudinal studies is a very serious concern, and a great deal of effort needs to be directed toward retaining each participant longitudinally over time. Participants disappear for a number of reasons: losing interest in study participation, the condition or phenomenon being studied changes, moving out of the area, becoming incarcerated or otherwise being lost to study continuation/follow-up. Furthermore, there exists some degree of investigator turn-over/attrition possible, as well.

Cost. Longitudinal studies may be more costly than cross-sectional studies to implement because of the need to maintain participants' involvement over time (retention).

Statistical analyses. Different types of analyses need to be conducted on data that are longitudinal compared to cross-sectional data. This is neither an advantage nor disadvantage, just a difference worthy of note.

Descriptive, Correlational, and Experimental Research

Back in Module 2, you were introduced to research questions of an exploratory, descriptive, or explanatory nature. Here we examine the traditions in quantitative research that help address these different types of research questions: descriptive, correlational, and experimental studies.

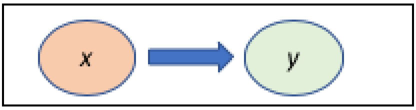
Descriptive Research. The aim of a ***descriptive research*** study is either to create a profile or typology of a population, or to describe a phenomenon or naturally occurring process (Grinnell & Unrau, 2014; Yegidis, Weinbach, & Meyers, 2018). Descriptive research contributes important information for understanding diverse populations, social work problems, and social phenomena. An example of descriptive research was a study aimed at identifying different substance use patterns among pregnant adolescents (Salas-Wright, Vaughn, & Ugalde, 2016). The study investigators started with an understanding from previous research that there exists an overall pattern of elevated substance use levels prior to conception compared to non-pregnant peers, and of reduced use during pregnancy. However, these investigators were concerned that this general, aggregate picture of pregnant adolescents might mask different patterns (heterogeneity) having important social work practice implications. The team analyzed substance use data for 810 pregnant adolescents, aged 12-17. They were able to discern and describe four different pattern types: abstaining from any substance use, using only alcohol, using both alcohol and cannabis, and polydrug use. Those who only drank alcohol prior to conception tended to cut back or eliminate drinking during pregnancy. Those who used alcohol and cannabis or engaged in polydrug use were fewer in number than those who only used alcohol, but also were less likely to stop using substances during pregnancy. Those who engaged in polydrug use were the youngest group and most likely to meet criteria for a substance use disorder. The different substance use types were not evenly distributed by demographics (ethnicity, age, family income) or pregnancy stage across study participants. The results of this descriptive study have social work and public health implications related to intervening in different ways with adolescents of different types to prevent or reduce prenatal exposure to substances.

Correlational Research. The aim in some studies is to evaluate the existence and nature ofrelationships that might exist between variables—looking to see if *variable x*and *variable y*are associated or correlated with each other. These are called ***correlational research*** studies. As mentioned in reference to descriptive studies, the investigator does not manipulate variables experimentally: naturally occurring relationships among variables is the focus of study. An interesting approach to understanding social work problems and social phenomena involves the use of geospatial data. You may have heard the term GIS—short for geographic information system—which is about data that have a geographical, spatial component. This type of data is useful for social workers to understand how a problem or phenomenon might be experienced in the environmental contexts where people live, work, recreate, or otherwise function.

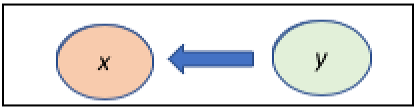
A correlational study using geospatial data examined relationships between use of and physical access to marijuana with child maltreatment (Freisthler, Gruenewald, & Price Wolf, 2015). First, the investigators demonstrated that parental use of marijuana was positively correlated with the frequency of child physical abuse, negatively correlated with physical neglect, and not correlated with supervisory neglect. Parents reporting marijuana use reported engaging in physical abuse three times more frequently than parents who did not use marijuana, with boys more likely than girls to be physically abused, and older children more likely to be physically abused than younger children. The data for these results came from telephone interviews with 3,023 persons in a state with "legalized" medical marijuana dispensaries and delivery services. The geospatial data came into play in analyses indicating that greater geographical density of medical marijuana dispensaries and delivery services was positively related to greater frequency of physical child abuse. These findings have important implications for social work practice and social policy, including that marijuana use may be problematic even if a marijuana use disorder is not diagnosed in the child welfare context, and that there may exist differences in the impact of marijuana use in communities where distribution is legitimized, relegated to illicit “street-level dealers,” or not present.

Correlation versus Causality. The major limitation of correlational research is that the tradition is unable to definitively determine causality in observed relationships between variables of interest. When a correlation is observed between two variables, x and y, at least four possibilities exist:

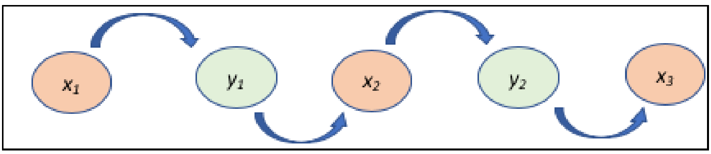
- x influences y



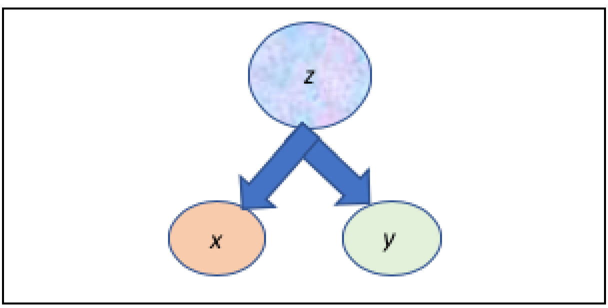
- y influences x



- x and y influence each other in a more complex, circular, iterative manner, where x influences y which influences x which influences y and so forth



- some third variable, z, causes both x and y but these two do not directly influence each other

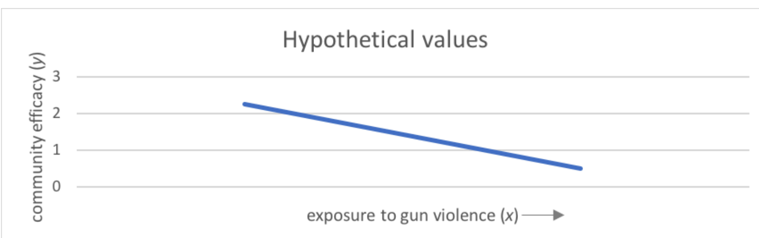


To logically draw conclusions about causality, at least three conditions need to be met.

1. Evidence exists that the two variables are related—as values for the "cause" change, values for the "effect" also change systematically in relation to those changes in the "cause." This is what the correlation statistic can demonstrate: changes in $x$relate to changes in y. In addition, that the two are related must be consistently demonstrated—that changing the "cause" consistently (not

arbitrarily) results in "effect" changes (Cozby, 2007). Consider the example presented earlier about exposure to gun violence and perception of community efficacy (Riley et al., 2017). The investigators observed that as ratings of exposure to gun violence increased (the *x*or hypothesized "cause" variable), ratings of perceived collective efficacy decreased (the y or hypothesized "effect" variable). This kind of relationship between the variables is called a negative or inverse correlation–as one goes up the other goes down. Depending on the actual values of the relationship, it might look something like Figure 3-2 where more exposure to gun violence (moving from left to right) is associated with lower community efficacy.

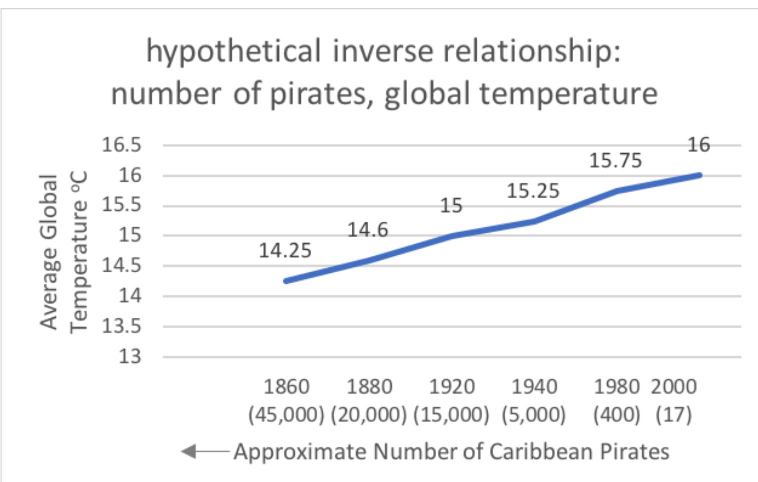
Figure 3-2. Depiction of negative (inverse) correlation using hypothetical values



2. The "cause" and "effect" must be sequenced in time, with the "cause" preceding (coming before) the "effect."For example, consider the earlier-mentioned Bowen (2017) study concerning conduct disorder among four-year-old children who have been exposed to a mother's experience of intimate partner violence. Before concluding that the exposure caused the conduct disorder symptoms, it is

important to demonstrate that the children's conduct disorder symptoms did not predate the exposure to intimate partner violence. If the conduct disorder came first, it is quite possible that the stress of managing a child with these behaviors influenced the intimate partner violence (y caused x) instead.

3. There is not another, third factor that "causes" both other factors. A demonstrative example comes from a purposefully absurd graph attributed on the internet to Bobby Henderson, used to demonstrate that correlation does not imply causation. The erroneous conclusion drawn from the data is that global warming (y variable) is a direct effect of the shrinking number of pirates (variable x). As you can see from Figure 3-3, between the years 1860 and 2000 the number of Caribbean pirates declined from 45,000 to 17, and in the same period, the average global temperature climbed from 14.25 degrees Celsius to 16 degrees Celsius ("data" retrieved from https://en.wikipedia.org/wiki/Flying_Spaghetti_Monster).



While the "data" indicate an inverse association between the two variables (notice that the approximate number of pirates is declining as year is increasing on the *x* axis), it is absurd to conclude that the reduction in pirate numbers caused global warming. A third variable might be causing both: for example, the mechanism of change in both might be related to the global spread of industrialization.

In summary, correlational research studies allow investigators to identify possible associations between factors, helping to develop better understanding about diverse populations, social work problems, and social phenomena. However, additional research is necessary to determine if there are causal relationships involved. This is important if we are to use evidence to inform the interventions we develop and test. If we do not understand the causal nature of the relationships and the mechanisms by which observed effects are caused, then we cannot determine the best places and ways to intervene. This understanding is enhanced through experimental research.

Experimental Research

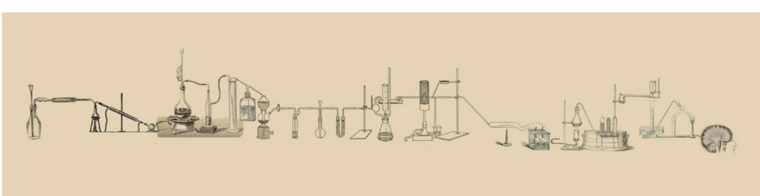
The aim of ***experimental research*** is to answer explanatory questions: to test hypotheses or theory about phenomena and processes. This includes testing theory and social work interventions. Our next course, SWK 3402, emphasizes research and evaluation to understand interventions. Here we focus on experimental research to understand diverse populations, social work problems, and social phenomena. Experiments are designed to help develop our understanding of causal relationships between factors, the etiology of social work problems and social phenomena, and the mechanisms by which change might occur. Investigators develop and discover this kind of information through systematic manipulation of specific factors and observing the impact of those manipulations on the outcomes of interest. The study designs and methods applied in experimental research are selected to eliminate as many alternative explanations for the results as possible, thereby increasing confidence that the tested mechanism or explanation is accurately accepted: "Explanatory research depends on our ability to rule out other explanations for our findings" (Engel & Schutt, 2013, p. 19). This observation leads naturally to a discussion of ***internal validity*** and ***external validity*** in experimental research.

Internal Validity. The concept of internal validity refers to the degree of confidence that can be applied to the results of an experimental study. Remember, the purpose of an experiment is to develop an understanding of causal relationships between variables—in other words, to test explanatory hypotheses. Strong internal validity means greater confidence in drawing causal or explanatory conclusions from the data; this is sometimes referred to as study integrity. Experimental design is sometimes referred to as "the gold standard" against which all others are compared. This, however, would only be true if the aim of those being compared is explanatory—many other designs are important, depending on the nature of the research questions and aims. If, however, the aim is explanatory, designing a study that allows for investigator control over the greatest number of possible explanatory factors, ruling out alternative explanations for the observed experimental results, is most desirable—it means the study has strong internal validity.

External Validity. External validity is about the degree to which a study's findings can be generalized to the population represented by that sample. External validity is again about confidence in the results, but this time it is about how well the results from a particular study represent a general reality for other similar people and other similar settings. For example, on attitudes about social issues, studies conducted with college student samples from private universities may poorly represent the nation's population of college-aged adults, or even the population of the nation's college students. Similarly, studies about a social work problem conducted with people in treatment or receiving services might not be representative of people experiencing the same problem but who are not in treatment and may never have sought treatment for the problem. This is called the clinical sample problem related to external validity concerns. As you will see in Chapter 6, external validity is powerfully influenced by the procedures involved in developing and retaining a study sample, including (but not limited to) problems of small sample size.

The Monitoring the Future study represent an example of strong cross-sectional survey methodology, contributing to the study's high external validity. The Monitoring the Future study has been conducted with 12thgrade students every year since 1975, and with 8th, 10th, and 12thgraders since 1991—you may very well have participated in the study at some point in your life. Some young adults are being surveyed in more recent years, as well. The purpose is to study trends in U.S. adolescents' beliefs, attitudes, and behavior. To enhance generalizability of the annual findings, about 50,000 students are surveyed each year. These students are from about 420 public and private high schools across the nation, and include about 18,000 8thgraders from about 150 schools, about 17,000 10thgraders, and about 16,000 12thgraders from about 133 schools. A randomly selected subset of participants has been followed longitudinally every two years since 12thgrade, as well. The geographic regions are randomly selected each year, with random selection of the schools in the selected areas, and random selection of the classes to be included within each school. Random selection is a strong external validity contributor. You may be interested to see some of the trend results reported by the study investigators, located at http://www.monitoringthefuture.org/.

If you have enjoyed television shows like *Bill Nye the Science Guy*and *Mythbusters*, then you have witnessed systematic experiments being implemented. Their demonstrations typically include elements of scientific process designed to enhance rigor of the experiments under way. A great deal of social science, behavioral science, and social work research follows scientific methods and logic developed in physical and natural sciences. You might wonder why all social work research is not conducted under these scientific, experimental procedures. The simple answer is that a great deal of social work research is designed to address other types of questions: experimental research is designed to answer explanatory questions, not the other types that we have discussed.



A more complex answer includes that many possible experiments cannot be practically or ethically conducted. For example, we cannot ethically expose people to conditions that might cause them harm. To understand the impact of disrupted parent-child attachment bonds we cannot experimentally manipulate this condition to see what happens; instead, we need to rely on observing the effects through what happened "naturally," such as when policy led to children and immigrant/refugee parents being separated at U.S. borders while parents were detained for having entered the country illegally. Similarly, we could not ethically conduct controlled experiments where children are randomly assigned to being raised by parents either with or without substance misuse problems, where families are randomly assigned to living in homes with or without lead contamination, or where communities are randomly assigned to having or not having access to affordable, healthy food. Abuse and ethical violations in experiments conducted with unknowing, under-informed, unempowered, and otherwise vulnerable persons led to the development of policies, guidelines, and practices discussed in your CITI training about research involving human subjects—including the Belmont Report and Institutional Review Boards.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=213

Chapter Summary

This chapter introduced the traditions of quantitative research. You read about the distinctions, advantages, and disadvantages of cross-sectional and longitudinal study designs. You also read about descriptive, correlational, and experimental studies, and were reminded that correlation does not imply causation. Finally, the topics of internal and external validity were examined. The next chapter further develops a foundation for understanding the wide range of available options and decisions to be made by investigators planning experimental research. The focus includes understanding the nature of variables involved in quantitative studies, whether they are descriptive, exploratory, or experimental in nature.

Module 3 Chapter 4: Overview of Quantitative Study Variables

The first thing we need to understand is the nature of variables and how variables are used in a study's design to answer the study questions. In this chapter you will learn:

- different types of variables in quantitative studies,
- issues surrounding the unit of analysis question.

Understanding Quantitative Variables

The root of the word ***variable*** is related to the word "vary," which should help us understand what variables might be. Variables are elements, entities, or factors that can change (vary); for example, the outdoor temperature, the cost of gasoline per gallon, a person's weight, and the mood of persons in your extended family are all variables. In other words, they can have different values under different conditions or for different people.

We use variables to describe features or factors of interest. Examples might include the number of members in different households, the distance to healthful food sources in different neighborhoods, the ratio of social work faculty to students in a BSW or MSW program, the proportion of persons from different racial/ethnic groups incarcerated, the cost of transportation to receive services from a social work program, or the rate of infant mortality in different counties. In social work intervention research, variables might include characteristics of the intervention (intensity, frequency, duration) and outcomes associated with the intervention.

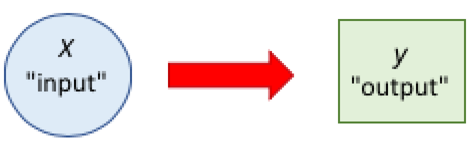
Demographic Variables. Social workers are often interested in what we call ***demographic variables.*** Demographic variables are used to describe characteristics of a population, group, or sample of the population. Examples of frequently applied demographic variables are

- age,
- ethnicity,
- national origin,
- religious affiliation,
- gender,
- sexual orientation,
- marital/relationship status,
- employment status,
- political affiliation,
- geographical location,
- education level, and
- income.

At a more macro level, the demographics of a community or organization often includes its size; organizations are often measured in terms of their overall budget.

Independent and Dependent Variables. A way that investigators think about study variables has important implications for a study design. Investigators make decisions about having them serve as either ***independent variables*** or as ***dependent variables***. This distinction is not something inherent to a variable, it is based on how the investigator chooses to define each variable. Independent variables are the ones you might think of as the manipulated "input" variables, while the dependent variables are the ones where the impact or "output" of that input variation would be observed.

Intentional manipulation of the "input" (independent) variable is not always involved. Consider the example of a study conducted in Sweden examining the relationship between having been the victim of child maltreatment and later absenteeism from high school: no one intentionally manipulated whether the children would be victims of child maltreatment (Hagborg, Berglund, & Fahlke, 2017). The investigators hypothesized that naturally occurring differences in the input variable (child maltreatment history) would be associated with systematic variation in a specific outcome variable (school absenteeism). In this case, the independent variable was a history of being the victim of child maltreatment, and the dependent variable was the school absenteeism outcome. In other words, the independent variable is hypothesized by the investigator to cause variation or change in the dependent variable. This is what it might look like in a diagram where "*x*" is the independent variable and "*y*" is the dependent variable (note: you saw this designation earlier, in Chapter 3, when we discussed cause and effect logic):



For another example, consider research indicating that being the victim of child maltreatment is associated with a higher risk of substance use during adolescence (Yoon, Kobulsky, Yoon, & Kim, 2017). The independent variable in this model would be having a history of child maltreatment. The dependent variable would be risk of substance use during adolescence. This example is even more elaborate because it specifies the pathway by which the independent variable (child maltreatment) might impose its effects on the dependent variable (adolescent substance use). The authors of the study demonstrated that post-traumatic stress (PTS) was a link between childhood abuse (physical and sexual) and substance use during adolescence.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=222

Types of Quantitative Variables

There are other meaningful ways to think about variables of interest, as well. Let's consider different features of variables used in quantitative research studies. Here we explore quantitative variables as being categorical, ordinal, or interval in nature. These features have implications for both measurement and data analysis.

Categorical Variables. Some variables can take on values that vary, but not in a meaningful numerical way. Instead, they might be defined in terms of the categories which are possible. Logically, these are called ***categorical variables***. Statistical software and textbooks sometimes refer to variables with categories as ***nominal variables***. Nominal can be thought of in terms of the Latin root "nom" which means "name," and should not be confused with number. Nominal means the same thing as categorical in describing variables. In other words, categorical or nominal variables are identified by the names or labels of the represented categories. For example, the color of the last car you rode in would be a categorical variable: blue, black, silver, white, red, green, yellow, or other are categories of the variable we might call car color.

What is important with categorical variables is that these categories have no relevant numeric sequence or order. There is no numeric difference between the different car colors, or difference between "yes" or "no" as the categories in answering if you rode in a blue car. There is no implied order or hierarchy to the categories "Hispanic or Latino" and "Not Hispanic or Latino" in an ethnicity variable; nor is there any relevant order to categories of variables like gender, the state or geographical region where a person resides, or whether a person's residence is owned or rented.

If a researcher decided to use numbers as symbols related to categories in such a variable, the numbers are arbitrary—each number is essentially just a different, shorter name for each category. For example, the variable gender could be coded in the following ways, and it would make no difference, as long as the code was consistently applied.

Coding Option A	Variable Categories	Coding Option B
1	male	2
2	female	1
3	other than male or female alone	4
4	prefer not to answer	3

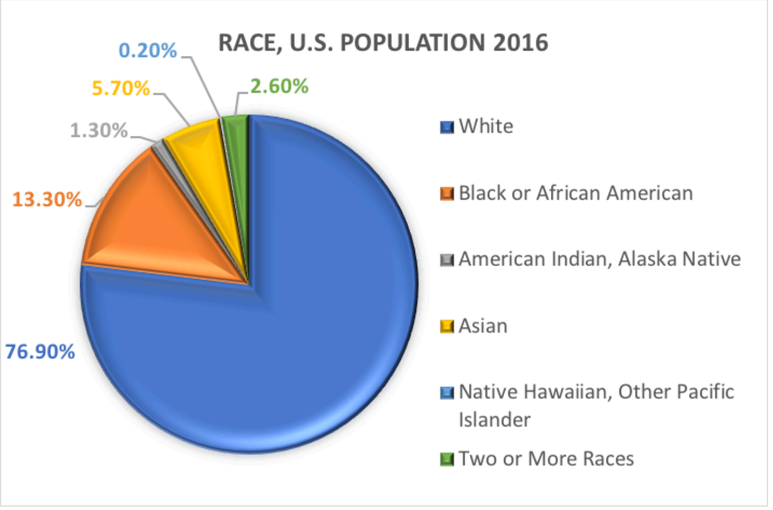
*Race and ethnicity.*One of the most commonly explored categorical variables in social work and social science research is the demographic referring to a person's racial and/or ethnic background. Many studies utilize the categories specified in past U.S. Census Bureau reports. Here is what the U.S. Census Bureau has to say about the two distinct demographic variables, race and ethnicity (https://www.census.gov/mso/www/training/pdf/race-ethnicity-onepager.pdf):

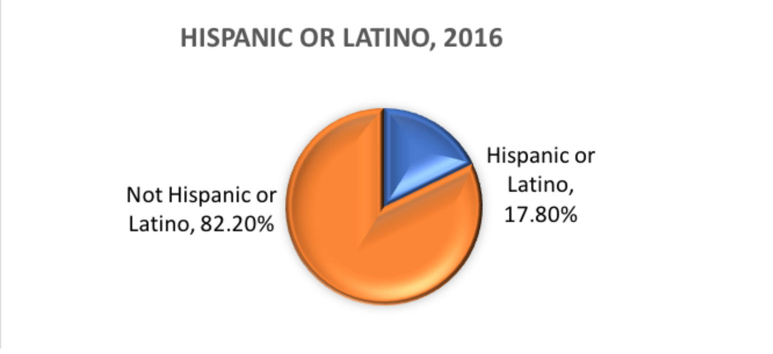
> *What is race? The Census Bureau defines race as a person's self-identification with one or more social groups. An individual can report as White, Black or African American, Asian, American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, or some other race. Survey respondents may report multiple races.*

> *What is ethnicity? Ethnicity determines whether a person is of Hispanic origin or not. For this reason, ethnicity is broken out into two categories, Hispanic or Latino and Not Hispanic or Latino. Hispanics may report as any race.*

In other words, the Census Bureau defines two categories for the variable called ethnicity (Hispanic or Latino and Not Hispanic or Latino), and seven categories for the variable called race. While these variables and categories are often applied in social science and social work research, they are not without criticism.

Based on these categories, here is what is estimated to be true of the U.S. population in 2016:





*Dichotomous variables.*There exists a special category of categorical variable with implications for certain statistical analyses. Categorical variables comprised of exactly two options, no more and no fewer are called ***dichotomous variables***. One example was the U.S. Census Bureau dichotomy of Hispanic/Latino and Non-Hispanic/Non-Latino ethnicity. For another example, investigators might wish to compare people who complete treatment with those who drop out before completing treatment. With the two categories, completed or not completed, this treatment completion variable is not only categorical, it is dichotomous. Variables where individuals respond "yes" or "no" are also dichotomous in nature.

The past tradition of treating gender as either male or female is another example of a dichotomous variable. However, very strong arguments exist for no longer treating gender in this dichotomous manner: a greater variety of gender identities are demonstrably relevant in social work for persons whose identity does not align with the dichotomous (also called binary) categories of man/woman or male/female. These include categories such as agender, androgynous, bigender, cisgender, gender expansive, gender fluid, gender questioning, queer, transgender, and others.

Ordinal Variables. Unlike these categorical variables, sometimes a variable's categories do have a logical numerical sequence or order. Ordinal, by definition, refers to a position in a series. Variables with numerically relevant categories are called ***ordinal variables.*** For example, there is an implied order of categories from least-to-most with the variable called educational attainment. The U.S. Census data categories for this ordinal variable are:

- none
- 1st-4thgrade
- 5th-6thgrade

- 7th-8thgrade
- 9thgrade
- 10thgrade
- 11thgrade
- high school graduate
- some college, no degree
- associate's degree, occupational
- associate's degree academic
- bachelor's degree
- master's degree
- professional degree
- doctoral degree

In looking at the 2016 Census Bureau estimate data for this variable, we can see that females outnumbered males in the category of having attained a bachelor's degree: of the 47,718,000 persons in this category, 22,485,000 were male and 25,234,000 were female. While this gendered pattern held for those receiving master's degrees, the pattern was reversed for receiving doctoral degrees: more males than females obtained this highest level of education. It is also interesting to note that females outnumbered males at the low end of the spectrum: 441,000 females reported no education compared to 374,000 males.

Here is another example of using ordinal variables in social work research: when individuals seek treatment for a problem with alcohol misuse, social workers may wish to know if this is their first, second, third, or whatever numbered serious attempt to change their drinking behavior. Participants enrolled in a study comparing treatment approaches for alcohol use disorders reported that the intervention study was anywhere from their first to eleventh significant change attempt (Begun, Berger, Salm-Ward, 2011). This change attempt variable has implications for how social workers might interpret data evaluating an intervention that was not the first try for everyone involved.

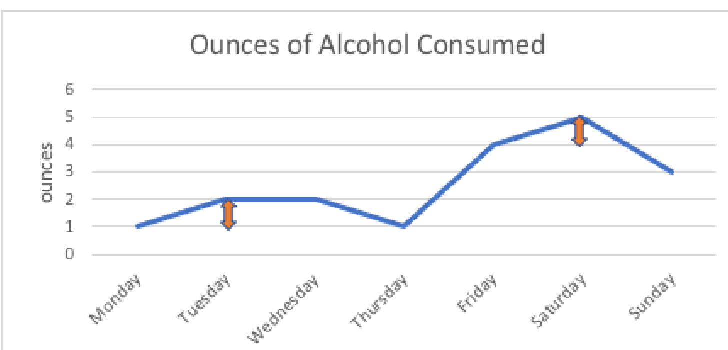
Rating scales. Consider a different but commonly used type of ordinal variable: rating scales. Social, behavioral, and social work investigators often ask study participants to apply a rating scale to describe their knowledge, attitudes, beliefs, opinions, skills, or behavior. Because the categories on such a scale are sequenced (most to least or least to most), we call these ordinal variables.

Examples include having participants rate:

- how much they agree or disagree with certain statements (not at all to extremely much);
- how often they engage in certain behaviors (never to always);
- how often they engage in certain behaviors (hourly, daily, weekly, monthly, annually, or less often);
- the quality of someone's performance (poor to excellent);
- how satisfied they were with their treatment (very dissatisfied to very satisfied)
- their level of confidence (very low to very high).

Interval Variables. Still other variables take on values that vary in a meaningful numerical fashion. From our list of demographic variables, age is a common example. The numeric value assigned to an individual person indicates the number of years since a person was born (in the case of infants, the numeric value may indicate days, weeks, or months since birth). Here the possible values for the variable are ordered, like the ordinal variables, but a big difference is introduced: the nature of the intervals between possible values. With ***interval variables*** the "distance" between adjacent possible values are equal. Some statistical software packages and textbooks use the term ***scale variable***: this is exactly the same thing as what we call an interval variable.

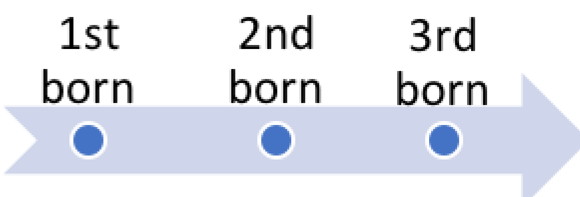
For example, in the graph below, the 1 ounce difference between this person consuming 1 ounce or 2 ounces of alcohol (Monday, Tuesday) is exactly the same as the 1 ounce difference between consuming 4 ounces or 5 ounces (Friday, Saturday). If we were to diagram the possible points on the scale, they would all be equidistant; the interval between any two points is measured in standard units (ounces, in this example).



With ordinal variables, such as a rating scales, no one can say for certain that the "distance" between the response options of "never" and "sometimes" is the same as the "distance" between "sometimes" and "often," even if we used numbers to sequence these response options. Thus, the rating scale remains ordinal, not interval.

What might become a tad confusing is that certain statistical software programs, like SPSS, refer to an interval variable as a "scale" variable. Many variables used in social work research are both ordered and have equal distances between points. Consider for example, the variable of birth order. This variable is interval because:

- the possible values are ordered (e.g., the third-born child came after the first- and second-born and before the fourth-born), and
- the "distances" or intervals are measured in equivalent one-person units.



Continuous variables. There exists a special type of numeric interval variable that we call ***continuous variables.*** A variable like age might be treated as a continuous variable. Age is ordinal in nature, since higher numbers mean something in relation to smaller numbers. Age also meets our criteria for being an interval variable if we measure it in years (or months or weeks or days) because it is ordinal and there is the same "distance" between being 15 and 30 years old as there is between being 40 and 55 years old (15 calendar years). What makes this a continuous variable is that there are also possible, meaningful "fraction" points between any two intervals. For example, a person can be 20½ (20.5) or 20¼ (20.25) or 20¾ (20.75) years old; we are not limited to just the whole numbers for age. By contrast, when we looked at birth order, we cannot have a meaningful fraction of a person between two positions on the scale.

The Special Case of Income. One of the most abused variables in social science and social work research is the variable related to income. Consider an example about household income (regardless of how many people are in the household). This variable could be categorical (nominal), ordinal, or interval (scale) depending on how it is handled.

Categorical Example: Depending on the nature of the research questions, an investigator might simply choose to use the dichotomous categories of "sufficiently resourced" and "insufficiently resourced" for classifying households, based on some standard calculation method. These might be called "poor" and "not poor" if a poverty line threshold is used to categorize households. These distinct income variable categories are not meaningfully sequenced in a numerical fashion, so it is a categorical variable.

Ordinal Example: Categories for classifying households might be ordered from low to high. For example, these categories for annual income are common in market research:

- Less than $25,000.
- $25,000 to $34,999.

- $35,000 to $49,999.
- $50,000 to $74,999.
- $75,000 to $99,999.
- $100,000 to $149,999.
- $150,000 to $199,999.
- $200,000 or more.

Notice that the categories are not equally sized—the "distance" between pairs of categories are not always the same. They start out in about $10,000 increments, move to $25,000 increments, and end up in about $50,000 increments.

Interval Example. If an investigator asked study participants to report an actual dollar amount for household income, we would see an interval variable. The possible values are ordered and the interval between any possible adjacent units is $1 (as long as dollar fractions or cents are not used). Thus, an income of $10,452 is the same distance on a continuum from $9,452 and $11,452—$1,000 either way.

The Special Case of Age. Like income, "age" can mean different things in different studies. Age is usually an indicator of "time since birth." We can calculate a person's age by subtracting a date of birth variable from the date of measurement (today's date minus date of birth). For adults, ages are typically measured in years where adjacent possible values are distanced in 1-year units: 18, 19, 20, 21, 22, and so forth. Thus, the age variable could be a continuous type of interval variable.

However, an investigator might wish to collapse age data into ordered categories or age groups. These still would be ordinal, but might no longer be interval if the increments between possible values are not equivalent units. For example, if we are more interested in age representing specific human development periods, the age intervals might not be equal in span between age criteria. Possibly they might be:

- Infancy (birth to 18 months)
- Toddlerhood (18 months to 2 ½ years)
- Preschool (2 ½ to 5 years)
- School age (6 to 11 years)
- Adolescence (12 to 17 years)
- Emerging Adulthood (18 to 25 years)
- Adulthood (26 to 45 years)
- Middle Adulthood (46 to 60 years)
- Young-Old Adulthood (60 to 74 years)
- Middle-Old Adulthood (75 to 84 years)
- Old-Old Adulthood (85 or more years)

Age might even be treated as a strictly categorical (non-ordinal) variable. For example, if the variable of interest is whether someone is of legal drinking age (21 years or older), or not. We have two categories—meets or does not meet legal drinking age criteria in the United States—and either one could be coded with a "1" and the other as either a "0" or "2" with no difference in meaning.

What is the "right" answer as to how to measure age (or income)? The answer is "it depends." What it depends on is the nature of the research question: which conceptualization of age (or income) is most relevant for the study being designed.

Alphanumeric Variables. Finally, there are data which do not fit into any of these classifications. Sometimes the information we know is in the form of an address or telephone number, a first or last name, zipcode, or other phrases. These kinds of information are sometimes called ***alphanumeric variables***. Consider the variable "address" for example: a person's address might be made up of numeric characters (the house number) and letter characters (spelling out the street, city, and state names), such as 1600 Pennsylvania Ave. NW, Washington, DC, 20500.

Actually, we have several variables present in this address example:

- the street address: 1600 Pennsylvania Ave.
- the city (and "state"): Washington, DC
- the zipcode: 20500.

This type of information does not represent specific quantitative categories or values with systematic meaning in the data. These are also sometimes called "string" variables in certain software packages because they are made up of a string of symbols. To be useful for an investigator, such a variable would have to be converted or recoded into meaningful values.

A Note about Unit of Analysis

An important thing to keep in mind in thinking about variables is that data may be collected at many different levels of observation. The elements studied might be individual cells, organ systems, or persons. Or, the level of observation might be pairs of individuals, such as couples, brothers and sisters, or parent-child dyads. In this case, the investigator may collect information about the pair from each individual, but is looking at each pair's data. Thus, we would say that the ***unit of analysis*** is the pair or dyad, not each individual person. The unit of analysis could be a larger group, too: for example, data could be collected from each of the students in entire classrooms where the unit of analysis is classrooms in a school or school system. Or, the unit of analysis might be at the level of neighborhoods, programs, organizations, counties, states, or even nations. For example, many of the variables used as indicators of food security at the level of communities, such as affordability and accessibility, are based on data collected from individual households (Kaiser, 2017). The unit of analysis in studies using these indicators would be the communities being compared. This distinction has important measurement and data analysis implications.

A Reminder about Variables versus Variable Levels

A study might be described in terms of the number of variable categories, or levels, that are being compared. For example, you might see a study described as a 2 X 2 design—pronounced as a two by two design. This means that there are 2 possible categories for the first variable and 2 possible categories for the other variable—they are both dichotomous variables. A study comparing 2 categories of the variable "alcohol use disorder" (categories for meets criteria, yes or no) with 2 categories of the variable "illicit substance use disorder" (categories for meets criteria, yes or no) would have 4 possible outcomes (mathematically, 2 x 2=4) and might be diagrammed like this (data based on proportions from the 2016 NSDUH survey, presented in SAMHSA, 2017):

	Illicit Substance Use Disorder (SUD)		
		No	Yes
Alcohol Use Disorder (AUD)	No	500	10
	Yes	26	4

Reading the 4 cells in this 2 X 2 table tells us that in this (hypothetical) survey of 540 individuals, 500 did not meet criteria for either an alcohol or illicit substance use disorder (No, No); 26 met criteria for an alcohol use disorder only (Yes, No); 10 met criteria for an illicit substance use disorder only (No, Yes), and 4 met criteria for both an alcohol and illicit substance use disorder (Yes, Yes). In addition, with a little math applied, we can see that a total of 30 had an alcohol use disorder (26 + 4) and 14 had an illicit substance use disorder (10 + 4). And, we can see that 40 had some sort of substance use disorder (26 + 10 + 4).

Back to Basics

Mathematical symbols can save a lot of words. Remember these?

Greater than: >
  Example: $5 > 4$
Less than: <
  Example: $4 < 5$
Greater than or equal to: ≥
  Example: age ≥ 45
Less than or equal to: ≤
  Example: IQ score ≤ 85

To make this distinction between variables and variable levels or categories crystal clear, let's consider one more example: a 2 X 3 study design. First, doing the math, we should see 6 possible outcomes (cells). Second, we know that the first variable (age group) has 2 categories (under age 30, 30 or older) and the other variable (employment status) has 3 categories (fully employed, partially employed, unemployed). This time the 6 cells of our design are empty because we are waiting for the data.

		Employment Status		
		Fully Employed	Partially Employed	Unemployed
Age Group	<30			
	≥30			

Thus, when you see a study design description that looks like two numbers being multiplied, that is essentially telling you how many categories or levels of each variable there are and leads you to understand how many cells or possible outcomes exist. A 3 X 3 design has 9 cells, a 3 X 4 design has 12 cells, and so forth. This issue becomes important once again when we discuss sample size in Chapter 6.

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.3-4.1 Beginning Data Entry

Chapter Summary

In summary, investigators design many of their quantitative studies to test hypotheses about the relationships between variables. Understanding the nature of the variables involved helps in understanding and evaluating the research conducted. Understanding the distinctions between different types of variables, as well as between variables and categories, has important implications for study design, measurement, and samples. Among other topics, the next chapter explores the intersection between the nature of variables studied in quantitative research and how investigators set about measuring those variables.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=222

Module 3 Chapter 5: Overview of Methods for Data Collection and Measurement

Remember that one defining characteristic of scientific inquiry is that it involves observation. This chapter is concerned with the many different approaches to observation that are commonly utilized in social, behavioral, and social work research. Each has its advantages and disadvantages, and investigators are faced with significant challenges in planning the measurement/observation aspects of their studies. In this chapter you will read about:

- validity and reliability principles related to quantitative measurement
- qualitative and quantitative measurement and data collection approaches

Measurement Validity & Reliability in Quantitative Research

The point of quantitative measurement is to objectively measure characteristics, features, or variables of interest. While this may seem obvious, it is important not to confuse variables and measures—measurement tools provide information about variables, they are not variables themselves. This measurement topic is focused on quantitative studies because it is not relevant to the aims of qualitative research. Two measurement traits are important in achieving the quantitative measurement task: ***measurement validity*** and ***measurement reliability.*** While neither trait can be realized 100%, the aim is to maximize both to the greatest extent possible. Let's look more closely at each of these traits and what they mean in quantitative research.

Measurement Validity. Measurement validity is about the extent to which a measurement tool or instrument adequately measures the concept, construct, or variable of interest (Perron & Gillespie, 2015). In other words, does it really measure what it purports to be measuring (Grinnell & Unrau, 2014). It may help to note that this use of the term "validity" is getting at a very different point than did the terms internal and external validity that your read about earlier in terms of experimental study designs. There we were concerned with either how confident we were in the results of the study (internal validity) or the conclusions of the study being generalizable (external validity). Here, we are concerned with the measurement procedures employed in generating those results. Measurement validity plays a role in both internal validity and external validity of a study, but operates at a more specific level of the research process.

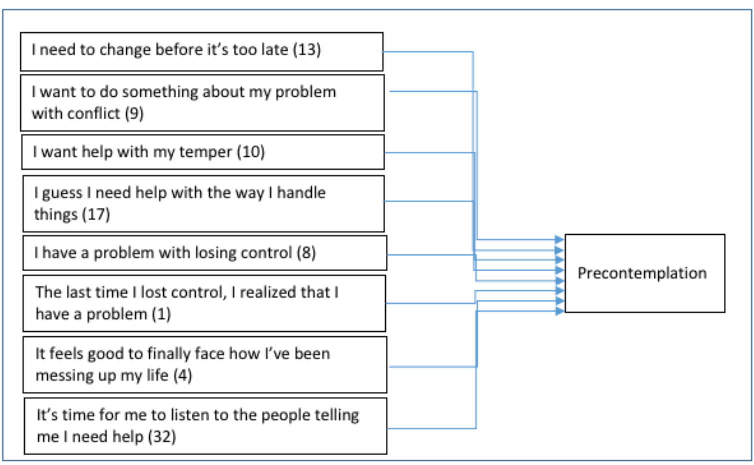
Our general definition of measurement validity includes aspects or components that you might read about: construct, content, and face validity, for example. While these are not all the same thing, they all contribute to a general assessment of measurement validity. Validity is sometimes assessed by convergent validity analysis, as well. In convergent validity, measurement scores or values are compared with those obtained on other measures of the same construct or variable, measured at the same time. For example, the URICA (University of Rhode Island Change Assessment) is a 32-item tool used by practitioners to assess clients' motivational readiness to change (McConnaughy, Prochaska, & Velicer, 1983). If investigators or clinicians desired a shorter, briefer version or a version for self-administered use on mobile devices, they would first need to demonstrate that the new version or format led to the same results as the original form.

Additionally, it is important that investigators recognize limitations of instruments and measures for work with different target populations. For example, the need to know how validly an instrument or measure translates into other languages. The problem with direct translation into another language is that nuances relevant to different cultures may play a role in how the measure is experienced. For example, translating an instrument developed in English into an instrument for Spanish-speaking persons is not simply a matter of using an English-Spanish dictionary and substituting words. There are many differences between the Spanish spoken and used in daily life among persons from Spain, Mexico, Chile, Cuba, and other areas, and there are many colloquialisms and slang terms that do not translate word-for-word. For example, the word "dope" has many meanings even in English: someone who does something stupid, marijuana or heroin (depending on region and cohort), insider information, or something that is awesome. A recommended practice is to have a native speaker of the language conduct a translation, then have another native speaker reverse translate it back into English and compare the original to the double-translated version to see where meaning was altered or lost.

Furthermore, it is important to be aware that language interpretation is not the same thing as cultural interpretation. In other words, just because the words are literally translated correctly does not mean that they have the same cultural meaning across cultures. Careful attention must be directed toward creating or selecting culturally relevant or culturally grounded approaches to measurement, instruments and measures that accurately capture the lived experiences of the studied population.

Latent variables. One important difference between the physical and natural sciences compared to social, behavioral, and social work science is related to whether the phenomena of interest are directly or only indirectly observable. Much of the science directed towards understanding diverse populations, social work problems, and social phenomena involves studying internal mental processes, attitudes, beliefs, and interpretations.A variable that is not directly observed but is mathematically or statistically inferred from other variables that are directly measuredis called a ***latent variable***. For example, in creating the Safe At Home instrument for measuring readiness to change battering behavior (Begun et al., 2003; 2008), investigators inferred the variables for Precontemplation, Contemplation, and Preparation/Action from the study participants' ratings of their agreement with 35 readiness statements. They created these 3 latent variables through a statistical approach called exploratory factor analysis (EFA)—see Figure 5-1. Validity of these sorts of inferred variables is an important issue to consider, and evaluating a research study may involve evaluating the evolution of inferred, latent variables that may have been used.

Figure 5-1. Items loaded in Contemplation (latent variable) through EFA (constructed from results presented in Begun et al., 2003)



Proxy variables. A commonly encountered validity issue arises with the use of ***proxy variables***. The word proxy means something (or someone) is representing something (or someone) else. A proxy voter is authorized to vote on behalf of someone else. A proxy variable is a variable that stands for or represents something else. A proxy variable is not directly relevant in itself, but serves as a close approximation of the variable of interest. Proxy variables are used when the variable of interest is difficult (or impossible) to measure; they are easily measured variables that can be used in place of more difficult-to-measure variables.

For example, birth order can be either a variable of interest used on social research, or it can act as a proxy variable for something more difficult to measure, such as parental attention or parenting experience. When we see a correlation between birth order and a trait (like intelligence or personality), we do not necessarily understand that relationship. Could it be the result of a biological phenomenon related to health of the prenatal, uterine environment which may change with each subsequent birth? Or, is it a function of how much time, undivided attention, and social/intellectual attention each subsequent child receives from parents? This is the problem with proxy variables—we do not know how validly they fit the actual construct for which they are standing in. Factors other than actual birth order number (1st, 2nd, 3rd and so forth) are associated with how much parental time, attention, and stimulation is received, making birth order a less than ideal proxy variable for the hypothesized mechanism of effect on child intelligence. Investigators generally do better to find direct measures of the constructs in which they are interested than to rely on proxy variables. Individuals reviewing a research report need to critically analyze the degree to which they believe a proxy variable is a good representative for the variable of interest. By the way, you may be interested to learn the results of a large-scale study about birth order (Rohrer, Egloff, & Schmukle, 2015): the authors concluded that "birth order does not have a lasting effect on broad personality traits outside of the intellectual domain" (p. 14224).

Also consider the problem of using age or income as a proxy variable for something else. For example, age is a relatively unreliable proxy for physical condition or health. Consider research that indicates 20% or more of U.S. children and adolescents have abnormally high cholesterol levels, a condition associated with poor cardiovascular health among adults (Nguyen, Kit, & Carroll, 2015). Or, consider older adults whose condition is similar to average adults in much younger age groups—marathons have age groupings of 70 and over or 80 and over—while many younger adults could not meet the physical challenges involved. Age is not a great proxy for fitness. Similarly, income might not be a good proxy for families consuming nutritious food. Some persons reside in “food deserts” where nutritious food is difficult to obtain within a reasonable geographic distance regardless of their available economic resources. Income is not entirely predictive of this experience since some families with low incomes are able to grow much of their own healthful food, while others cannot, and some families with higher incomes spend their resources on poor quality “fast” or “junk” foods instead of healthful foods.

Proxy variables are risky to apply to populations where there exists a great deal of ***heterogeneity***. Unfortunately, investigators are sometimes forced to rely on proxy variables when measuring the true variable of interest is not possible. This problem is often encountered when working with administrative or secondary data.

Measurement Reliability. As a complement to measurement validity, measurement reliability indicates the degree of accuracy or precision in measuring the construct or variable of interest. This accuracy indicator is concerned with stability in measurement, for one thing: if you use the tool or instrument to measure the same thing multiple times, you would hope to get the same result. Imagine trying to measure your cereal box with a rubber band—you want to know how many stretched rubber bands big it is. If you measured it 5 times in a row, the chances are that you would not get the same answer all 5 times because the rubber band is such a flexible tool and you cannot be sure you applied the same amount of pressure each time .

Reliability problems arise in real-world research, as well. Imagine, for example, that your goal was to measure aggressive behaviors exhibited by a toddler. It might be easy to count kicking, biting, shoving, and hitting behaviors. But how would you count instances where the toddler yells or spits at someone, raises a hand to strike but doesn't follow through, or kicks the person's chair rather than the person? Would everyone observing that behavior reliably count it the same way? This gets at the issue of precision. An important dimension of measurement reliability is called ***inter-observer reliability***. This might also be called inter-rater or inter-coder reliability, depending on the nature of the measurement data. In studies where qualitative, open-ended questions, or observational data are being evaluated for analysis, strong research methodology involves applying methods to assess and ensure that different observers, raters, or coders would get the same result when evaluating the same data. For example, the team might spend time up front training a group of observers to rate or code a set of sample data. Once everyone has developed a strong consensus on the practice data (at least 90% agreement is one standard), pairs of observers might rate or code a percentage of the actual data, with their evaluations being compared. Again, the expectation is that their level of agreement is very high, and that it remains high over time (not just when the training was still fresh in their minds).

There exists a tremendous need to ensure reliability in measurement, particularly in social work research where variables might be difficult to directly observe and quantify. For example, social work practitioners and researchers might be concerned with reliably measuring anxiety. Asking different individuals to rate how anxious they feel right now on a scale of 0-10 might not generate very reliable answers. While zero might be clear to everyone (no anxiety at all), for a person who often experiences very high levels of anxiety, a "5" might mean something very different than for a person who seldom experiences anxiety or whose anxiety is mild when experienced.

The goal with measurement reliability is to minimize measurement error. Measurement error is a recognized problem in social work, social science, and behavioral science research. It concerns the difference between value obtained through measurement of a variable and the variable's true (unknown) value. Investigators do their best to be accurate in measurement, but a certain amount of "noise" is always a concern in the data collected. Measurement error leads investigators to either over- or under- estimate reality. Turning back to our anxiety measurement example, measurement error or bias may be introduced by the context in which anxiety is measured. Measuring immediately before or after a quiz might lead to different results than measuring first thing in the morning after a good night's sleep. Or, women being interviewed about experiences of sexual harassment might respond differently to male versus female interviewers. Or, data collected on Mondays might be either more accurate after participants have had a restful weekend or might be less accurate after their partying weekends, compared to data collected on other days of the week.

Cultural Competence. An issue integral to both measurement validity and reliability concerns ***cultural competence in measurement*** procedures. You may be familiar with the long-standing controversy concerning cultural bias in standardized testing, particularly intelligence tests. A demonstrative example comes from an intelligence test purportedly administered to Koko, the gorilla who learned to communicate with humans using American Sign Language. When asked where you should go if it starts to rain, Koko "failed" the item by responding "up a tree" instead of "inside a house." The test is clearly "human-centric" and biased against gorillas. Similarly, there are many examples of ethnic, cultural, geographic, gender, and other bias built into testing materials. For example, the T-ACE (Tolerance, Annoyance, Cut Down, and Eye Opener) and TWEAK (Tolerance, Worry, Eyeopener, Amnesia, Kut Down) are two alcohol screening tests developed for use with women when it was determined that existing screening measures were more appropriate for use only with men (Russell, 1994).

Cultural equivalence/nonequivalence and avoiding ethnocentrism are important consideration when evaluating or planning measurement in quantitative research. The issue is that concepts or constructs may have very different meaning to persons living in different cultural contexts. Consider, for example, the meaning of marriage to individuals in same- versus different-sex relationships, particularly in light of the unstable history of marriage policy across the nation and the globe. Also consider that equivalents for a phenomenon that exists in one cultural group may not exist in another. An example might be the sibling-like cousin relationships that exist among cultural groups characterized by high degrees of cooperation and community compared to the nuclear family conception of siblings.

One type of research question concerns measurement science: validating research measurement and clinical assessment tools for use with diverse populations. This may begin with the initial development of a measurement approach/instrument, or involve recommended modifications of existing tools based on evidence. An important "how to" resource is a book called Developing cross-cultural measures in social work research and evaluation, 2ndedition (Tran, Nguyen, & Chan, 2016).

A concern beyond the actual measurement tools is the measurement procedures involved in a study. Of great concern is the way that the measurement process might put study participants in jeopardy. One example is measurement that asks about illegal behaviors, such as underage drinking, illicit use of drugs at any age, or substance use during pregnancy. Asking for honest, truthful answers to these kinds of questions can place participants in jeopardy with legal or child welfare systems if the confidentiality is breached. One tool for managing this risk is the Certificate of Confidentiality that an investigator might request from the National Institutes of Health (NIH) in connection with securing local IRB approval to conduct this type of research. The Certificate of Confidentiality is designed to protect investigators from being forced by court order to disclose individually identifiable data.

Another scenario to consider is discussed in an article about safety in research involving unauthorized immigrant women experiencing intimate partner violence:

> *"Social science research with vulnerable populations is necessary in order to address social and health problems among those most in need. Yet, this research is fraught with inherent risks to its participants" (Kyriakakis, Waller, Kagotho, & Edmond, 2015, p. 260).*

The authors expressed concern over the potential physical, emotional, psychological, economic, and legal harms to which these women could be exposed by virtue of the research process. The risk to the 29 Mexican immigrant women was related to their recent history of exposure to intimate partner violence, and for 26 of the women the risk was compounded by their unauthorized immigrant status. The authors described various strategies used to minimize the women's risks of participation, including cautious recruitment, screening, informed consent, interviewing, data management/storage, resource referral, and data reporting procedures.

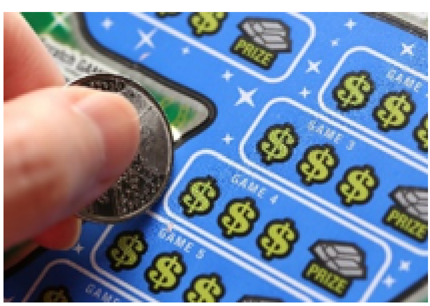
Ways of Measuring and Collecting Data

Whether a scientific study adopts a qualitative or quantitative approach, investigators need to determine a strategy for collecting data about the phenomenon being observed. Options vary along several dimensions, one of which is the degree to which strategies are obtrusive. Think about your family's home videos and the pictures you have on your phone. There is a great deal of difference between the candid and posed images—the candid images capture more natural behavior than do the posed images.

© R. Veenstra

The same is true of scientific measurement: behavior observed under natural conditions is more representative than is behavior observed under laboratory conditions or conditions where those being observed are aware that their behavior is being studied. This is the first dimension we explore with regards to ways of measuring and collecting data.

Naturalistic Observation. Both qualitative and quantitative studies can utilize ***naturalistic observation*** techniques in data collection. This means that information about behavior is collected in normally experienced contexts. If an investigator wishes to learn about aggression between preschool-aged children, the behavior might be observed in naturally occurring situations and natural environments—daycare, playgrounds, at home, or riding in the car. The advantage of this approach is that the data are representative of natural behavior. The disadvantages of this approach include lack of investigator control over circumstances or variables that might interfere with the behavior being observed. For example, home observations might be interrupted by the doorbell or telephone ringing and the dog barking, which might alter the behavior of those being observed. Laboratory conditions can minimize these confounding factors, but observations of behavior expressed in the laboratory may not accurately reflect or generalize to natural behavior conditions. Another factor which investigators take into account is the possibility that the behavior of interest might not occur often enough to actually be observed under natural conditions. Investigators may feel the need to create laboratory conditions to "trigger" the behavior of interest so that it can be observed. For example, imagine that investigators are interested in observing how families negotiate the distribution of resources across members and for the family as a whole. This may be difficult to observe during a conveniently scheduled period of observation. The investigators might, instead, provide the family members with a specific distribution task and observe how they negotiate a solution—perhaps providing them with a scratch-off State lottery ticket and asking them to decide what to do if were a $1,000 winner. While this is not a typical, routine resource distribution problem that families encounter, the task is designed to "trigger" resource distribution behavior from the family members being observed. The limitation being that family members will need to advise the investigators as to how typical their behavior was compared to naturally occurring situations.



The type of data collected from naturalistic observation differs between qualitative and quantitative studies. Qualitative data might include rich, detailed descriptions of who does/says what, when they do/say it, how they do/say it, and the sequential context of what happens–the precursors and consequences of each behavior. These data may later be coded for themes or other qualitative analyses. Quantitative data might include numeric counts of the frequency of certain behaviors or the duration (in seconds or minutes) of specific behaviors of interest. To accomplish this, the investigators need to have predetermined what behaviors will be recorded as data, being very precise as to the definition and criteria for those behaviors. In both cases, qualitative and quantitative, the ability to review the action is helpful to ensure data accuracy and completeness. Thus, recording the observed events is common so that the events can repeatedly be replayed. Furthermore, working with recordings allows investigators to cross-check data decisions made by individuals coding, scoring, or rating the observed behaviors.

Direct observational data is particularly useful in studies where participants are not able to provide information about themselves or their usual behavior. This is generally true of infants and young children, for example. As their language/communication skills develop and mature, individuals' ability to answer questions about their internal thought processes and behavior improves, at least somewhat. Observation may be a strong option in studying persons who have cognitive or intellectual barriers to consistently providing information—persons with dementia, for example.

Photovoice. An approach to qualitative data collection that differs from observing individuals in their natural environments involves having individuals themselves actively engage in observing their natural environments. ***Photovoice*** is a powerful participatory action strategy for data collection (Wang, 1999). The approach begins with having individuals take photographs of what is meaningful to them in their daily lives. It is particularly important as a method used with traditionally marginalized groups or communities (Fisher-Borne & Brown, 2018). The second phase of photovoice methodology is equally, if not more, important—recording the participants' explanations of the images they chose to capture or create. The use of photovoice methodology was presented as a strategy for engaging young, Black men as researchers in community-based participatory research (CBPR) with the goal of shaping an HIV prevention and care project, ensuring that they are empowered to engage and have their voices heard (Fisher-Borne & Brown, 2018). In this study, participants documented engaged in action-oriented, problem-solving group discussion sessions about their photos. This feature distinguished the study from documentary photovoice projects—the emphasis of the discussions was on helping the men identify issues and action steps toward personal and societal change. Each week of the project had a theme around which the men took pictures (e.g., stereotypes, Black masculinity, sexuality, self-esteem, emotions, multiple "identities"). The pictures were discussed in group sessions, with each photographer selecting their own discussion photo, providing a name or label for it, and leading discussion about it in relation to the theme. The analyses which included the study participants resulted in 4 themes: identity, stereotypes reinforce oppression, expectations, and breaking free/embracing self. The authors argued in favor of using this approach as a qualitative, action-oriented research method:

> "*Embedded within the Photovoice technique is the belief that people can participate and define for themselves issues that will shape their lives. While Photovoice is a vehicle for personal expression, its structure provides for the vocalization of community and public issues. As such, the visual images and*

participatory process serve as a tool for advocacy and collective action" (Fisher-Borne & Brown, 2018, p. 171).

Artifact Analysis. Another form of unobtrusive observation involves analyzing the remains or artifacts that remain after naturally occurring behavior. For example, alcohol researchers may wish to record data concerning the alcohol purchases made by study participants. Participants could be asked to save all receipts (paper and electronic) for all purchases made in a month. These receipts can then be content analyzed for purchases involving alcohol. This ***artifact analysis*** approach may be more representative of actual alcohol purchasing behavior than asking about the behavior—individuals may be unaware of or mistaken about their actual purchasing behavior, either under- or over-estimating reality. Other types of artifacts that might be analyzed include personal diaries and social media postings, as well as data recordings from devices such as fitness and GPS trackers. As in the case of Photovoice methodology, other arts-based approaches might be appropriate: eliciting and analyzing dance, poetry, drawings/painting, sculpture, and other artifacts.

Administrative and Secondary Data Sources. Client records and

administrative data are forms of artifacts that might usefully be analyzed to answer some research questions, as well. The information may not have originally been collected for research purposes, but may be informative in research all the same. Administrative data includes information routinely recorded in programs, agencies, and institutions as part of normal operations. For example, investigators analyzed data routinely collected at the state level as child abuse and neglect (CA/N) reports (Lanier at al., 2014). The investigators calculated disproportionality ratios for each state by comparing the state's child maltreatment rate for black and white groups, and for Hispanic and white groups. They then related these state-by-state data to statistics about each state's poverty, teenaged mother, and single mother rates. The results of this study using administrative data indicated that disproportionality in child maltreatment rates was correlated with disproportionality in poverty rates, and this was observed for both black and Hispanic groups. Furthermore, a significant amount of the child neglect disproportionality between black and white groups (66%) was accounted for by disproportionality in teen motherhood rates. This kind of study and the observed disproportionate poverty-maltreatment links have profound social justice implications for social work practice and policy.

Administrative data differs from what you may have heard called ***secondary data*** analysis. Secondary analysis concerns information collected for a research purpose, but analyzed differently than in the primary study for which the information was originally collected. The original study involved ***primary data*** collection, the later study involves secondary data analysis. For example, numerous investigators have secondarily analyzed data from the original Fragile Families and Child Wellbeing study. The original is an ongoing, multi-year, longitudinal study of child development outcomes in relation to numerous parent, family, and neighborhood factors. Barnhart and

Maguire-Jack (2016) reported on their analyses concerning the role of parenting stress and depression in the relationship between single mothers' perception of community social cohesion and child maltreatment. They analyzed data from a subset of 1,158 single mothers from the full original study. The investigators reported that single mothers' perception of community social cohesion has an effect on child maltreatment, and the nature of this effect is influenced by maternal depression: greater social cohesion was associated with lower depression rates, and lower depression was associated with lower child maltreatment rates. These findings demonstrate the importance of social workers attending to both individual and community level factors when developing interventions to prevent or reduce child maltreatment; intervening around community social cohesion might make a difference in child maltreatment because it can influence maternal depression. This is also important because social workers not only care about child well-being, we care about the well-being of mothers, as well.

The two examples (Lanier, et al., 2014 and Barnhart & Maguire-Jack, 2016) demonstrate the greatest advantage associated with administrative and secondary data. The investigators addressed their research questions without having to expend tremendous time and resources in generating new primary data. They avoided duplicating the data gathering efforts that have already been expended. That does not mean that administrative and secondary data analysis studies are easier or less effort-intensive than primary data studies: the effort and time is distributed differently. These studies still need to undergo IRB review for the protection of human subjects to ensure that the risks are minimal or less. These two studies also demonstrate one of the most significant limitations involved with administrative and secondary data. The primary data collection may not have included the best, direct measures of the variables in which the investigators were interested. There may exist a need for faith in the leap of logic that extends between the available data and the variables that the investigators wish to analyze.

Furthermore, using administrative or secondary data leaves the new study vulnerable to any limitations in methodology associated with the primary study or the administrative data. For example, in a study of prisoner visitation patterns by family members, investigators were able to access data concerning every visit made to every state prisoner in Ohio over the preceding five years (Begun, Hodge, & Early, 2017). The study was limited, however, by imprecise and high variability in how family relationships were reported in the administrative data: mother, mother-in-law, step-mother, and parent were all used differently by different recorders and at different institutions. Similarly, someone might be described as a sibling, brother, brother-in-law, step-brother, half-brother, step-sibling, or half-sibling. Thus, the investigators were limited in the degree of detail they could apply to the variable of family relationship—the possible categories had to be combined into more general, less specific categories for the sake of accuracy.

Geographic Information Systems (GIS). You were introduced to ***geographic information systems* (GIS)** research in the earlier discussion about correlational studies, and the example about marijuana dispensary density in relation to child maltreatment rates. This is considered a non-intrusive form of data collection, since it is not directly experienced by the individuals whose environments are mapped. The approach can be combined with other, more intrusive approaches, to generate answers to important questions about diverse populations, social work problems, and social phenomena. For example, investigators examined the nature of parents' engagement with "activity spaces" using both qualitative interview data and GIS methods (Price Wolf, Freisthler, Kepple, & Chavez, 2017). The aim of the study was to better understand how physical and social environments shape parenting behaviors, which in turn, influence child development and health outcomes. They found that parenting had a powerful influence on the places where parents go (child-centric places with their children, "adults-only" places to escape from parenting demands). In addition, day of the week, children's ages, role as primary or other caregiver, and income all influenced parents' use of activity spaces by imposing boundaries that limited options. This information provides social work with insights concerning where community-based interventions for parents might best be situated to be accessible for these families, and when different spaces might best be targeted.

Key Informants. One method of collecting data is to have knowledgeable others provide descriptions rather than collecting information directly from individuals. Sometimes this approach is combined with data collection from study participants, as well. ***Key informants*** are helpful when studying individuals who cannot effectively share reliable information about themselves—young children, for example, or persons experiencing significant short- or long-term cognitive/intellectual impairments. Child development studies often involve assessment tools completed by parents and teachers. This approach is reasonable when the studied behaviors are observable, but not with studies of internal mental or emotional processes. Great care must also be taken to protect the rights of the individuals being studied.

Surveys. A great deal of social research is conducted using survey tools for data collection. It is important to distinguish between surveys as tools for data collection and ***survey methodology***. Survey methodology refers to how survey study samples are drawn from the population (more about this in Chapter 6) and the use of survey instruments. Surveys are administered through a variety of mechanism: hard copy, paper and pencil surveys; telephone or in-person interview surveys; and, on-line, e-mail or web-based surveys. Most likely, you have participated in surveys from product marketing, course or instructor evaluations, public opinion research, and others.

A great deal has been written about constructing "good" survey instruments and questions. First, it helps to consider the type of information to be provided for a survey question: objective, factual data or information about a person's subjective opinions, beliefs, attitudes, perceptions, ideas, or states (Fowler & Cosenza, 2009). Asking effective factual questions is not always as simple and straightforward as it may seem. Imagine a study where investigators want to know about a person's alcohol consumption. They plan to ask study participants:

> *How often and how much do you drink?*

You have probably already identified one problem with this question: drink what? It is very important to clearly define terms in survey questions. The revised question might be:

> *How often and how much do you drink beverages containing alcohol?*

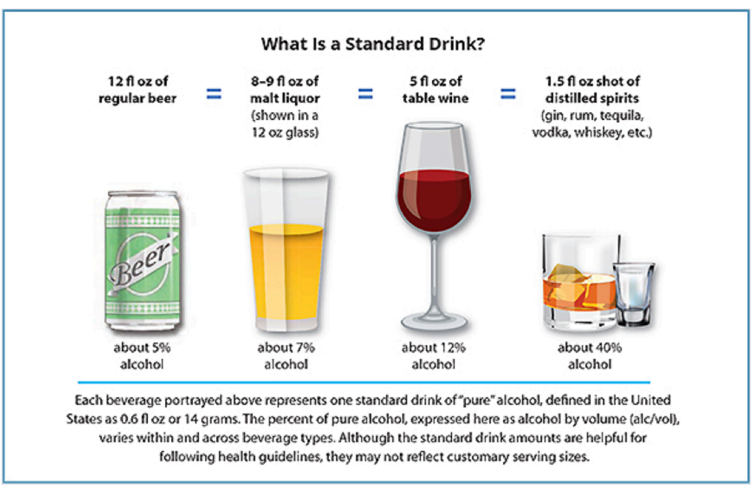
While this question is better defined than the first draft, it still suffers from another common problem: it is a ***double-barreled question***. A double-barreled question is presented as a single question, but actually asks two different questions needing two distinct answers. In this case, one question is about frequency of alcohol consumption, the other is about amount. Recognizing this, the investigators might change their survey to ask:

> *How often do you drink beverages containing alcohol?*

When you drink beverages containing alcohol, how much do you drink?

There remains a significant problem with the second of these two questions—a lack of precision about how amounts might be reported. One person might respond "a lot," another might respond "it depends." The investigators need to make clear that they are asking for a specific, concrete, objective number here. Still, if a person responds "3" in reference to the number of drinks, the investigators still do not know how much alcohol was consumed because different types of alcoholic beverages contain different amounts of alcohol, and it is unclear how large the 3 drinks might be. Figure 4-1 shows how questions about drinking amounts could be asked.

Figure 4-1. Using the chart* below, on a day that you drink alcohol, how many standard drinks do you consume?



Retrieved from https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/what-standard-drink

Asking questions about a person's subjective internal states can also be complicated. As we saw with asking factual questions, defining terms and avoiding double-barreled question both remain important. It is important that the wording of a question not be biased at the outset. Bias can be introduced unintentionally by using specific labels—consider the difference between the labels "victim" and "survivor" of intimate partner violence, for example. Another potential source of bias comes from the phenomenon recognized as ***social desirability*** responding. Sometimes the answers people provide reflect what they believe is the most favorable or "correct" response rather than a response that truly reflects their own reality—aiming to be "PC" ("politically correct" rather than accurate. The literature contains several measures of social desirability, questions that can be integrated into a survey, to determine the extent to which an individual's answers might be suffering from social desirability bias.

In addition, the options provided in the survey for responding to questions need to be carefully considered. Response options include (but are not limited to):

- *Open ended or fill-in-the-blank*. This option provides the greatest degree of latitude and potential variability in responding. Depending on the nature of the questions asked, the responses can be qualitatively (rather than quantitatively) analyzed. If the goal of the study is quantitative, these types of response are more difficult to work with. Problems include difficulty with interpreting handwriting and spelling errors, or non-standard uses of words or terms. Responses may also vary greatly in terms of how much information and detail different participants provide—some may give a word or two while others provide great amounts of detail. This response variability can be difficult to analyze.
- *Yes/no*. This option offers only two possibilities, meaning that it

is “forced choice” question. This is sometimes used as a screening question to determine the appropriateness of asking further questions on the topic. For example, our investigators asking about alcohol consumption amounts might first ask: Have you consumed any beverages containing alcohol in the past 12 months? Then, they would as the frequency and amount questions only of individuals who responded “yes” to the screening questions; those responding “no” would skip the frequency and amount questions. One problem with force choice questions is that study participants may become frustrated, especially when the answer is not quite so cut and dried. For example, a screening question that asks, “Do you drink beverages containing alcohol?” would be difficult to answer with either “yes” or “no” by a woman who drank until she became pregnant and now does not, but will again after the baby is born.

- *Rating scales*:Survey questions often ask participants to apply a continuum or rating scale in answering a subjective question. This might be in the form of how much they agree with specific statements, their evaluation of something (degree of goodness, fairness, attractiveness, or other qualities), or estimations about how often something occurs. Two general types of scales exist. One is a single, unidimensional scale, such as rating happiness from low to high. The other is a two-dimensional scale, such as rating affect from unhappy to happy with a neutral value in the middle. It is important when deciding which type of scale to apply that both ends of the two-dimensional scale cannot occur at the same time for the same person. For example, a person might be unhappy about some things and happy about others at the same time. In this case, it would be more informative to have two separate unidimensional scales (one for happiness and one for unhappiness) instead of relying on a single two-dimensional scale, since feeling neutral is very different from feeling both unhappy and happy.

It might be wise, to offer participants options outside of the main choices. Response options like “unsure,” “don’t know,” or “prefer not to answer” might keep someone from either skipping the question or providing an inaccurate response. On the other hand, it may be that a forced choice is desired by the investigators and an alternative, unscaled option is not offered.

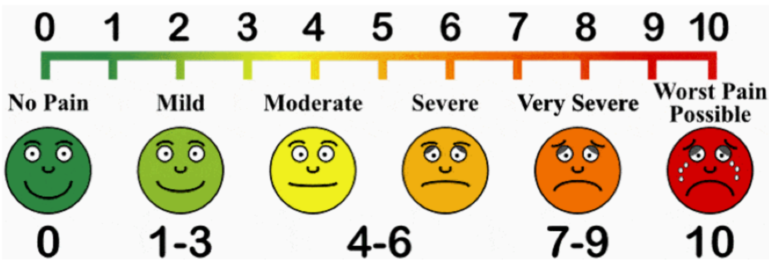
Another consideration with rating scales is how many points on the scale are offered. Too few options (e.g., a 3-point scale being 0, 1, 2) may lead to there being very little variation in the responses provided—it would be too easy for everyone to pick the same value, perhaps the middle of the range (ratings of 1). On the other hand, having too many options makes it difficult for people to decide on an appropriate rating—for example, when rating something on a 100-point scale, would there be a meaningful difference between ratings of 70 and 71? It would be easier for someone to use a 10-point scale (where the response might be 7). A common solution is the 5-point scale with two extremes, a neutral midpoint, and two modified options between the midpoint and extremes. For example:

1	2	3	4	5	9
Not At All	A Little	Somewhat	A Lot	Extremely Much	Unsure/ Don’t Know

Notice in this example that each possible numeric response has an associated description. In survey terms, these ratings are “anchored” by the words provided. Sometimes a scale is presented with some anchors being unspecified. For example:

1	2	3	4	5	6	7
Never						Always

This leaves the meaning of the ratings between 2 to 6 open to individual interpretation. Additionally, the anchors do not have to be words or phrases. For example, symbols or emoticons might be used, especially in working with children. The pain scale is an example that might be usefully modified for surveys (retrieved from www.disabled-world.com/health/pain/scale.php):



A note about Likert scales. A common error in describing survey scales revolves around use of the term "Likert scale." By definition, a Likert scale is a set of individual items that, in combination, form a single index (or scale). A Likert scale is made up of Likert items. Likert items generally have five or seven options on the rating continuum, but the midpoint is always neutral. Just because items have five rating options arrayed on a continuum does not make it a Likert item—a unidimensional item has no neutral midpoint, so it is not a Likert item. A Likert scale score is calculated by mathematically combining the ratings for individual Likert items, usually as either a total (sum) or as an average (mean) score. In some cases, the developer of a scale might provide alternative scoring instructions (such as multiplying some item ratings to provide weighted scores if some items are more important than others to the meaning of the final score). For example, ratings on the following four Likert items about readiness to succeed in a social work course about research and statistics are simply added together to create a composite Likert scale score for each study participant—one participant's hypothetical responses are scored in this example, receiving a score of 14 on a possible range of 4 to 20 (see Table 4-1).

Table 4-1. Likert items and Likert scale scoring example

How much do you agree or disagree with each statement below? 1=completely disagree 2=somewhat disagree 3=neither disagree nor agree 4= somewhat agree 5=completely agree	**Item Rating**
Research evidence is important for social work practice	4
Math is relatively easy for me	2
Intellectual challenges and puzzles are enjoyable to work on	3
I am confident that I can master the content in this course	5
Scale Score	**14**

Reverse scoring scale items. In order to compute a single, composite score by adding ratings together requires all of the items to run in the same direction with regard to the construct being measured. Consider what would happen if the second item (B) had been phrased differently: Math is relatively difficult for me. In this case, the item would be negative with respect to the construct of readiness to succeed in the course. Thus, the rating scale would need to be reversed before the rating values could be added together. In other words, ratings of 1 and 5 would be flipped so that ratings of 1 become 5, and ratings of 5 become 1; similarly, ratings of 2 and 4 would be flipped so that ratings of 2 become 4, and ratings of 4 become 2.

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.3-5.1 Summing Scores

Screening and Assessment Tools. Very often, investigators utilize existing ***screening and assessment tools*** for measuring the phenomena or constructs of interest in a research study. The literature is rich with measures designed to screen for or assess many different behaviors, conditions, or social work problems. It is usually preferable to use pre-existing measurement instruments rather than creating new ones for a particular study, particularly if the literature supports their reliability and validity for the job (more about this later in this chapter). Electronic search terms that help investigators and practitioners locate pre-existing measures on a specific topic include: tests, measures, measurements, assessment(s), scale(s), and screening.

Locating measures that can be applied in research studies is facilitated by various catalogues, compilations, guides, and handbooks that review them. For example, the *Mental Measurements Yearbook*(Carlson, Geisinger, & Jonson, 2017) is a frequently updated published resource with information to facilitate the selection of measurement tools by practitioners and investigators. An investigator interested in measures concerning attention deficit with hyperactivity disorder (ADHD) would find at least 4 options reviewed in the series (from 1985 to the present). A listing of all reviewed measures is presented on the publication website (https://buros.org/tests-reviewed-mental-measurements-yearbook-series). Many of the measures are commercially available, meaning that there is potentially a fee associated with their use—they may not be available in the public domain—and many require specialized training or certification.

Additional guides to research measures, including other than clinical behaviors, are:

- *Handbook of clinical rating scales and assessment in psychiatry and mental health*(Baer & Blais, 2010)
- *Handbook of family measurement techniques, v.* 1-3(Touliatos, Perlmutter, Straus, & Holden, 2001)
- *Handbook of psychiatric measures, 2nded.* (Rush, First, & Blacker, 2008)
- *Handbook of research design and social measurement*, 6*thed.*(Miller & Salkind, 2002)
- *Measures for clinical practice and research*, 5*thed*, *v.* 1*&*2(Corcoran & Fischer, 2013)
- *Positive psychological assessment*: A *handbook of models and measures*(Lopez & Snyder, 2003)
- *Rating scales in mental health*(Sajatovic & Ramirez, 2012)

Important cautions revolve around the use of screening and assessment results as research measures. First, investigators must develop procedures for providing test results to study participants, especially those whose results indicate a clinical problem or high risk of developing one. Second, the investigator needs to understand the measurement tool sufficiently well to determine whether actual scores can be used as a scale variable or if the test results can only be relied on for the yes/no dichotomous categories related to whether or not the person meets criteria for the condition being studied. For example, a measure of depression might help determine if a person is "over" or "under" the threshold for meeting depression criteria, but actual raw depression scores cannot be treated as a scale: the difference between a score of 38 and 40 is not the same as the difference between a score of 48 and 50. The AUDIT (Alcohol Use Disorder Identification Test) is an example of a screening instrument that has been used both ways (Babor, Higgins-Biddle, Saunders, & Monteiro, 2010). This screening measure is in the public domain and has 10 items, each of which is rated from 0 to 4 points based on the rubric presented for each item. An AUDIT score of 8 or higher for men, 7 or higher for women, is considered a "positive" screening result, indicating that an alcohol use disorder possibly exists. On the other hand, higher AUDIT scores are indicative of greater risk, and the actual scores can be used as a scale (ranging from 0 to 40). Or, the scores can be broken into 4 categories:

- Zone 1 scores 0-7 intervene with alcohol education alone,
- Zone 2 scores 8-15 intervene with simple advice about alcohol misuse,
- Zone 3 scores 16-19 intervene with simple advice plus brief counseling and continued monitoring,
- Zone 4 scores 20-40 intervene with referral to specialist for diagnostic evaluation and treatment (Babor, Higgins-Biddle, Saunders, & Monteiro, 2010, p. 22).

Interviews. Whether they are engaged in a qualitative, quantitative, or mixed-methods study, investigators may rely on interviews to gather data. ***Interview protocols*** vary in terms of the extent to which they are pre-scripted and structured. Qualitative studies in the grounded theory tradition typically are relatively unstructured since the purpose is to learn about the individual's own perspectives and interpretations about events, processes, or phenomena. The "flow" of such an interview follows from what the study participant shares rather than from a script predetermined by the investigator. At the other end of the continuum, quantitative study interviews typically are quite structured and scripted so as to elicit similar types of information (though varying in content) from each study participant. The semi-structured interview format is situated between the unstructured and highly structure interview formats. The questions in a semi-structured interview are general and open-ended and the "script" may include specific prompts to help guide the process. The research interview requires an investigator to employ many of the same skills used in social work clinical practice.

An important characteristic defining interview questions is whether they are open- or close-ended questions. Questions that can be answered with simple statements of fact or with yes/no responses are close-ended questions. An example might be:

> *How many brothers and sisters lived with you 50% of the time or more when you were growing up?*

The answer to this question is likely to be short and factual (objective). On the other hand, an open-ended question requires the person to elaborate on the information being shared. For example:

> *If you had to pick the one brother or sister to whom you feel emotionally closest, who would it be and why?*

A common error that interviewers make is failing to frame questions in open-ended format when they are seeking rich, descriptive responses. Even with open-ended questions, prompts to keep the story expanding may be necessary:

> *Can you tell me more about that?*

What else should I know about it?

I am not sure I understand. Can you explain it to me differently?

Interview data might be recorded in notes kept by the interviewer, audio recordings, video recordings, or notes kept by a third-party observer. Ideally, more than one strategy is utilized in order to ensure the most complete data—any of these approaches can fail or data from one approach might be unintelligible/unreadable later. It is helpful to have multiple forms for cross-checking and completion purposes. Furthermore, they complement each other. For example, tone of voice and meaningful pauses may be detected in an audio recording but not be evident in written notes or transcripts of the interview dialogue. Video recordings may pick up nonverbal cues that are missed in audio-only or written transcript records.

Focus Groups. A ***focus group*** is essentially a group interview. There are two main reasons for using focus group procedures rather than individual interviews. The first is that it might be more practical, feasible, or advisable to meet with small groups rather than individual study participants. For example, information might be gathered from 6 individuals in a 90-minute group session more quickly than in 6 separate 30- or 60- minute interviews. The second (and perhaps more important) reason is that the focus group participants may build off each other's comments in such a way that a more rich, in-depth picture emerges from the whole group than would emerge from assembling information provided by 6 separate individuals. A third reason is that individuals may feel more comfortable discussing certain topics in a group context than in individual interviews with an unfamiliar investigator. Of course, the opposite may also be true:

some may feel more comfortable discussing the topic individually rather than in a group.

Focus group methodology was used in a study of the importance of HIV/AIDS prevention education among older adults (Altschuler & Katz, 2015). The investigators hosted five 90-minute focus groups with adults aged 50 to 82 years. The study was prompted by the observation that increasing numbers of older adults are living with HIV/AIDS, prevention education efforts are seldom targeted/ delivered to older adults, and programs serving older adults seldom address HIV/AIDS prevention education. They wanted to learn from study participants what they recommend, what would motivate them to participate in HIV/AIDS prevention education or to seek such information. This information was shared in responding to 5 general questions in a semi-structured group interview context. Transcripts of the recorded focus group sessions were thematically coded by the investigators. Analyses led to the identification of 4 themes:

- Wanting to learn about the topic was fostered by knowing someone who has HIV/AIDS or a general sense of compassion for these individuals.
- Recommending that HIV/AIDS prevention education be

delivered in settings where older adults work or volunteer.
- Reminding others that despite stereotypes, older adults remain sexually active "both within and outside of monogamous relationships," and therefore remain at risk for exposure (Altschuler & Katz, 2015, p. 694).
- Ambivalence about discussing the topic with their medical care providers and discomfort with how providers responded to their questions about this topic.

The investigators concluded that HIV/AIDS prevention education is appropriate and needed among older adults, and the study participants provided insight as to how this might best be undertaken.

Journaling and Ecological Momentary Assessment (EMA). One problem with relying on surveys, questionnaires, clinical instruments, and interviews is that much of the information individuals are asked to provide is retrospective prospective in nature. In other words, people are asked to describe what has happened in the past or what they believe will happen in the future. Unless the questions are about the “here and now,” people are not providing information about what is happening in real time, as it happens. We learned back in Module 1 that recalled information is only as good as the person’s memory and how memories are constantly being revised by experiences and interpretation processes.

To improve the quality of information provided by study participants, investigators might ask them to maintain a log, diary, or journal on a frequent, routine, automatic basis, or as the phenomenon of interest occurs. For example, if the study is about anxiety, the person might write down answers to a set of questions about each anxiety episode as it is happening or as soon after as possible. Not only would this provide investigators with the ability to count the frequency of anxiety episodes, they would also have information about episode intensity, what preceded each episode, and what might have helped the person resolve each episode. The data can be analyzed more contextually than would be possible without these event records, using more global, retrospective accounts.

Another format for collecting event data is called ***ecological momentary assessment* (EMA)**. This approach minimizes recall and memory bias by sampling an individual's behavior or experiences at periodic intervals (Shiffman, Stone, & Hufford, 2008). A study participant may be signaled by the investigators at either regular or random times throughout the day or week to immediately record information about what is happening. This approach has been used in a number of alcohol-related studies to help investigators understand how alcohol use is influenced by situational, contextual circumstances—cravings, stress, affect, motivation to drink, motivation to abstain from drinking, concurrent other substance use, social and place factors (Wray, Merrill, & Monti, 2014). The approach is also presented as a possible technology assist to treating alcohol use problems by providing real-time support rather than supporting individuals' change efforts on a clinic or office schedule (Morgenstern, Kuerbis, & Muench, 2014). EMA data are collected from individuals as they function in their natural environments. The data are sensitive to naturally occurring fluctuations and patterns in the behavior of interest, rather than relying on a single, retrospective summary of the behavior. This is particularly important when what is being studied is dynamic—frequent changes in mood, behaviors influenced by environmental context, transitory thoughts or feelings, for example. EMA is facilitated by technology—cell phone access, text messaging, and web-based survey tools that can be accessed from most locations. The method is sometimes limited by certain contexts, such as schools, jails, bathrooms, locker rooms, driving, and other situations/settings where use of these technologies is discouraged or banned.

Concept Map. An individual or group process approach to understanding social phenomena is called a ***concept map***. The result of this process is a diagram depicting how ideas about a topic are related, possibly as a web of ideas. While concept mapping is useful in social work practice with individuals, families, small groups, communities, and organizations, it also can be useful in conducting exploratory or descriptive research studies. The underlying research philosophy and methodology are essentially qualitative in nature, involving a structured step-by-step approach to achieve the end goals; the approach also involves both qualitative grouping of concepts and the use of statistical methods to analyze participant input in generating the resulting concept map.

> *"Concept mapping is a structured process, focused on a topic or construct of interest, involving input from one or more participants, that produces an interpretable pictorial view (concept map) of their ideas and concepts and how these are interrelated"* (*Trochim & Donnelly*, 2007, *p.*27).

A team of investigators utilized concept mapping as a method for exploring types of support needs identified by a group of diverse sexual minority youth (Davis, Saltzburg, & Locke, 2010). The 20 study participants were served by a GLBTQ youth-focused. The participating youth generated 58 statements of need in the "idea generation" phase of the process. Next, each individual participant sorted the collection of statements in clusters that were conceptually meaningful to them, and provided a label for each of their item clusters. This step concluded with each participant rating each statement on a 7-point scale of importance (not important to extremely important) in meeting their emotional or psychological needs and again in meeting their social needs as a GLBT youth. The investigators used these data in multidimensional scaling statistical analysis, and generated concept maps.

> *"Concept maps were produced, providing graphic representations of relationships and relevance of the ideas. The maps illustrated how the youths' ideas clustered together and the rating data reflected the value given to each conceptual idea and cluster" (Davis, Saltzburg, & Locke, 2010, p. 232).*

The study participants next engaged with the investigators in group discussions about how to interpret the preliminary concept mapping results. A combination of additional qualitative and statistical methods led to the final results. The youth's statements fit into 5 clusters of needs (educating the public, policies, community support & involvement, individual responsibilities, and GLBTQ acceptance & individuality). In terms of importance, the highest ratings were assigned to "policies" statements, lowest to "community support & involvement" statements. Results from studies like this have important implications for planning culturally and developmentally relevant services in programs, organizations, communities, and service delivery systems—both in terms of process and products.

Social networks. Another method that allows graphical depiction of relationships comes from asking questions about ***social networks*** and using research methods designed to gather social network data. Social networks are the patterns of formal and informal connections that exist between individual units in complex social systems. Social network maps depict the strength and/or frequency of connections between individual units (individual persons or groups). A great deal of such research is conducted using social media interaction data, for example. Questions about how ideas, rumors, or innovations are disseminated or diffused can be answered in this manner. Social network analysis allows investigators to identify key "nodes" or "nodal points" where different networks connect, as well. Questions about how people use their social networks might rely on other forms of data collection (survey or interview questions, for example). This type of research helps us understand how social relationships and social contexts influence human behavior.

Social network data collection and analysis was used in a study to understand the phenomenon of suicide diffusion among youth social networks—how suicidality might spread among adolescents (Mueller & Abrutyn, 2015). These investigators used data previously collected in the National Longitudinal Study of Adolescent to Adult Health to map how suicidal ideation and suicide attempts by individual adolescents ("ego" individuals) might be influenced by suicide attempts of their friends ("alter ego" individuals). Ego individuals who knew about an alter ego's attempted suicide (disclosed suicide attempt) during the past year experienced higher levels of emotional distress and suicidality themselves compared to egos with an alter ego whose suicide attempt was undisclosed (ego did not know about it) or whose alter egos had no previous suicide attempts. The authors did find evidence of "the social contagion of negative emotions" but that emotional contagion alone is not relevant to the social contagion of suicide (Mueller & Abrutyn, 2015, p. 144).

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=223

Chapter Summary

In summary, there exist many options for collecting qualitative and quantitative data in social work research. The issues of measurement validity and reliability are important concerns for quantitative research. Ensuring the appropriateness and cultural competence of measurement approaches and procedures is also critically important in any type of research involving people as participants.

Module 3 Chapter 6: Study Participants

Research informing our understanding of diverse populations, social work problems, and social phenomena is entirely dependent on the participants who provide the data and observations for analysis. Guidelines for engaging participants vary by research approach: qualitative, quantitative, or mixed-methods, as well as by study purpose (descriptive, exploratory, or explanatory questions). In this chapter you will learn about:

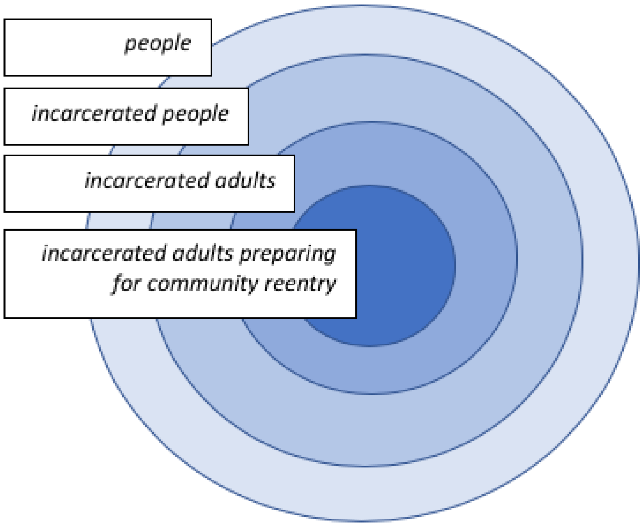
- Sample size implications
- Strategies for establishing study samples
- Participant recruitment
- Participant retention
- IRB concerns for research involving human participants

What Is a Sample?

First things first: understanding this chapter requires an understanding of what a sample is. In research, a ***sample*** is a group drawn from a population for the purposes of observation or measurement. If we measured an entire population, we would not be talking about a sample. It is usually unreasonable to attempt this feat of measuring an entire population to answer a research question—although it is sometimes possible in evaluating a small-sized program. Ideally, the sample drawn is a good representation of the population. This allows an investigator to offer generalizations based on the study results derived from the sample. Many of the quantitative statistics used in research are based on certain assumptions about the relationship of the sample to the population. If sampling error is small, then statistical values observed for the sample will be a good representation of the statistical values we would see for the entire population if we could really measure them all.

A note about populations is warranted here: a research population is not the same as the entire population. A research population is defined by certain variables of interest in the study being conducted. For example, consider a study of service needs experienced by adults following release from incarceration (Begun, Early, Hodge, 2016). The population relevant to this study was defined as men and women preparing for release from jail or (state) prison—it was not all people, or even all incarcerated people (not juveniles, not persons in federal prison, not individuals who are still serving sentences). The results of the study are generalizable only to this specific population, and since the data were collected only in the state of Ohio, generalizations to other states are made with caution (see Figure 6-1).

Figure 6-1. Sphere of generalizability to study population



Matters of Sample Size and Diversity

The question of how many study participants are needed for a specific study is complicated to answer. First, it depends on the study approach and aims. Most qualitative research is not intended to be generalizable to a larger population, but most quantitative research is. If the aim of a qualitative study is to develop a deep, rich understanding of a construct, event, experience, or other phenomenon, then the number of study participants needed might be relatively few. For example, in grounded theory studies, the need is to get to a point where all or most facets of the topic are addressed and where enough responses cluster together to have a picture of common themes emerge from the data. This ideal point of saturation is achieved when new ideas, not expressed by previous participants are no longer being presented by subsequent participants. The recommended number is in the 20-30 participant range (Creswell & Poth, 2018). Similarly, in ethnography, the number of participants depends on how many observations are required to achieve a clear picture of the culture-sharing group (Creswell & Poth, 2018). In narrative studies the aim is to develop the story of a single individual, or a few individuals involved in a single event. In focus group studies, the emphasis is on group process, so the recommendation is 5-10 persons per group—the number of groups varies from 1 to several.

The aim of quantitative research lies in sharp contrast: the aim is to generate information that is generalizable from the sample to the population. Therefore, the numbers needed are usually greater for studies designed around quantitative than around qualitative approaches. Table 6-1 presents an example generated from 12 social work students using samples of M&Ms® candies to estimate the proportion of green M&Ms® in the population of M&Ms® candies (adapted from Begun, Berger, & Otto-Salaj, 2018, p. 8).

12 Individual Students' M&M's® counts				Aggregated Teams' M&M's® counts			Aggregated Class M&M's® counts		
# green	not green	total #	% green	# green	total #	% green	# green	total #	% green
3	31	34	8.8%						
5	32	37	13.5%	19	115	16.5%			
11	33	44	25%						
4	34	38	10.5%						
5	26	31	16.1%	13	107	12.1%			
4	34	38	10.5%						
3	36	39	7.7%				67	447	15%
4	34	38	10.5%	13	119	10.9%			
6	36	42	14.3%						
11	37	48	22.9%						
9	18	27	33.3%	22	106	20.8%			
2	29	31	6.5%						
Reported green M&M's® population[a]								16%	

a This estimate is per https://www.exeter.edu/documents/mandm.pdf which also states that the Mars Candy Company no longer publishes the proportions of colors in the M&M's® product line.

As you can see, when the individual students' sample sizes were between 18 to 37, the range of values for proportion (%) of green candies was wide—as low as 6.5% and as high as 33.3%. As students combined their samples (teams of 3), the sample sizes from 106 to 119 had much less variability in the observed proportions: 10.9% to 20.8%. Finally, when the teams' samples were combined into one single class sample of 447 candies, the sample's observed proportion of 15% was close to the actual proportion for the population reported by the company (16%). (Yes, they ate the candy after completing the demonstration.)

Second, the diversity of study participants is at least as important as numbers in qualitative research studies. To ensure rich coverage of a topic, investigators seek participants with differences likely to be important to that topic—an approach that maximizes variation among study participants so that different perspectives will arise in the data. The approach to selecting participants is purposeful. For example, in a grounded theory qualitative study of benefits experienced by older adults in adult day services programs, investigators interviewed 28 service recipients (Dabelko-Schoeny & King, 2010). More important than the numbers, however, is that the investigators intentionally, strategically engaged participants from 4 different programs. The group of study participants ended up being Caucasian and African American women and men, differing in marital status, ranging in age around a mean of 78 years, experiencing different types and numbers of medical conditions, and ranging in income ($0-$88,000 annual income). The investigators' conclusions were that five themes expressed by the study participants fit into two general categories: program experiences and perceived impact of experiences. The program experience themes included social connections with other participants, empowering relationships with staff, and participation/enjoyment of activities and services. The perceived impact of experiences themes included perceived improvement in psychosocial well-being and perceived decrease in dependence and burden on their primary caregiver. These results contribute to an understanding of the perceived benefits experienced by persons who participate in adult day services programs.

Just as sample diversity is important in qualitative research, the issue of sample heterogeneity is an important consideration in quantitative research. The challenge played out in quantitative research concerns the sample having diversity that is representative of the population to which results will be generalized. In contrast to the purposeful selection process used in qualitative research, strategies for randomly selecting participants into quantitative studies are more common. More about this topic is presented in the next two sections.

Filling a Quantitative Study Design

Let's revisit a study design example from Chapter 4—the 2 X 3 design with empty cells. The original example was about waiting to see how individuals sorted themselves out in terms of the two variables of interest. This time, we are going to use the diagram to help figure out how many people of each type investigators need to include in our sample for a quantitative study. Imagine that a planned study is about participants' ratings of an outcome variable—maybe likelihood of engaging in additional college education—and the investigators wish to determine how age group and employment status relate to this outcome variable. The statistical analyses planned by the investigators require a minimum of 10 persons in each cell (more about this in Module 4).

The result would be the need for ending up with a minimum of 60 study participants, recruited from each of the 6 types (10 x 6=60).

		Employment Status		
		Fully Employed	Partially Employed	Unemployed
Age Group	<30	10	10	10
	≥30	10	10	10

Now imagine that the investigators, based on their review of literature, have decided that they really need 3 age groups: 18-25, 26-35, and over 35. Let's see what this does to the design and number of participants needed. As you can see, we now have a 3 X 3 design, which makes 9 cells. With 10 participants needed per cell, the minimum sample size is now increased to 90.

		Employment Status		
		Fully Employed	Partially Employed	Unemployed
	18-25	10	10	10
Age Group	26-35	10	10	10
	>35	10	10	10

Thus, one answer to the question of how many study participants are needed in a quantitative study is that it depends on a combination of the study design and the intended statistical analysis requirements. In our next course, SWK 3402, you will also learn about matters of statistical power and effect size in determining how many participants are needed in an experimental study about interventions. For now, it is sufficient to understand that the more complex a study design, the greater the number of study participants are required.

Sampling Strategies: Probability and Nonprobability

Earlier, when discussing the strategy used for generating the sample in a qualitative study we used the word "purposeful." This means that selection of study participants was not left to chance. In quantitative studies, however, investigators need to minimize the extent to which bias enters into the participant selection process if the sample drawn from the population is to be reflective of that population. They need every individual in the population to have the same chance, or probability, of being selected. This means using random selection strategies.

Random Selection. ***Random selection*** refers to strategies used to recruit participants into a quantitative study with the aim of maximizing external validity (the extent to which results from the sample can be generalized to the larger population). Random selection is considered a ***probability sampling*** strategy since the probability for each and every member of a population to be included in the sample is the same. This use of the word "random" should not be confused with the concept of random assignment in an experimental study. ***Random assignment*** refers to dividing the randomly selected sample into assigned experimental groups—like who receives the experimental, new intervention being tested and who receives the usual form of intervention (a lot more about this in our next course, SWK 3402).

Random selection is never perfect—there is always some error in the degree to which a sample represents a population. The aim is to minimize this error, especially to minimize systematic error or bias in the sample. Simple random sampling begins with a known pool of possible participants: they are the sampling frame. The investigator is going to select a specific number of those possible participants, at random, so that each possible participant has the same probability of being selected as any other participant. For example, imagine that a social worker wants to use the residential zipcodes for a randomly selected group of individuals in a program for men and women court-ordered to receive treatment to stop their intimate partner violence behaviors to determine the best locations to place a new branch of the program. The social worker decided that 20% of the zipcodes should be utilized. The sampling frame, or available pool, is all those served over the past year—for this example, let's say it was 1,200 persons. This means that the need is to select 240 participants for the study (1,200 x .20=240). The social worker in this situation could assign each "case" in the program a number from 1 to 1200. Then, a statistical program could be used to randomly select 240 of those numbers. An example is the free random number generator called Stat Trek (https://stattrek.com). The social worker enters 240 in the box for how many random numbers are desired, 1 for the minimum value, 1200 for the maximum value, and does not allow duplicate entries ("false" means that once a number has been selected it is removed from the pool so every number generated is unique.) The resulting table tells the social worker to use the zipcodes from "cases" numbered 186, 251, 58, 635 and so forth to the last one, 246.

240 Random Numbers

0186	0251	0058	0635	0105	0815	0448	0348	0084	0558	0335	0122	0443	0681	0284	0276	1181
0866	0294	0810	1100	0230	1002	0204	1079	1173	0571	0297	0238	0097	0520	0451	0976	0281
0848	0984	0895	1164	0037	0379	0874	0907	1125	0153	0033	0712	1074	0307	0771	0015	1084
0840	0371	0579	0592	0553	0669	0323	0161	0327	1028	0127	0110	0481	0566	0630	0438	1015
0166	1194	0827	0409	0464	0938	0396	0502	0822	1061	0345	0656	0361	0045	0674	1189	1161
0609	0182	0584	0259	0353	0951	0358	0617	0476	0899	0830	0156	0661	0028	1045	0956	0025
0417	0758	0053	0968	1186	0533	0412	1092	1135	0686	1151	0076	0263	0020	0750	0959	0971
0933	1048	0384	0540	0707	0207	0507	0489	0861	0946	1010	0818	0194	0545	0374	0007	0789
0843	0118	0776	0882	0002	1122	0725	1036	0422	0425	1053	0050	0340	0989	0243	0963	0319
0733	0130	0738	0997	0856	0079	0891	0217	1040	0089	0225	0135	0404	0797	1138	0114	0148
0366	0912	0792	0271	0315	1066	0012	0456	0643	0399	1130	1020	1032	0994	1109	0763	0920
1087	0268	0887	0869	0041	1007	1071	0879	0255	0925	0435	0387	1169	0904	0178	1156	0943
0063	0302	1104	1097	0802	0805	0233	0430	0720	0169	0622	0143	0699	1112	0191	1117	1058
0917	0140	0071	0597	0220	0468	0605	0515	0784	1177	1199	0494	0528	0746	0092	0853	0332
0694	0246															

Specs: This table of 240 random numbers was produced according to the following specifications: Numbers were randomly selected from within the range of 1 to 1200. Duplicate numbers were not allowed. This table was generated on 7/14/2018.

Essentially, this is like drawing numbers from a hat or generating lottery numbers on television. Unfortunately, sometimes the size of the starting pool is unknown or changes. For example, a social worker may want to obtain client satisfaction information from 10% of families served by an agency. There is no pre-existing pool to draw from, the need is for 10% of whoever comes into the program moving forward in time. In this case, investigators need to rely on systematic random sampling strategies–retaining an element of random selection, but more vulnerable to systematic bias. For example, an investigator might decide to collect mental health and substance use screening data from every 10thperson who comes to the emergency department for care. Assuming that there exists no systematic bias in the pattern by which people come to the emergency department, this plan is effective. Where bias enters the picture is if the sample selection is conducted only during "office hours" (M-F, 8-5). The people who come in on weekends and late night may be very different than those who are sampled.

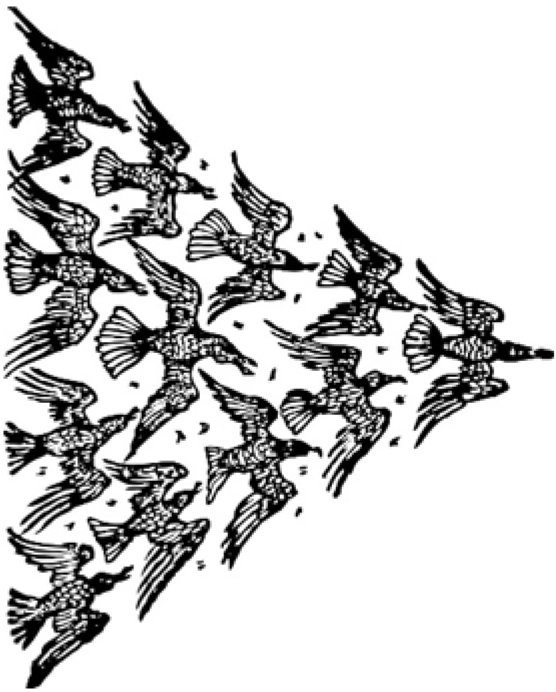
As a different example, for a study about family recreational time, taking every 10th child enrolled in after school programs is fraught with potential bias problems. First, families with more than one child in the programs have a greater chance than single-child families of being selected. Second, this type of list is already systematically arranged from the sequencing of when children are enrolled—there may be some systematic differences between those who enroll early and those who enroll later. Working from alphabetical listings has a similar problem: certain last names are more common than others in the U.S.—Smith, Johnson, Williams, and Brown are the top 4. This alters the chances of the persons with names in that part of the alphabet. Sampling bias is potentially introduced when the pool from which selection is made is systematically arranged.

Convenience Sampling. Probability sampling through random selection is not always essential because sometimes the research aims do not include generalizability back to the population. An example of this scenario is a pilot or feasibility study being conducted in preparation for a larger scale study with external validity being addressed. Another example is a study where it simply is not feasible to draw a representative sample from the general population of interest. Investigators might rely on selection from among those most easily accessed—called ***convenience sampling***. In these instances, potential bias from ***non-probability sampling*** is recognized as a generalizability limitation in interpreting the results.

For example, consider the comparative, exploratory study conducted with social work students in 4 countries: United States, Greece, Cyprus, and Jordan (Kokaliari, Roy, Panagiotopoulos, & Al-Makhamreh, 2017). The purpose of the study was to extend our understanding of social workers' perceptions concerning non-suicidal self-injury. The selection of these countries was intended to maximize heterogeneity (diversity) in terms of cultural values and religious belief systems: inclusion of students with Christian, Greek Orthodox, and Muslim religious backgrounds, and nations differing in social/political orientation toward individual and women's rights. The students participating in the survey study were those easily accessible to the investigators who were associated with social work programs in those countries. The investigators identified differences and similarities in the students' responses which are suggestive of the need to better address the problem of non-suicidal self-injury in social work education; however, they also identified the potential limitations of the study related to the convenience sampling approach. Student surveys may not be a good representation of social work practitioners, either. The observed differences in this study were that Jordanian students viewed the problem behaviors as being associated with weak/absent ties to religion, while U.S. and Greek/Cypriot students attributed the problem to mental illness and socioeconomic factors. Students from each nation were equally unprepared to understand treatment of the problem. One concern with how to better educate social workers about non-suicidal self-injury is that the behaviors may differ in the varied cultures/countries, and marked differences in stigma exist, as well.

Snowball Sampling. Social work investigators are often asking research questions related to difficult-to-locate populations—questions about uncommon problems, stigmatized phenomena, and intersectionality issues. Quite possibly, individuals who meet the study criteria know others like themselves. In these cases, investigators may rely on ***snowball sampling*** strategies: the sample builds itself from a few core individuals who identify other participants, and they identify still others. The name comes from how you might build the base of a snowman: start with a small snowball, roll it in the snow and as more snow adheres to it, the snowball grows larger (assuming it is good packing snow). This is another form of non-probability sampling.

A more complex version of snowball sampling is called respondent driven sampling. Not only does this approach involve the study participants helping to identify other potential participants, it involves some analysis of the social networks that are tapped into throughout the snowball process. For example, investigators interested in addressing questions related to the experience of human trafficking might begin with identifying a small number of individuals willing to participate in their study (the sample nucleus). These individuals would each identify one or more others who could be invited to participate. Those who accept the invitation and become participants would each identify one or more others. The hard work for study investigators begins with identifying a sufficiently diverse initial core: if they identify too many within the same social network, these members will all be identifying each other or the same few and the snowball will cease to grow. The hard work continues in terms of ensuring protection of privacy rights for those nominating others and for those being nominated for invitation. Then, the investigators must take into consideration the potential bias built into the snowball approach—these people are sufficiently similar to know each other. Entire networks of diverse others may be missed in the snowball process. This limitation is based on the principle of ***homophily***: the observation that people tend to associate with others similar to themselves in ways that are meaningful to them (i.e., race, ethnicity, social class, gender identity, sexual orientation, political or religious belief systems, athletic team loyalty, drinking habits, and many more). The saying "birds of a feather flock together" is true of human social networks, as well.



Participant Recruitment

Successfully engaging sufficient numbers and diversity of participants in a study is critically important to the study's success. ***Participant recruitment***can be conceptualized as a 3-step process (adapted from Begun, Berger, & Otto-Salaj, 2018):

- Generating initial contacts is about identifying and soliciting potential participants. This step is accomplished through a variety of advertising strategies: media advertisements (including social media outlets), posters and flyers, and mass mailings (postal and e-mail). Some success is achieved through telephone calls, as well. In recent years, the rate at which people respond to many of these approaches has declined sharply—people are finding themselves saturated with claims for their attention, and are ignoring many of these messages as "spam" or "junk" mail. Investigators also need to critically consider the potential sources of bias related to different possibilities. For example, individuals' responsiveness different types of on-line social media is influenced by age group (Snapchat, twitter, and Facebook, for example, are used differently by adolescents, emerging adults, and older adults); reading newspapers is heavily influenced by geography and social class; receiving church newsletter postings is dependent on affiliation; and, reading bus advertisements is dependent on who rides buses and which routes they frequent. In-person recruitment might be effective in specialized locations (e.g., clinics) but may not be effective with the general public (e.g., stores and shopping malls). These, too, are subject to bias in terms of who will agree to listen to the recruitment message and who will not.

Like any advertising effort, the nature of the message is important to consider. The direct contact tool needs to be crafted with consideration of why someone would WANT to engage with the study. The message needs to capture attention first, then needs to capture interest next. It also needs to be easy for someone to respond—a 24/7 phone line or e-mail address or website is more accessible than a phone number with an answering machine. Furthermore, cultural competence of the message needs to be taken into consideration: images and language need to convey a sense that "this study is about people like me."

- Screening is part of the recruitment process, as well. If strong effort is placed on the first step (generating contacts), there is strong likelihood that individuals who do not meet the study criteria will be contacted. The study investigators need to have clearly defined criteria for study inclusion/exclusion. This process can be as simple or as complex as the criteria for participation. If the study is limited to adults, then the screening can simply be a question: Are you at least 18 years of age? Anyone who answers "no" to this question fails to meet the inclusion criteria and is excluded from participation. Consider the example of a study concerning the relationship of social work education and self-esteem on social work students' social discrimination of persons with disabilities (Bean, & Hedgpeth, 2014). The direct contact method used in this survey study was to distribute the survey only to individuals who met the study criteria: being in the last semester of their social work program. On the other hand, had they distributed the survey to the entire school, they would have needed to impose a screening strategy to ensure that other students did not become enrolled in the study.

Four difficulties arise in regard to screening. First, investigators need to identify reliable, valid screening measures for each of the study criteria. Second, the screening information cannot be treated as data in the study unless consent was secured for screening and for the main study. Third, investigators need to have an appropriate response for individuals who become excluded. Consider, for example, individuals with a substance use problem who are interested in participating in a treatment study, but who fail to meet the study criteria. Ethically, investigators cannot simply turn them away without offering strong alternatives. Fourth, screening is an important part of the participant-investigator relationship. If the screening experience is tedious or otherwise unpleasant, the chances of retaining participants in the actual study diminish.

- Consenting study participants is a process, not an outcome. A signed consent form simply provides documentation that the process was engaged—the process is the goal. This means that investigators studies need to engage in interactions with potential study participants that ensures they are fully informed about the study and what should be expected—including any potential risk and benefits, as well as steps the investigators take to minimize risk. Making sure of potential participants' comprehension of the consent information is a responsibility of study investigators. The IRB review process helps ensure that the proposed procedures meet expectations for protecting potential participants rights. In addition to providing participants with sufficient information to make an informed choice about engaging with the study, investigators need to ensure that participants are making the decision without influence of coercion. Some potential coercion scenarios are more obvious than others. For example, instructors who wish to engage their students in a study should not be in the position of knowing who does or does not consent to participate, since students may believe that this knowledge could affect their grade in the course. This is a problem to be addressed in social work intervention research: clients may feel obligated to participate in a study conducted by the person delivering them

> services, out of concern that the decision might affect the services they receive or out of a desire to help the person who has been helping them. Problems might arise in relation to incentives offered to study participants, too. The amount offered might be considered coercive: $25 may not be coercive to people in some circumstances, but is a lot of money to people in other circumstances, potentially affecting their decision about participation. Legally, children and adolescents cannot provide consent—they have not achieved the age of majority. Thus, persons responsible for ensuring their welfare need to provide consent (parents, legal guardians). At the same time, investigators have a responsibility to secure the child's/adolescent's assent to participate. This means that the child or adolescent agrees to the conditions of participation.

Exceptions do exist: sometimes the consent process or consent documentation is waived by an IRB. Sometimes it is not reasonable to secure assent. Consider, for example, studies conducted with infants who would not understand an investigator's description of what is expected to happen. The investigators need to engage parents/guardians in consent and then observe the infants for cues as to their willingness to continue—for example, crying is a clear sign that the baby is done participating, assent has ended.

Participant Retention

In longitudinal research, it is critical that study participants remain engaged with the study through all phases of data collection. While participant recruitment is about people initially enrolling, ***participant retention*** is about them staying involved over time. Obviously, retention is not an issue in cross-sectional research studies. An important point for investigators to keep in mind is that retention begins with recruitment—retention only happens when study participants decide that the experience continues to be worthwhile.

The opposite of participant retention is called***participant attrition***—people dropping out of a study before completion. The obvious problem with attrition is that the number of participants may drop below the level needed for a strong study design and generalizability from the smaller sample to the population. A less obvious problem is that attrition is seldom a random process. This means that sampling bias may become introduced into the study despite all the care and effort that investigators may have applied in their initial probability random selection. For example, men or persons of color may drop out at a higher rate than women or white participants. The sample at the end of the study is what matters, so minimizing attrition is an important issue for investigators to address. It is costly to replace missing participants—starting over with recruitment and all the data collection efforts that need to be repeated (not to mention paying incentives again). Investigators need to constantly address the question: why would participants WANT to remain involved with the study?

Researchers have demonstrated that the rate of participant attrition tends to be greatest at the start of a study: if someone is going to drop out, this is most probable early on (Begun, Berger, & Otto-Salaj, 2018). That does not mean they will certainly be retained once they are involved in middle and later phases of a study; the dropout rate tends to slow down to more of a trickle as time passes. This means that investigators might engage in their most strenuous, costly retention efforts early on. In addition to making the study protocol interesting and a positive experience for participants, developing good rapport and effective systems for tracking and maintaining contact are critical.

Chapter Summary

This is the final chapter of Module 3. Here you learned about sampling in quantitative research studies, and about participant numbers and heterogeneity (diversity) as important issues in both qualitative and quantitative research. This chapter also discussed aspects of participant recruitment and retention that are important in both qualitative and quantitative studies.

While it is not quite the same as recruiting you to participate in a research study, let's think about how the content of this chapter relates to your participation in end-of-semester course/instructor evaluations. For many courses, the response rate is 30% or less—meaning that 2/3 of your classmates do not participate.

What is the greatest potential bias risk to the program's evaluation results of failing to recruit participants like you (and many of your peers)?

In a class of 25 students, how many do you think would be a good number to recruit for a strong sample?

What are 3-5 things the program could do to increase your participation in the evaluation process—what makes you WANT to participate?

Does reading this chapter and thinking about this topic make the likelihood of your participation in evaluating this course greater, less, or about the same? Why?

Module 3 Summary

This module was concerned with the many research methodology choices faced by social work, social science, and behavioral science investigators. You read about the ways in which research questions lead to research approaches: specifically, about qualitative, quantitative, and mixed-methods approaches. Next, you learned a great deal about different qualitative and quantitative research traditions and the many available options: each has different advantages, disadvantages, and goodness-of-fit with study aims to consider. In Chapter 4, you read about variables and their role in quantitative research. Next you learned about many different strategies that investigators use to collect data, and again learned about their relative advantages and disadvantages. Finally, you were presented with issues related to participants in social work, social science, and behavioral science research. This included issues of sample size and diversity, IRB concerns with involving humans as study participants, and issues in participant recruitment and retention (for longitudinal studies). This Module 3 content has great relevance to the next course module, Module 4, where we look at data analytic approaches.

Module 3 Key Terms and Definitions

administrative data: information routinely recorded in programs, agencies, and institutions as part of normal operations, not with a research intent.

alphanumeric variables: data in words or numbers that have no mathematical meaning (for example, telephone numbers, addresses, names or descriptions).

categorical variables: variables with categories of a non-numeric type, characterized by their names/labels rather than mathematical numbers (for example, race, religion, or national origin categories)

cohort: a group of individuals with a shared, common history or experience (for example a birth cohort like "baby boomers," "Gen X," "Gen Y," or "Gen Z."

***community-based participatory research* (CBPR):** a collaborative inquiry process involving an egalitarian, mutual learning partnership between community partners and researchers, both as experts in their own ways, sharing that expertise toward a common research-related goal.

concept map: the diagram resulting from a systematic research process where concepts, ideas, or constructs are identified and organized in terms of their mutual relationships.

content/artifact analysis: research analysis of information, records, recordings, or objects created in the normal course of events/living, outside of a research intent.

continuous variables: numeric variables on a continuum of values where possible values are equally distant (see interval variables).

convenience sampling: a type of non-probability sampling, subject to sampling bias, that draws a group of participants (sample) in a systematic, non-random manner based on ease of access to recruit them (for example, college students, clinic populations, club members).

correlational research: research designed to evaluate the existence and nature of relationships between variables (whether they are statistically correlated, and if so, the strength and direction of the correlation).

cross-sectional research: research designs involving only one data collection time per individual element or unit being measured, providing a single-time "snapshot" of data.

cultural competence in measurement: using measures and measurement procedures that are appropriate for the group being measured, as free from bias as possible; "cultural" is broadly defined to include gender, race, ethnicity, language, religion, national origin, ability/disability, and other factors relevant to a study population.

demographic variables: variables used to describe specific characteristics of a population, group, or sample (for example, age, gender, race, ethnicity, or income level)

dependent variable: the variable in a study design presumed or hypothesized to vary as a function of variation in the independent variable; its value is presumed to depend on the independent variable values; the "outcome" or "output" variable in a study design (i.e., usually denoted as the "y" variable, where "x" denotes the independent variable).

descriptive research: research with the purpose of creating a profile or typology (description) of a population, phenomenon, or process.

dichotomous variables: a special type of categorical variable with only two possible categories or values (for example, yes/no, meets/does not meet clinical criteria, completed/did not complete treatment).

double-barreled question: a single question that asks for information representing two different questions, may not be accurately answered in a single response; often uses the word "and" to connect the two different questions being asked.

ecological momentary assessment (EMA): a form of real-time data collection, often using technology to record events as they are happening (rather than retrospective accounts)

ethnographic research (ethnography): one of the major qualitative research traditions with the aim of understanding a population, social work problem, social phenomenon, or experience from the collective perspective of a group (rather than discrete individual perspectives).

experimental research: research designed to answer explanatory questions, testing hypotheses or theory about a phenomenon or process.

external validity: the degree to which quantitative research results based on a study sample can accurately be generalized to the larger population or other settings.

focus group: a research tradition involving interview data collected in a group setting (rather than from individual participants).

geographic information systems (GIS): a form of data representing geospatial relationships between places where events, experiences, or behaviors occur and the nature of those events, experiences, or behaviors (for example, locations of state prisons and prisoner visitation patterns).

generalizability: the extent to which findings based on sample data can be expected to reflect what is true of the population from which the sample was drawn (see external validity).

grounded theory research: one of the major qualitative research traditions, aimed at developing theoretical explanations about a phenomenon or process from data provided by intentionally/purposively selected individuals.

heterogeneity: differences, diversity, variability in a group or population.

homophily: observation that people tend to associate with others similar to themselves.

independent variable: the variable in a study design presumed or hypothesized to vary on its own, but that determines variation observed in the dependent (outcome or output) variable; the "input" variable in a study design (i.e., usually denoted as the "*x*" variable, where "*y*" denotes the dependent variable).

internal validity: the degree of confidence that can be applied to the results of an experimental, explanatory study; study integrity; ability to rule out alternative explanations for the experimental results.

inter-observer reliability: different observers, raters, or coders getting the same result when evaluating the same qualitative, open-ended question, or observational data; also called inter-rater or inter-coder reliability.

interval variable: continuous variable (on a continuum) where the distance between possible values (intervals) are equivalent (for example, age where the 1-point interval between 40 and 41 is the same as the 1-point interval between 49 and 50, 1 year).

interview protocols: data are systematically collected by investigators interact with participants to provide information in a "live" dialogue (may be in-person, by telephone, or other interactive media).

key informants: persons identified as close and knowledgeable about someone for whom data are being recorded (for example, parents of young children, spouse or caregiver to a person with special care needs)

latent variable: a variable not directly observed but mathematically or statistically inferred from other variables that are directly measured.

longitudinal research: study design where data are collected repeatedly from the same individuals/elements/units over time.

measurement reliability: indicates the degree of accuracy or precision in measuring the construct or variable of interest.

measurement validity: the extent to which a measurement tool or instrument adequately measures the concept, construct, or variable of interest, comprised of several types of validity (construct, concurrent, predictive, and others).

mixed-methods research: a study approach that integrates aims, questions, and methods in both qualitative and quantitative traditions.

narrative inquiry: one of the major qualitative traditions, eliciting and analyzing the "stories" and their contexts as experienced by one or a few individuals.

naturalistic observation: a form of data collection concerning individuals' behavior in their naturally occurring environments; the observer is a non-participant (objective observer).

nominal variables: another name for categorical variables (categories have names, are not numeric)

non-probability sampling: forms of participant selection that are not random; participants have unequal chances of being selected from the population.

ordinal variables: categories of a variable have numeric relationships, can be sequenced in numeric order but are not continuous interval variables (for example, ratings of frequency from never to always)

participant attrition: the opposite of participant retention, refers to participants dropping out of a (longitudinal) study before completion

participant recruitment: process of initially engaging potential participants in a research study

participant retention: the opposite of attrition, refers to maintaining participation for the duration of a (longitudinal) study

***participatory action research* (PAR):** investigators are active both as researchers and change agents in the group/process under study.

phenomenological research (phenomenology): one of the major qualitative traditions, involving the constructing of shared, common meaning about an event, experience, or phenomenon out of the experiences and perceptions of a group of study participants.

photovoice: a data collection approach involving a combination of participants recording images (photographing) aspects of their experiences and environments along with qualitative data about their explanations of those images/photographs.

primary data: data collected with the intention of meeting aims of the specific research study for which the data are collected.

probability sampling: selecting research participants using methods that ensure every member of the original population has an equal chance (probability) of being selected; (see random selection).

proxy variables: variables that represent or stand for something not directly measured (for example, birth order representing amount of parental attention experienced).

qualitative research: research with the aim of describing or exploring populations and phenomena as they naturally occur; approaches flow from inductive reasoning process; multiple qualitative traditions exist.

quantitative research: research with descriptive, exploratory, or explanatory aims where the data collected are in quantifiable terms; approaches flow from deductive reasoning process; usually involves some form of statistical analysis or reporting; built from a

random assignment: refers to dividing the randomly selected sample into assigned experimental groups

random selection: refers to strategies used to recruit participants into a quantitative study with the aim of maximizing external validity (the extent to which results from the sample can be generalized to the larger population)

sample: a group drawn from a population for the purposes of observation or measurement.

scale variable: another name for an interval variable.

screening and assessment tools: instruments or tools used in professional practice to evaluate clients, but which may also be used in research to collect data.

secondary data: research using data originally collected for a different research purpose (a secondary use of the data).

snowball sampling: building a study sample from having enrolled participants identify others who may be recruited as participants, too.

social desirability: a person's tendency to provide answers that seem to be more acceptable/correct than would be their true answers; a potential type of bias in data.

social networks: the formal and informal connections that exist between individual units in complex social systems.

survey methodology: research approach involving probability sampling and data collection by use of survey tools (whether by in-person interview, telephone, paper-and-pencil, on-line, or otherwise collected).

triangulation: strategy of applying multiple methods to study a single phenomenon, generating a unified understanding by combining the results.

unit of analysis: data may be collected from individuals, but the unit of analysis might be the individuals themselves, or higher order units to which they belongdyads, couples, families, classrooms, neighborhoods, and others.

variable: elements, entities, or factors that can change (vary) in a research study.

Module 3 References

Altschuler, J., & Katz, A. (2015). "Of course it's relevant!": A focus group study of older adults' perceived importance of HIV/AIDS prevention education. *Qualitative Social Work*, 14(5), 687-701.

Auerbach, D.M., Darrow, W.W., Jaffe, H.W., & Curran, J.W. (1984). A cluster of cases of the acquired immune deficiency syndrome: Patients linked by sexual contact. American Journal of Medicine, 76, 4878-492.

Babor, T.F., Higgins-Biddle, J.C., Saunders, J.B., & Monteiro, M.G. (2010). The Alcohol Use Disorders Identification Test (AUDIT): Guidelines for use in primary care, 2nded. Geneva: The World Health Organization (WHO), Department of Mental Health and Substance Dependence.

Baer, L., & Blais, M.A. (2010). *Handbook of clinical rating scales and assessment in psychiatry and mental health.*NY: Humana Press.

Barnhart, S. & Maguire-Jack, K. (2016). Single mothers in their communities: The mediating role of parenting stress and depression between social cohesion, social control and child maltreatment. *Children and Youth Services Review*, 70, 37-45.

Begun, A.L. (*unpublished master's thesis*). Content analysis of sibling relationships as presented in preschool children's books. University of Michigan, Department of Psychology (1980).

Begun, A.L., Berger, L.K., & Salm Ward, T.C. (2011). Using a lifecourse context for exploring alcohol change attempts and treatment efforts among individuals with alcohol dependency. *Journal of Social Work Practice in the Addictions*, 11(2), 101-123.

Begun, A. L., Brondino, M. J., Bolt, D., Weinstein, B., Strodthoff, T., & Shelley, G. (2008). The revised Safe At Home Instrument for assessing readiness to change intimate partner violence. *Violence and Victims*, 23(4), 508-524.

Begun, A.L., & Carter, J. (2017). Career implications of doctoral social work student debt load. *Journal of Social Work Education*, 53(2), 161-173.

Begun, A.L., Early, T.J., & Hodge, A. (2016). Mental health and substance abuse service engagement by men and women during community reentry following incarceration. *Administration and Policy in Mental health and Mental Health Services Research*, 43(2), 207-218.

Begun, A.L., Hodge, A.I., & Early, T.(2017).A family systems perspective in prisoner reentry. In S. Stojkovic (Ed.), *Prisoner reentry: Critical issues and policy directions*, (pp.85-144).NY: Palgrave Publishing.

Begun, A. L., Murphy, C., Bolt, D., Weinstein, B., Strodthoff, T., Short, L., & Shelley, G. (2003). Characteristics of the Safe At Home Instrument for Assessing Readiness to Change Intimate Partner Violence,*Research on Social Work Practice*, 13(1), 80-107.

Begun, A. L., Shelley, G., Strodthoff, T., & Short, L. (2001). Adopting a stages of change approach for individuals who are violent with their intimate partners. *Journal of Aggression, Maltreatment & Trauma*, 5(2), 105-127.

Begun, S. (2015). A mixed-methods examination of family planning and social work: Implications for research, education, and the role of personal beliefs. Contraception, 92(4), 389-390.

Bowen, E. (2017). Conduct disorder symptoms in pre-school children exposed to intimate partner violence: Gender differences in risk and resilience. *Journal of Child & Adolescent Trauma*, 10, 97-107.

Cameron, R., & Miller, P. (2007). Mixed methods research: Phoenix of the paradigm wars. Paper presented at 21stannual Australian & New Zealand Academy of Management (ANZAM) conference, Sydney, Australia (December). Retrieved from https://www.researchgate.net/publication/37357300_Mixed_method_research_phoenix_of_the_paradigm_wars

Carlson, J.F.. Geisinger, K.F., & Jonson, J.L. (Eds.). (2017). *The twentieth mental measurements yearbook*.Lincoln, NE: Buros Center for Testing.

Cash, S.J., Thelwall, M., Peck, S.N., Ferrell, J.Z., & Bridge, A. (2013). Adolescent suicide statements on MySpace. *Cyberpsychology, Behavior and Social Networking*, 16(3), 166-174.

Corcoran, K., & Fischer, J. (2013). *Measures for clinical practice and research*, 5*thed., v.* 1-2. NY: Oxford University Press.

Cozby, P.C. (2007). *Methods in behavioral research, ninth ed.*Boston: McGraw-Hill.

Creswell, J.W. (2014). A concise introduction to mixed methods research. Thousand Oaks, CA: Sage.

Creswell, J.W., & Poth, C.N. (2018).*Qualitative inquiry & research design: Choosing among five approaches, 4thed*. Thousand Oaks, CA: Sage.

Curran, J.W., & Jaffe, H.W. (2011). AIDS: The early years and CDC's response. *Morbidity and Mortality Weekly Report* (MMWR) *Supplements*, 60(4), 64-69 (October 7).

Dabelko-Schoeny, H., & King, S. (2010). In their own words: Participants' perceptions the impact of adult day services.*Journal of Gerontological Social Work*, 53(2), 176-192.

Davis, T. S., Saltzburg, S., & Locke, C. R. (2010). Assessing community needs of sexual minority youth: Modeling Concept Mapping for service planning. *Journal of Gay & Lesbian Social Services*, 22(3), 226- 249.

Denzin, N.K., & Lincoln, Y., Eds. (2011). *The SAGE handbook of qualitative research, 4thed.* Thousand Oaks, CA: Sage.

Engel, R.J., & Schutt, R.K. (2013). *The practice of research in social work, third ed.* Thousand Oaks, CA: Sage.

Fisher-Borne, M., & Brown, A. (2018). A case study using photovoice to explore racial and social identity among young Black men: Implications for social work research and practice. *Journal of Ethnic & Cultural Diversity in Social Work*, 27(2), 157-175.

Fowler, F.J., & Cosenza, C. (2009). Design and evaluation of survey questions. In L. Bickman & D.J. Rog, Eds.,*The Sage handbook of applied social research methods, 2nded.*,(pp. 375-412). Thousand Oaks, CA: Sage.

Freisthler, B., Gruenewald, P.J., & Price Wolf, J. (2015). Examining the relationship between marijuana use, medical marijuana dispensaries, and abusive and neglectful parenting. *Journal of Child Abuse & Neglect*, 48,170-178.

Glaser, B.G., & Strauss, A.L. (1967). *The discovery of grounded theory: Strategies for qualitative research.*Chicago: Aldine Publishing.

Grinnell, R.M., & Unrau, Y.A. (2014). *Social work research and evaluation: Foundations of evidence-based practice, 10thed.* NY: Oxford University Press.

Hagborg, J.M., Berglund, K., & Fahlke, C. (2018). Evidence for a relationship between child maltreatment and absenteeism among high-school students in Sweden. *Child Abuse & Neglect*, 75, 41-49.

Helm, D. (2016). Sense-making in a social work office: An ethnographic study of safeguarding judgements.*Child & Family Social Work*, 21(1), 26-35.

Holosko, M.J. (2006). *Primer for critiquing social research: A student guide*. Belmont, CA: Brooks/Cole.

Jirek, S.L. (2017). Narrative reconstruction and post-traumatic growth among trauma survivors: The importance of narrative in social work research and practice. Qualitative Social Work, 16(2), 166-188.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research*, 1(2), 112-133.

Kaiser, M. (2017). Redefining food security in a community context: An exploration of community food security indicators and social worker roles in community food strategies. *Journal of Community Practice*, 25(2), 213-234.

Kettle, M. (2018). A balancing act: A grounded theory study of the professional judgement of child protection social workers. *Journal of Social Work Practice: Psychotherapeutic Approaches in Health, Welfare, and the Community*, 32(2), 219-231.

Kokaliari, E.D., Roy, A.W., Panagiotopoulos, C., & Al-Makhamreh, S. (2017). An exploratory comparative study of perspectives on non-suicidal self-injurious behaviors among social work students in the United States, Greece, Cyprus, and Jordan: Implications for social work practice and education. *International Social Work*, 60(4), 1015-1027.

Kyriakakis, S., Waller, B., Kagotho, N., & Edmond, T. (2015). Conducting safe research with at-risk populations: Design strategies from a study with unauthorized immigrant women experiencing intimate abuse. *Qualitative Social Work*, 14(2), 259-274.

Lanier, P., Maguire-Jack, K., Walsh, T., Drake, B., & Hubel, G. (2014). Race and ethnic differences in early childhood maltreatment in the United States. *Journal of Developmental & Behavioral Pediatrics*, 35(7), 419-326.

Lawson, H.A. (2015). Introduction. In H.A. Lawson, J.C. Caringi, L. Pyles, J.M. Jurkowski, & C.T. Bozlak, Eds. (2015). *Participatory action research, (pp. ix-xxvii).*NY: Oxford University Press, Pocket Guides to Social Work Research.

Lee, S.J., Hoffman, g., & Harris, D. (2016). Community-based participatory research (CBPR) needs assessment of parenting support programs for fathers.*Children and Youth Services Review*, 66, 76-84.

Lopez, S.J., & Snyder, C.R. (2003). *Positive psychological assessment: A handbook of models and measures: A handbook of models and measures.*Washington, DC: American Psychological Association.

McConnaughy, E.A.., Prochaska, J.O., & Velicer, W.F. (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research & Practice*, 20(3), 368-375.

Miller, D.C., & Salkind, N.J. (2002). *Handbook of research design & social measurement, 6thed.*Thousand Oaks, CA: Sage.

Minnes, S., Min, M.O., Short, E.J., Wu, M., Lang, A., Yoon, S., & Singer, L.T. (2106). Executive functioning in children with prenatal cocaine exposure (12-15 years). *Neurotoxicology and teratology*, 57, 79-86.

Morgenstern, J., Kuerbis, A., & Muench, F. (2014). Ecological momentary assessment and alcohol use disorder treatment,*Alcohol Research: Current Reviews*, 36(1), 101-109.

Mueller, A.S., & Abrutyn, S. (2015). Suicidal disclosures among friends: Using social network data to understand suicide contagion. *Journal of Health and Social Behavior*, 56(1), 131-148.

National Highway Transportation Safety Administration (NHTSA). (2010). *National survey of drinking and driving attitudes and behaviors*: 2008. *Volume* I: *Summary report.*Washington, DC: U.S. Department of Transportation. Retrieved from https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/811342.pdf

Nguyen, D., Kit, B., & Carroll, M. (2015). *Abnormal cholesterol among children and adolescents in the United States*, 2011-2014. NCHS *data brief*, *no* 228. Hyattsville, MD: National Center for Health Statistics.

Nilsson, K.K., Straarup, K.N., & Halvorsen, M. (2015). Early maladaptive schemas: A comparison between bipolar disorder and major depressive disorder. *Clinical Psychology and Psychotherapy*, 22, 387-391.

Oktay, J.S. (2012). *Grounded theory*. NY: Oxford University Press, Pocket Guides to Social Work Research series.

Olate, R., Salas-Wright, C., & Vaughn, M.G. (2012). Predictors of violence and delinquency among high risk youth and youth gang members in San Salvador, El Salvador. *International Social Work*, 55(3), 383-401.

Padgett, D.K. (2008). *Qualitative methods in social work research*, *2nded.*Thousand Oaks, CA: Sage.

Perron, B.E., & Gillespie, D.F. (2015). *Key concepts in measurement*. NY: Oxford University Press. Pocket Guides to Social Work Research.

Price Wolf, J., Freisthler, B., Kepple, N.J., & Chavez, R. (2017). The places parents go: Understanding the breadth, scope, and experiences of activity spaces for parents. *GeoJournal*, 82,355-368.

Riley, C., Roy, B., Harari, N., Vashi, A., Violano, P., Greene, A., ...Wang, E.A. (2017). Preparing for disaster: A cross-sectional study of social connection and gun violence. *Journal of Urban Health*, 94, 619-628.

Rohrer, J.M., Egloff, B., & Schmukle, S.C. (2015). Examining the effects of birth order on personality. *Proceedings of the National Academy of Sciences*, *112*(46), 14224-14229.

Rush, A.J., First, M.B., & Blacker, D. (2008). *Handbook of psychiatric measures, 2nded.*Washington, DC: American Psychiatric Publications.

Russell, M. (1994). New assessment tools for risk drinking during pregnancy: T-ACE, TWEAK, and others. *Alcohol Health & Research World*,*18*(1), 55-61.

Sajatovoc, M., & Ramirez, L.F. (2012). *Rating scales in mental health,3rded.* Baltimore, MD: Johns Hopkins University Press.

Salas-Wright, C.P., Vaughn, M.G., & Ugalde, J. (2016). A typology of substance use among pregnant teens in the United States. *Maternal & Child Health Journal*, 20(646-654).

Substance Abuse and Mental Health Services Administration (SAMHSA). (2017). *Key substance use and mental health indicators in the United States: Results from the 2016 National Survey on Drug Use and Health*(HHS Publication No. SMA 17-5044, NSDUH Series H-52). Rockville, MD: Center for Behavioral Health Statistics and Quality. Retrieved from https://www.samhsa.gov/data/

Shiffman, S., Stone, A.A., & Hufford, M.R. (2008). Ecological momentary assessment.*Annual Review of Clinical Psychology*, 4, 1-32.

Thomas, K.A., & Rickwood, D.J. (2016). One woman's journey of recovery from mental illness—Hopes, back-up plans, rebuilding self and service support. *Qualitative Social Work*, 15(4), 501-517.

Touliatos, J., Perlmutter, B.F., Straus, M.A., & Holden, G.W. (2001). *Handbook of family measurement techniques, vols. 1, 2 & 3.*Thousand Oaks: Sage.

Tran, T., Nguyen, T., & Chan, K. (2016). *Developing cross-cultural measures in social work research and evaluation, 2ndedition.* NY: Oxford University Press, Pocket Guides to Social Work Research series.

Trochim, W.M.K., & Donnelly, J.P. (2007). *The research methods knowledge base, 3rded.* Mason, OH: Thomson Custom Solutions.

Vaughn, M.G., Perron, B.E., Abdon, A., Olate, R., Groom, R., & Wu, L.T. (2012). Correlates of handgun carrying among adolescents in the United States. Journal of Interpersonal Violence, 27(10), 2003-2021.

Wang, C. (1999). Photovoice: A participatory action research strategy applied to women's health. *Journal of Women's Health*, 8(2), 185-192.

Wray, T.B., Merrill, J.E., & Monti, P.E. (2014). Using ecological momentary assessment (EMA) to assess situation-level predictors of alcohol use and alcohol-related consequences. *Alcohol Research: Current Reviews*, 36(1), 19-27.

Yoon, S., Kobulsky, J.M., Yoon, D., & Kim, W. (2017). *Children and Youth Services Review*, 82(1), 271-279.

MODULE 4

Module 4 Introduction

This module builds on earlier modules about understanding and working with data concerning diverse populations, social phenomena, and social work problems. First, we learn basics about coding and analyzing qualitative data. Then, we learn basics about analyzing quantitative data; these basics are further developed through completing several linked Excel workbook learning activities. In this module, we learn about conducting and interpreting basic univariate and bivariate statistical analyses related to proportion, frequency, probability, percentage, mean, median, mode, variance, normal distribution, standard deviation, bar graphs, pie charts, histograms, and scatterplots. We also learn about comparing different demographic groups using 5 statistical analytic approaches to test hypotheses (null and alternative hypothesis testing): single sample t-tests, independent sample t-tests, one-way analysis of variance, chi-square, and correlation. We become familiar with Type I and Type II errors, and with basic nonparametric principles, as well.

Reading Objectives

After engaging with these reading materials and learning resources, you should be able to:

- Describe how different types of qualitative data might be interpreted or analyzed;
- Recognize and practice univariate data analysis for different types of quantitative variables (categorical, dichotomous categorical, ordinal, and interval/continuous variable types);
- Explain the logic underlying inferential statistics and null hypothesis testing, as well as Type I and Type II errors;
- Identify underlying principles and practice 5 inferential

statistics approaches with quantitative data (univariate and bivariate);

- Interpret results of one-sample t-test, independent samples t-test, one-way analysis of variance (Anova), chi-square, and correlation analyses;
- Recognize basic principles of non-parametric data analysis;
- Identify and practice steps in data analysis conducted using Excel software;
- Define key terms related to working with qualitative and quantitative data.

Module 4 Chapter 1: Working With Qualitative Data

Prior modules introduced qualitative approaches and methods for study design and data collection. You learned that exploratory and descriptive research questions are often addressed using qualitative methodologies—naturalistic observation, interview, focus group, social network, GIS, or open-ended survey data, to name a few. Qualitative studies do not necessarily test hypotheses about the data (although they can test hypotheses generated by prior theory (Glaser & Struss, 1967). Qualitative studies often use data to develop an understanding of social work problems, social phenomena, or diverse populations. The focus of this chapter is on what investigators do with the collected qualitative data to begin answering their research questions.

In this chapter you will learn about:

- moving from qualitative data collection to data preparation;
- coding qualitative data.

Qualitative Data Preparation

Investigators engaged in qualitative research are careful to collect data in a manner that preserves, to the closest possible degree, the specific wording and context or what their study participants share. Technology tools can assist in capturing participants' statements ***verbatim***. For example, digital audio or video recording is possible with small, portable recorders, cell phones, or dictation devices. Digital recording is preferable to recordings based on "tapes" because they are more durable and less vulnerable to accidental destruction.

Furthermore, digital software exists to help transcribe these recordings into written text (more about this later). The entire process of data collection and preparation (recording and ***transcription***) needs to be approved by the institutional review board (IRB) for human subject participation and consented to by the study participants.

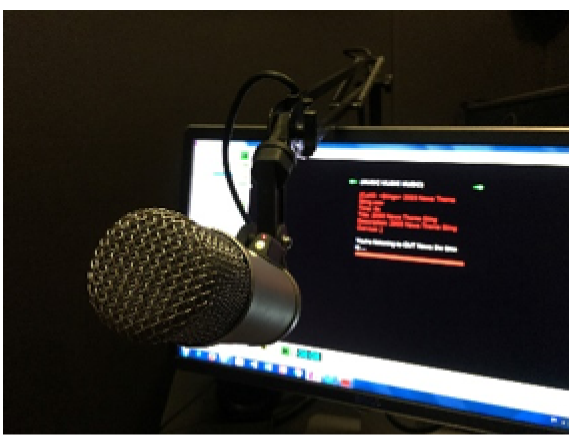
However, seasoned investigators do not rely on technology alone. There exist many painful stories concerning data lost through technology failures: not picking up the narrative in the first place because microphones or cameras were not sufficiently sensitive or were improperly placed, batteries running out, or damaged recording devices; and, including accidentally erased recordings or other epic failures. Audio and video recordings are usually supplemented with ***field notes***—these are written (or typed) either by the person conducting the observation, interview, or focus group, or by a collaborating observer/recorder. It is difficult to effectively play both roles, interviewer and recorder, and maintain strong interview rapport with participants; engaging a collaborator as recorder/note-taker is well-advised.

There exists another, important reason for these field notes: they contribute rich descriptive detail about the context of statements made, supplementing the recorded and transcribed participant statements, infusing the record with greater meaning. For example, field notes can clarify who was the speaker when recorded voices sound similar. And, the notes can describe changes in body language, long pauses, facial expressions, making or losing eye contact, or other events that can help investigators interpret meaning from the ***context*** of what is said.

To demonstrate, consider an early study of children's emerging sibling relationships (Nadelman & Begun, 1982). The investigators engaged firstborn preschool aged children (2 ½ to 5 years old) from families expecting their second baby in ***projective*** doll play, a structured format for observational data. Each of the children, in their own homes, were introduced to the standardized, portable doll house and doll family (parents, child, and baby matching their own race and family composition). What the child said was audio-recorded throughout the doll play session. The investigator also kept field notes describing each child's behavior with the dolls with comments made every 30 seconds. These field notes were particularly helpful in understanding child utterances, since the children's language skills were emerging. The notes also explained long silences that occurred when children, upon discovering the toy toilet, ran off to the potty themselves—potty training was an active part of the children's reality at this age, and a reminder to go was often triggered by play with the toy toilet. Field notes also helped when the doll play sessions were repeated after the birth of the younger sibling. The children often ran to the baby when playing with the doll baby, and the investigators had notes characterizing a child's touching, eye contact, and other interactions with the real baby brother or sister. Combining these different types of qualitative data allowed investigators to develop a rich description of emerging sibling relationships that extended far beyond the preexisting 'sibling rivalry' paradigm.

<u>Qualitative Data Transcription</u>. Investigators may choose to work with observational data in real time—***coding*** observed behaviors as they occur. For the sake of reliable and verifiable data, however, they often choose to work from recordings. One challenge with recordings is the necessity to replay them, over and over, to analyze the data. Instead, investigators often choose to work from transcripts of the recorded interview, focus group, or observation sessions; reading the transcript is often faster than repeatedly rewinding and replaying content for coding purposes. As previously noted, however, some of the rich context may be lost in translation from audio or video recording to written transcription.

Data transcription is time consuming. Transcribing a one-hour interview could require four or more hours of transcription time—assuming the recording is clear and easy to interpret and there is only one person speaking. It could require upwards of ten hours in real time to transcribe a one-hour group interview (a family with multiple members or a focus group, for example). Digital transcription software can help (e.g., Dragon Speak®) but often introduces inaccuracies, necessitating a great deal of time re-checking the transcription to be certain of its accuracy. These software packages MAY have contractual agreements that violate research standards for confidentiality and data security—reading the fine print before clicking on "ACCEPT" is critical, especially for "free" transcription software packages. Investigators need to consider the relative cost of their own time, the time and experience of study team members, or professional transcription (e.g., ranges $1 to $5 per minute across several websites visited in August, 2018). Furthermore, transcription services need to be approved by the institutional review board (IRB) when a study involving human participants is proposed, and participant consent is required, as well.



Try this out for yourself: watch a brief Youtube video, such as "Charlie Bit My Finger–Again." While listening to the video but not watching the screen, write down everything you hear the children vocalize. Now replay the video, listening while reviewing your transcription.

- How accurately did you record the children's vocalizations?

Replay the video again, this time both watching and listening. Compare your audio transcription to what you see and hear–the seeing being a context for what you hear.

- How accurately did your audio transcription reflect what happened–the children's vocalizations and the context?
- How much meaning do you think you lost without tone of voice and body language/gestures?
- What do you think would improve the quality of the data you transcribed, moving from the video to the written word?

Qualitative Data Coding

The process of coding qualitative data is systematic and should be replicable–that being one hallmark of empirical evidence. This does

not mean that different investigators would draw the same conclusions about the data–it means that others would be able to repeat the process, following in a systematic fashion, use of the same raw data. This represents a major distinction between qualitative and quantitative research: in qualitative studies, the investigator is recognized and accepted as part of the data interpretation process; in quantitative studies, investigator impact is minimized to enhance internal validity. Either way, however, the procedures used are clearly identified and replicable, even if the results in qualitative coding differ with different investigators.
Qualitative data coding and analysis is about grouping terms, concepts, ideas, images, or other elements together into themes or categories. The themes and categories provide a means of organizing participants' data in meaningful ways. Qualitative analysis:

> "aims to make sense of and give meaning to the data collected. In general, the process of qualitative data analysis involves the identification and organization of themes or patterns from the words, text, and narratives obtained in the data collection" (Corcoran & Secret, 2013, p. 166).

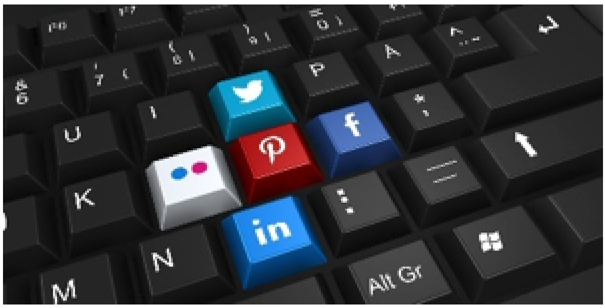
Coding themes, patterns, or categories derives from one of two sources, depending on the qualitative study approach adopted and the nature of the qualitative data collected. Contents may be analyzed based on predetermined categories (***a priori coding***) or the categories may emerge from the contents being analyzed. Predetermined categories come from hypotheses based on theory or literature. Categories emerging from the data is a process referred to as ***open coding***, which involves creating the categories or groupings, confirming that the contents of each have points of similarity or overlap, and providing the categories or groupings with meaningful labels or names.
Tools for helping interpret qualitative data may be as simple as

colored highlighters applied to a printed transcription, statements copied onto index cards which can be shuffled and re-arranged, or digital highlighting applied to a Word® or .pdf document. A number of sophisticated analysis-assistance software tools also exist: four often reported in the literature are Atlas.ti®, NVivo®, MAXQDA®, and Dedoose®. Several qualitative analysis assistance software packages are free to download, at least in their "lite" form (e.g., Provalis QDA Miner®).

When selecting qualitative data analysis programs, investigators need to consider several factors:

- These programs do not DO the analysis, they support the investigator doing analyses—remember, the investigator is the "tool" for determining themes and codes.
- Some programs only assist with text data, others assist with analysis of images and other forms of data, as well (e.g., Dedoose®).
- Cloud-based programs and some free packages may have practices that violate data confidentiality or security requirements established by an institutional review board (IRB).

As previously noted, an investigator may approach qualitative data with a pre-established list of categories that are applied to the data—deductive or *a priori* coding. The investigator applies the predetermined coding categories with the data in terms of whether each them appears or, possibly, how frequently the theme appears (depending on the research question). For example, investigators were interested in exploring the nature and extent of personal information shared by adolescents in the MySpace social networking site (Hinduja & Patchin, 2008). The study addressed a widespread concern that individuals were providing information that left them vulnerable to sexual predators. The team began with an *a priori* coding scheme for the type of information youth publicly post: they conducted a content analysis with a randomly sampled set of 9,282 MySpace profile pages. The coding included first name, full name, birth date, telephone number, address or city/state of residence, school attended, email address, instant messaging screen name, pictures, pictures in swimsuit or underwear, and evidence of alcohol/other substance use. In terms of identification, over 38% of pages sampled provided the adolescent's first name, almost 9% provided a full name, over 81% provided a city of residence, and almost 28% referenced their school. Furthermore, almost 57% of sampled profiles included a photograph, and over 15% had a friend's photograph in a swimsuit or underwear, while slightly over 5% had a photograph of themselves in swimsuit/underwear; over 18% presented evidence of alcohol use, 7.5% tobacco use, and 1.7% marijuana use. About 40% of youth set access to "private," meaning that the other 60% of profiles were potentially viewable by anyone. While single pieces of information alone might not lead to identifying an individual, in combination (especially with pictures included) information shared makes identification and/or personal contact possible. Since this research was completed, public education about safer Internet use may have had an impact on individuals' posting patterns, and organizations have introduced security measures that may reduce vulnerability. However, concern remains that individuals, including adolescents, are vulnerable to predators or exploitation based on what is shared across public domains of the Internet.



Open Coding. In contrast, when investigators approach qualitative data without predetermined or preconceived ideas about categories and themes present in participants' responses, they engage in open coding (an inductive process). This is the foundation of the qualitative approach called grounded theory. Open coding was defined by Rubin and Babbie (2013) as:

> A *qualitative data processing method in which, instead of starting out with a list of code categories derived from theory, one develops code categories through close examination of qualitative data*" (*p.* 337).

Accompanying the investigators' description of the resulting coding categories is a set of memos, notes, or journaling entries that depict what the investigator was thinking in making these coding decisions. This information helps guide others in how the results emerged to provide transparency and replicability. While qualitative research is science and results in empirical evidence, performing it well can be compared to the art of dance (Engel & Schutt, 2018). This is because it requires the investigator to maintain a state of openness in interpreting the variety of ideas shared by participants while concurrently maintaining objectivity in applying the methods and a subjective awareness and reflection about the dynamic processes by which themes and categories emerge.

For example, investigators reported on a study that utilized grounded theory methods to explore the ways that transgender persons are depicted in the media (McInroy & Craig, 2015). They conducted in-depth, semi-structured interviews of several hours length with 19 young adults who self-identify as LGBTQ persons. The interviews were audio-recorded and transcribed; coding of each interview was conducted by three separate coders. The investigators found, first of all, that participants described their experiences with offline media (e.g., television and movies) differently from their experiences with online media (e.g., websites and social media). Transphobic representations (negative reactions or opinions concerning transgender persons) emerged from the thematic coding as an issue with offline media more significantly than issues of homophobia (negative reactions of opinions concerning gay, lesbian, or bisexual persons). The participants also expressed that offline media exhibits very little in the way of positive representations of transgender people; this was experienced as a contrast with more positive representations of LGBTQ persons. The participants also experienced transphobia to a greater extent in the online environment, possibly because of the anonymity allowed in this environment. On the other hand, the online environment offered greater numbers and range of supportive, helpful options for transgender youth. The study authors shared a large number of verbatim quotes from their study participants to demonstrate the coding categories that emerged from the data.

Cross-Checking Coding Decisions

In the example just presented, notice that the authors reported three individuals coded each interview. This is an important aspect in preserving the integrity of qualitative analysis—that one person's coding decisions be confirmed by others' independent decisions. Furthermore, many studies take their findings back to a subsample of the original participants or a new sample to learn if their conclusions are a good fit with the participants' lived experiences. These activities are part of the "assessing interpretations" step in the qualitative "data analysis spiral" (adapted from Creswell & Poth, 2018, p. 186):

- data collection
- managing and organizing the data
- reading and making memo notes about emergent ideas (initial analysis)
- describing and classifying codes into themes (data reduction)
- developing interpretations (including how themes relate to one

another, how they are distinct)

- assessing interpretations (how themes/interpretations hold up in other examples, how well other investigators and study participants agree with interpretations)
- representing and visualizing the data
- presenting an account of the findings.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=428

Chapter Summary

In this chapter, you were introduced to general concepts and issues involved in managing and analyzing qualitative data. There exist excellent textbooks and online demonstrations for learning to master these skills. This content also applies to mixed methods research since many aspects of mixed methods data are qualitative.

Module 4 Chapter 2: Working With Quantitative Descriptive Data

Prior modules introduced quantitative approaches and methods, including study design and data collection methods. This chapter examines what investigators do with collected quantitative data to begin answering their research questions. As you work your way through the contents of this chapter, you will find it helpful to have a basic calculator handy (even one on your smart phone will do). In addition, you are presented with links to the Excel workbook for learning activities that need to be completed using a computer (data analyses do not work on most tablets or phones).

In this chapter you will learn:

- the distinction between univariate and bivariate analyses,
- descriptive statistics with categorical and numeric variables,
- computing and interpreting basic inferential statistics
- understanding graphs, figures, and tables with basic statistics.

Univariate and Bivariate Analyses

The way we think about quantitative variables has to do with how they are operationalized in a particular study. First, let us consider the distinction between univariate and bivariate work with quantitative variables, then explore distinctions between independent and dependent variables. When the goal is to describe a social work problem, diverse population, or social phenomenon along a particular dimension, investigators engage in ***univariate analyses***. The prefix “uni” means one (like a unicycle having one wheel).

Univariate Analysis. Univariate analysis means that one variable is analyzed at a time; if multiple variables are of interest, univariate analyses are repeated for each. Univariate analysis has the purpose of describing the data variable-by-variable, and are often referred to as ***descriptive statistics***, or simply as “descriptives.” See Table 2-1 for an example of how this might appear in a research report. The example is based on a subset of data about the Safe At Home instrument developed to assess readiness to change among men entering a treatment program to address their intimate partner violence behavior (Begun et al., 2003). The study participants are described in terms of several variables, with the descriptive statistics reported for each variable, one at a time.

Table 2-1. Univariate descriptive data example for N=520* men entering treatment

	N	%	Range	Mean (M)	sd
Referral Type					
nonvoluntary	424	81.5%	---	---	---
voluntary	96	18.5%	---	---	---
Race / Ethnicity					
African American	215	41.8%	---	---	---
White	205	39.9%	---	---	---
Latino/Hispanic	71	13.8%	---	---	---
Other	23	4.5%	---	---	---
Relationship Status					
Married/living with partner	234	45.6%	---	---	---
Single in a relationship	101	19.7%	---	---	---
Not in a relationship	178	34.8%	---	---	---
Age	---	---	18-72	33.2	8.8668
Time Since IPV Incident	---	---	0-520 weeks	40 weeks	56.4120 weeks
Readiness Score	---	---	.14-7.62	4.60	1.2113
***Note: totals may not equal 520 due to missing data or may not total 100% due to rounding.**					

Back to Basics: Rounding Numbers

Bivariate Analysis. When investigators examine the relationship between two variables, they are conducting ***bivariate analysis*** ("bi" meaning two, as in a bicycle having two wheels). This is contrasted to the descriptive work using univariate analysis (one variable at a time).

For example, investigators might be interested to know if two variables related to each other (correlated or associated). Turning back to the Safe at Home example (Begun et al., 2003), the variable indicating men's readiness to change their intimate partner violence behavior was positively correlated with the variable about their assuming responsibility for the behavior: men with higher readiness to change scores were more likely to assume responsibility for their behavior than were men with lower readiness scores. Looking at this pair of variables together, in relation to each other, is an example of bivariate analysis. Other examples of research questions leading to bivariate analysis include: having a history of child maltreatment victimization associated with school absenteeism in a Swedish study (Hagborg, Berglund, & Fahlke, 2018);

Rounding numbers: The note at the bottom of Table 1 mentions "rounding" as a reason why totals might not equal 100%. Rounding is often used when there are more digits than necessary to convey the important message. For example, the fraction 2/3 is the same as 0.666666 with the "6s" repeating indefinitely when 2 is divided by 3. In social work, social science, and behavioral research, it is seldom necessary to go beyond two or three decimal

> places. We "round up" when the next digit is 5 or higher (5-9) and "round down" when the next digit is less than 5 (0-4). In our example, 0.666666 would be rounded up to 0.67. The decimal version of 1/3, 0.333333 with the "3s" repeating indefinitely, would be rounded down to 0.33. Thus, with slight rounding error being introduced, we might end up with a total slightly lower or higher than 100%.

low empathy being associated with gang membership (as were a number of other variables) in an El Salvador study (Olate, Salas-Wright, & Vaughn, 2012); and mothers' smoking during pregnancy being associated with children's behavioral regulation problems (and other variables) at the age of 12 years (Minnes, eAdd Newt al., 2016).

How Univariate and Bivariate Analyses Are Used: In earlier modules you were introduced to the types of research questions that lend themselves to quantitative approaches (Module 2) and to different kinds of quantitative variables (Module 3). Likewise, the type of data analyses and reports that investigators generate are directly related to the nature of the questions the data answer. Univariate and bivariate analyses are used descriptively. In other words, when the question to be answered calls for a description of a population, social work problem, or social phenomenon, these are useful analyses to conduct and present. Univariate and bivariate analyses help develop a picture of how the numeric values for each quantitative variable are distributed among a sample of study participants. Bivariate analyses are also used to test hypotheses about the relationships between variables—the subject of Chapters 3, 4, and 5 in this module. The next section goes into greater detail as to the nature of descriptive analyses.

Univariate Descriptive Analysis

Here we explore ways that investigators might analyze specific types of variables in order to help answer their research questions. First, our focus is on descriptive statistics: statistics that help describe populations and groups, or the dimensions of social work problems and social phenomena. As we work through these Module 4 materials, you may find that the question "Statistics or sadistics?" asked in the book *Statistics for people who (think they) hate statistics* (Salkind, 2017) changes to the statement "Statistics, not sadistics!"

Descriptive Statistics

Descriptive statistics help develop the picture of a situation. Since the point of variables is that their values vary across individual cases, investigators need to understand the way those values are distributed across a population or in a study sample. The three main features about distribution are of great interest to investigators:

- frequency,
- central tendency, and
- variance.

These three features are reported for one variable at a time (univariate analysis). The descriptive statistics reported about a variable depend on the type of variable: categorical, ordinal, and interval variables that you learned about in Module 3. In Module 3 you also learned about descriptive research: quantitative descriptive research reports the results of descriptive statistical analyses. Descriptive analyses are reported for variables used in quantitative exploratory and explanatory studies as background for understanding the additional (bivariate and other) statistical analyses reported.

Frequency Analysis: Categorical Variables

The first descriptive analyses are concerned with frequency. Frequency is a count of how much, how often, or how many, depending on what the variable was measuring.

Frequency Count. Counting "how many" is among the simplest descriptive statistics. In the Safe At Home example above (Table 2-1), investigators wanted to know how many men participating in the study entered batterer treatment voluntarily and how many were mandated to enter treatment by court order (nonvoluntary clients). As you can see in Table 2-2, the total number of participating clients was 520 (this would be written as N=520, where the capital N reflects total sample size). In addition, you can see that the study included 96 "voluntary" clients and 424 "nonvoluntary" clients (this is sometimes written as n=96 and n=424, where the lower case "n" is about a group within the larger, full set where we used "N"). Not only do we see that more clients entered nonvoluntarily, we can calculate how many more nonvoluntary clients there were: 424 – 96 = 328 more nonvoluntary than voluntary clients participated in the study.

Table 2-2. Frequency report for voluntary and nonvoluntary participants.

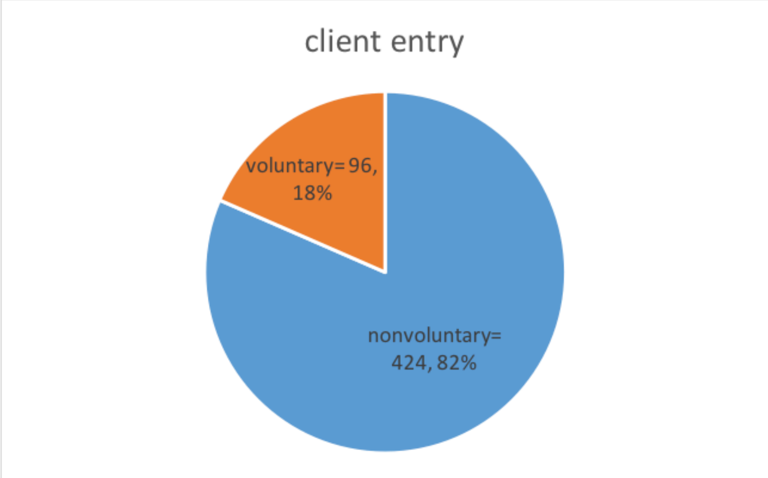
	N	**%**
nonvoluntary	424	81.5%
voluntary	96	18.5%
Total	**520**	**100%**

Interactive Excel Workbook Activities

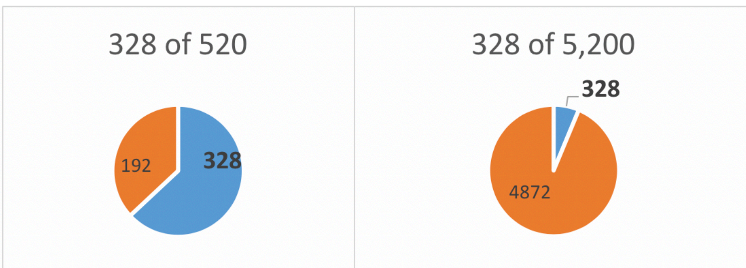
Complete the following Workbook Activity:

- SWK 3401.4-2.1 Frequency Counts

Pie-Chart. The pie-chart is a useful tool for graphically presenting frequency data. It is a quick, easily-interpreted visual device. Here is a sample pie chart depicting the nonvoluntary and voluntary client frequency (and proportion) data described in Table 2-2.



Frequency, Proportion, or Percentage. The investigators in our example not only wished to know how many voluntary and nonvoluntary clients were involved in their study, they also wanted to consider their relative ***proportions***. You might wonder why it matters since they already know how many men were in each group, and that there were 328 more involuntary than voluntary clients in the study. It matters because the investigators wish to evaluate how meaningful that difference might be. A difference of 328 clients is proportionately large in a sample of 520, but the same 328 would be proportionately small among a very large sample such as 5,200.



Back to Basics: Converting Fractions to Decimal Units

Mathematical operations are simpler when fractions are

A proportion is simply a fraction, or ratio, calculated by dividing the number in each group (fraction numerator) by the total number in the whole sample (fraction denominator). To simplify interpretation, the fraction is often converted to percent–the proportion of 100% represented in the fraction.

- For nonvoluntary clients, this would be computed as: (424/520) x 100%, which comes out to be (.815) x 100%, or

81.5%.

- For voluntary clients, this would be computed as: (96/520) x 100%, which comes out to be (.185) x 100%, or 18.5%.

The investigators provided these ***percentages*** in Table 2-2; the percentages total to 100% (81.5% + 18.5% = 100%). You can see that the two groups were not equivalent in actual size, relative proportion, or percentages of the total sample. In Chapters 4 and 5 you will learn about statistical analysis approaches that allow investigators to test if the size of the observed difference is statistically significant.

expressed in decimals. To make the conversion, think about what the fraction stands for: a top number (numerator) over a bottom number (denominator) means to divide the numerator by the denominator. Example: 3/4 means divide 3 by 4, so 3/4 = .75

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-2.2 Proportion and Percentage

Univariate Central Tendency Analysis: Numeric Variables

Knowing the frequency with which certain categories or levels of a categorical variable are reported in the data is useful information, as witnessed above. Frequency counts are less useful in understanding data for a numeric variable, however. For example, in the Safe At Home study, it would be difficult to interpret information about the participants' ages: 3 were age 18, 7 were age 19, 17 were age 20, 8 were age 21, 14 were age 22, and so forth for all 520 individuals between 18 to 72 years. A more useful way of looking at these numeric (interval type) data is to consider how the study participants' values for age were distributed and how any individual's values for age might differ from the rest of the group. Ideally, we would like to see a meaningful summary of the data.

The concept of ***central tendency*** refers to a single value that summarizes numeric data by characterizing the "center" values of a data set—a single value around which individuals' values tend to cluster. Central tendency helps make sense of the variability observed among participants on a numeric variable. The most commonly used central tendency ways to summarize data are:

- mean,
- median, and
- mode.

For example, within a classroom of public school students we might see many different values for the variable "number of days absent" during a school year. Most states in the U.S. require 180 student instruction days in the school year. Let's consider a hypothetical example (see Table 2-3) where hypothetically:

- we have a class of 28 students
- the number of days absent for each student ranges from 0 to 49 out of the year's 180 days
- the total number of absent days for the class members combined was 194.

Table 2-3: Number of days absent for each student in (hypothetical) class.

student	# days	student	# days
1	2	16	6
2	4	17	4
3	0	18	49
4	0	19	0
5	26	20	3
6	0	21	0
7	15	22	4
8	3	23	10
9	9	24	1
10	12	25	2
11	1	26	0
12	1	27	2
13	3	28	12
14	19	---	---
15	6	total	194

Our goal is to find a single central tendency value that helps represent the group on this days absent variable; summarizing the variable without having to rely on all 28 values at once. Let's look at each of the common central tendency options.

Back to Basics: Computing and Average or Mean Value

Mean. The first central tendency indicator we consider is the mean. You might recognize this as numeric

Step 1. Add up the values for each member of the group on the variable of interest
Step 2. Divide that total by the number of values there were in the group (how many individual members contributed to that total).
Step 3. The result is the "average" value for the group, also called the "mean" value (written as a capital, italicized M).

average; average and mean for a set of values are the same thing. The mean is computed by adding up all the values, then dividing by the actual number of values. In our example, we would first find the sum of the number of days absent across all the students together (194), then divide that total by the number of students contributing to that total (N=28). This would be the mean value for our students on the variable for number of days absent: 194 divided by 28 (194/28) = 6.93 with rounding. We would report this as M=6.93.

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-2.3 Computing a Mean

Median. The median is the next common central tendency feature, and it can be thought of as the half-way point in a distribution of values. In other words, it is the point where half of the values are higher and half of the values are lower. Another name for the median is the 50th percentile—since 50% is halfway to 100%. In other words, the 50th percentile is the value where 50% (half) of the scores are lower and the other 50% (other half) are higher. In our classroom absenteeism example, the median value is 3 days: three is the point where half of the students had more and half had fewer days absent. We can easily compute that half of the 28 students is 14, so we identify the median point as where the lowest 14 values are included and all of the higher values are excluded—14 students had scores of 3 or less. We would report this information as: *Mdn*=3.

Mode. The last common measure of central tendency to consider is the mode. Sometimes it is useful to know the most common value for a variable. This is simply a frequency count—how many times did each value appear in our data. For our student absenteeism data, the most common value was 0—six of our students did not miss any days of school (see Table 2-4). This would be written as Mode=0.

Table 2-4. Frequency of each value for days absent example.

# days absent	# of students with that value	# days absent	# of students with that value
0	6	10	1
1	3	12	2
2	3	15	1
3	3	19	1
4	3	26	1
6	2	49	1
9	1	total	28

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-2.4 Computing Descriptive Statistics (Mean, Median, Standard Deviation)

Comparing Mean, Median, and Mode. Why wouldn't we always use the mean to describe the variable of interest? The answer is demonstrated in our student absenteeism example: even a few extreme values (sometimes called ***outliers***) can seriously distort or skew the picture. The most common value (mode) was 0 but we had a couple of students with some really high values compared to the others: 26 and 49. When we add those two extreme values into the total number of absent days to compute the group mean, they have a big impact on the total (194): the mean for the 28 students was 6.93 days absent. Without those two extreme values, the total would have been 119, making the mean for the 26 remaining students (28 – 2 = 26) only 4.6 days (119/26=4.6). This mean of 4.6 days is considerably less than the mean of 6.93 days we observed with these two extreme students included. Thus, you can see that the two extreme values had a powerful skewing impact on the mean.

As a real-world example, consider data about the 2017 incomes (salary plus bonuses) for 1,999 NFL professional football players. The range was $27,353 to $23,943,600—a difference between 5-figure and 8-figure incomes! In 2017:

- mean player income was $1,489,042 (we would report this as *M*=$1,489,042),
- median income was $687,500 (*Mdn*=$687,500), and
- mode income was $465,000.

The mean and median incomes are very different, and the mean differs markedly from the mode, as well. This NFL player income example demonstrates what data look like when great disparities exist between individuals across a group or population—because the mean is highly sensitive to extreme values, there exist large differences between the mean, the median, and the mode. This is an important social work and social justice issue in poverty, health, incarceration, and other rates where large disparities between population groups exist.

In a relatively small sample (like our 28-student classroom), one or two extreme values made a big difference between the mean and median. Think about the mean income of people in a room where everyone is a social work student learning about social justice issues. Then, think about what would happen to that mean if Bill Gates arrived as a guest speaker invited to talk about social justice on a global scale. Just that one person would drastically change the summary picture based on mean income. However, the median would not really shift much; Bill Gates would be only one person out of the total number of people in the room; the halfway point is not affected much by one person being extreme, since the median is a frequency count of people on either side of the central midpoint value regardless of how far out the values are spread. The mode also is not affected since Bill Gates' one extreme value is certainly not going to be the most common value (otherwise it would not be extreme). This all leads to a discussion of the importance in understanding the nature of variance as it relates to distribution of values for a variable.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=429

Univariate Data Distribution: Numeric Variables

Central tendency is about how the set of values for a specific variable are similar, how they cluster around a mean, median, or mode value. However, investigators are also interested in how individual values differ. This is a curiosity about how those values are distributed or spread across a population or sample: the ***range*** and ***variance*** of the values. These elements are explained below, along with the more commonly reported ***standard deviation***.

Range. Turning once again to our school absenteeism example, we can see that there exist very great differences in the numbers of days absent among the students in this class. The minimum value observed was 0 and the maximum value observed was 49. The range of values in a set of data is computed as the highest value minus the lowest value. In our example, the range would be 49 – 0 = 49 days absent. It is just a coincidence that the range in our example equals the highest value: this is only because our lowest value was 0 (if the lowest value was different from zero, the range would be different from the highest value).

The range has its greatest meaning when we understand the context in which it exists: a 49-point range is large for the possible 180 days but would not be if the context for possible values was 1,800 instead. Therefore, it becomes helpful to consider the range in ratio or percent absent days, giving the range greater meaning. This is calculated as the observed value divided by the possible value for the ratio, then multiplying by 100% to determine percent.

- Missing 0 of 180 days possible is a division (ratio) problem: (0/180) x 100% = 0 x 100% = 0% days absent;
- Missing 49 of 180 days possible is computed as (49/180) x 100% = 27.2 x 100% = 27.2% days absent.
- Thus, the range is 0% to 27.2% days absent.

One statistical approach to assessing this variability in days absent is our next topic: understanding variance.

Variance. One way of thinking about variability in the data is to compare each individual's score to the mean for the group—how far from that central, summary value each individual's value falls. When individuals' scores are clustered very close to the mean, the variance is small; when individuals' scores are spread far from the mean, the variance is large.

This sequence of steps describes how variance is computed in a sample.

Back to Basics: Subtracting Negative Numbers

Computing the range can be a little tricky if the lowest value is a negative number and the highest value is a positive number–unless you remember that *subtracting a negative* number is the same as *adding it as a positive* number (the two negative signs cancel each other out). For example: if the lowest value is -4 and the highest value is 26, then the range would be computed as: (26) – (-4) = (26) + (4) = 30

- Step 1. Compute the mean of all the values together (overall mean).
- Step 2. Compute the distance of each individual value from that mean—this becomes a subtraction problem where the mean is subtracted from each individual score, showing the "difference" between the individual score and the group mean.
- Step 3. Multiply each of those distance values by itself (square the difference value). Why bother with this step, you might wonder? This is done to resolve potential problems with negative numbers from our earlier subtraction step. Some values were smaller than the mean and some were larger than the mean—when we subtract the mean from those smaller values, we end up with a negative number. In our next step (step 4) we are going to add the computed distances together. If we add in negative numbers, we are essentially subtracting values from the total which does not give an accurate picture of the total distances from the mean. To solve this negative number problem, we capitalize on the fact that multiplying two negative numbers together gives a positive value. Therefore, if we multiply a negative value by itself (square a negative difference score) we get a positive distance value (the distance squared).
- Step 4. This step is where we add up the squared distance values to get a single total (sum) of the squared distances.
- Step 5. Divide this sum by the number of values in the data set

minus 1—one less than the number of contributions to the variance calculation. Why is this not just divided by the number of values instead, you might wonder? It has to do with the fact that we are computing variance of a sample, not the whole population. Without this adjustment, the computed sample variance would not be quite the same as the population variance—more about this distinction between sample and population parameters later, when we talk about statistical analyses and the conclusions drawn from the statistical calculations.

Back to Basics: Multiplying With Negative Numbers

Remember that multiplying two numbers with the same sign results in a positive sign number, whether those two signs are both positive or both negative. So, when multiplying two negative numbers, the negative x negative signs cancel each other out.

positive x positive = positive

positive x negative = negative

negative x negative = positive

Here is what it looks like with our school absenteeism example.

- Step 1. We know that the mean is 6.93 from our earlier discussion about means.
- Step 2. Our first student missed 2 days, so the distance or difference score is:
  - (2 – 6.93) = -4.93 (a negative number).
- Step 3. Now we compute those differences for every one of our 28 students. Once we have done this, we compute the square of each difference by multiplying it by itself. For our first student, this is:
  - $(-4.93)^2 = (-4.93) \times (-4.93) = 24.3$ (rounding)
- Step 4. After computing each student's squared difference, we add them up (sum or total).
- Step 5. We take that total value and divide it by one less than the number of students contributing data: (28-1) = 27. The value we end up with for variance in the class is 109.25 (see Table 2-5).

Table 2-5. Computing variance in student absenteeism example: (2949.86)/(27)=109.25

student	# days	distance from mean	(difference)^2	student	# days	distance from mean	(difference)^2
1	2	-4.93	24.30	16	5	-.93	.86
2	4	-2.93	8.58	17	4	-2.93	8.58
3	0	-6.93	48.02	18	49	42.07	1769.88
4	0	-6.93	48.02	19	0	-6.93	48.02
5	26	19.07	363.66	20	3	-3.93	15.44
6	0	-6.93	48.02	21	0	-6.93	48.02
7	15	6.07	65.12	22	4	-2.93	8.58
8	3	-3.93	15.44	23	10	3.07	9.42
9	9	2.07	4.28	24	1	-5.93	35.16
10	12	5.07	25.70	25	2	-4.93	24.30
11	1	-5.93	35.16	26	0	-6.93	48.02
12	1	-5.93	35.16	27	2	-4.93	24.30
13	3	-3.93	15.44	28	12	5.07	25.70
14	19	12.07	145.68	------	------	------	------
15	6	-.93	.86	total	194	------	2949.86

Standard Deviation

Working with variances can be unwieldy, for two reasons. A practical reason is that the numbers can become very large, making them difficult to work with. Another practical reason is that variance is a bit difficult to interpret. Remember when we calculated the square of the differences in days absent minus the mean? We started with the unit of analysis being days absent. By squaring the values, our units are now the square of the difference in days absent. That just is not very intuitive. Instead, we find it easier to convert the figure back to the unit of days of absent. To do this is straightforward math: we take the square root of the variance of the squared distances. This gets both the value and the units back to a more interpretable, practical place.

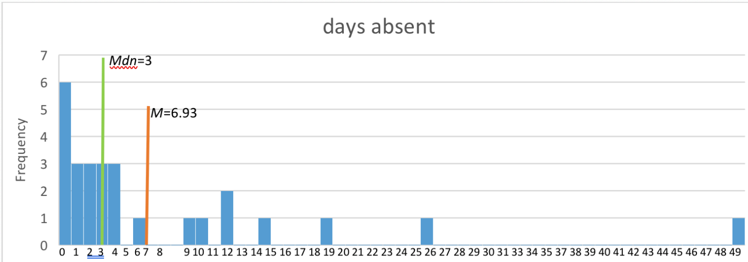
- Square root of 109.254 is 10.452 (using the square root calculator on www.math.com)
- Cross-check this yourself: multiply 10.452 x 10.452 and see what result you get—it is slightly different in the third decimal place because of rounding.
- Thus, the standard deviation in our example is reported as: *sd*=10.45 days absent.

The way to think about standard deviation is that the greater a standard of deviation is, the more variable the values are—the more they spread rather than cluster around the mean. This is an important aspect of understanding the distribution of values on the variable of interest. A graph is also very helpful.

Graphing Numeric Variables

The way the numeric values of a variable are distributed can be visualized using what is called a ***histogram***. The histogram is a graph where the *x*-axis represents the set of observed values for a specific variable (a univariate analysis), and the *y*-axis represents how many times each of those values was observed in the data (see Figure 2-1). You can see from the height of the bars that the frequencies begin to taper off after about 4 days absent and how far from the rest of the class those few extreme students are situated.

Figure 2-1. Histogram depicting distribution of days absent data, mean, and median.



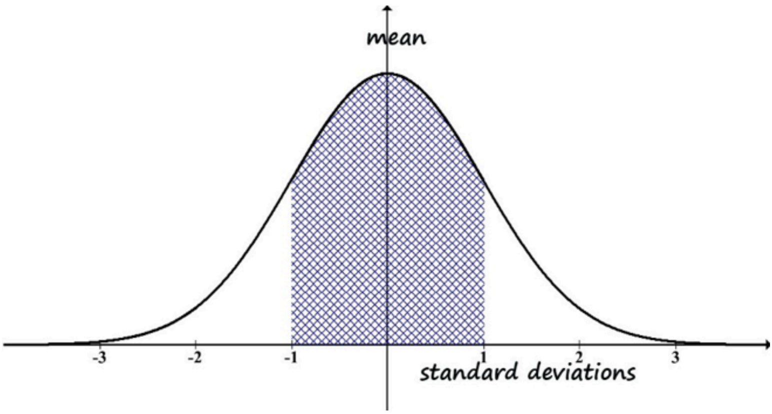
Another clue about the degree of spread or variability is the relative size of the sample's standard deviation compared to the sample mean. In our example, the *sd*=10.45 is greater than the M=6.93. This indicates that a great deal of spread exists in the data. This observation relates to understanding the characteristics of normal distribution.

NORMAL DISTRIBUTION

You have probably been exposed to at least the general concept of the bell curve or normal curve at some point in your education. A ***normal distribution*** curve is sometimes called a "bell" curve because of its shape: high in the middle, low in the "tails," and even on both sides of center. Bell-shaped curves do not necessarily meet all criteria of a normal distribution curve—you will see examples in this chapter. We focus on the true normal curve at this point, a fundamental concept related to many of the statistics used in social work, social science, and behavioral research.

Our in-depth exploration of its importance in understanding data begins with the picture of a normal distribution curve (Figure 2-2). The type of variable mapped onto a normal distribution curve is an interval or continuous variable—the units are equal intervals across a continuum.

Figure 2-2. Generic normal distribution curve (histogram).



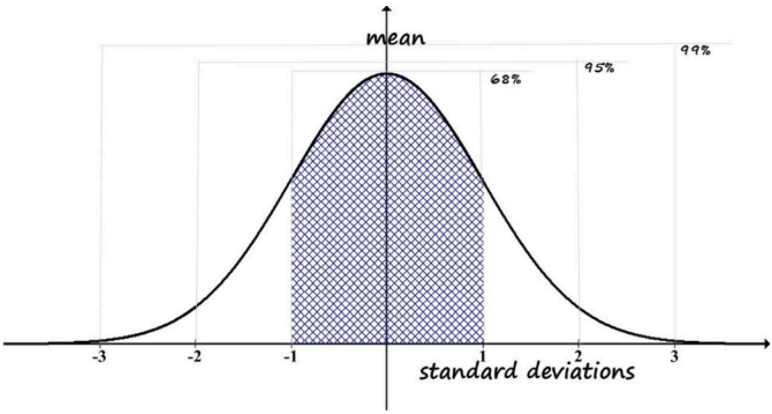
In a normal distribution curve, the values are arranged around the mean in very specific ways.

- The mean is the center of the curve.
- The observed data values are distributed equally on either side of the mean—the right and left sides are mirror images. This goes along with the mean being in the center; the right and left ends are called “tails” of the distribution curve.
- The median is equal to the mean. This also goes along with the mean being in the center of the distribution.
- The set of observed data values (the frequency each value is observed) are distributed in the following way, with regard to the standard deviation value (see figure 2-3):
    - 68% are within one standard deviation of the mean (34% on each side of the mean)—see the grid-marked area under the curve in Figure 2-3.
    - Combining all values observed for 2 standard deviations either side of the mean will include a total of 95% of the

values (all of those within 1 standard deviation plus those from 1 to 2 standard deviations from the mean).

- Combining all values within 3 standard deviations of the mean accounts for 99% of the values.

Figure 2-3. Proportion of values under the normal curve by standard deviations from the mean.

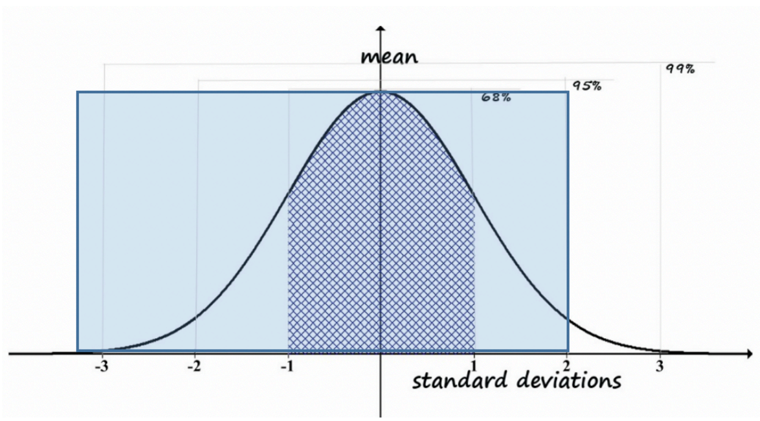


Percentile Scores. You may wonder how this relates to information such as someone having a 95th percentile score on some measure. Look at the right-side tail, out at 2 standard deviations from the mean. You can see that very few scores are way out in the tail beyond that point—in other words, it is rare to be more than 2 standard deviations away from the mean, and exceedingly rare to be more than 3 standard deviations away.

We saw that 95% of the scores fall between -2 and +2 standard deviations. This, however, is NOT the 95th percentile score. The reason it is not has to do with the definition of the 95th percentile: it is the point with 95% of the observed scores falling below and only

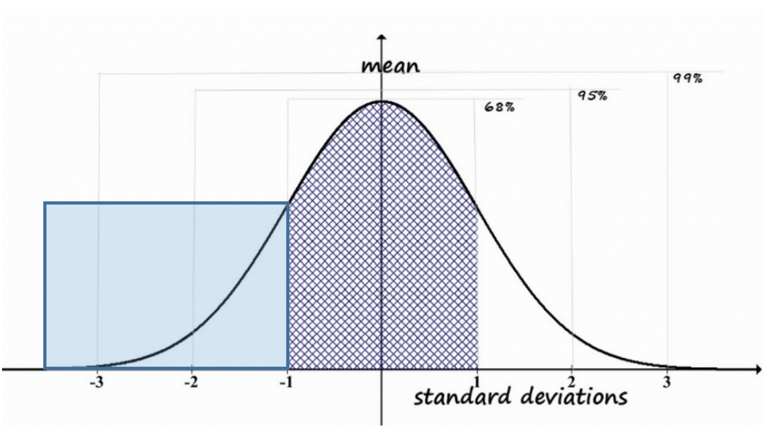
5% of the observed scores falling above. In this case, at 2 standard deviation above the mean (+2) we have 95% of scores in the -2 to +2 standard deviation range, which means 5% of scores are outside the range (100% – 95% = 5%). But, these are not all on one tail; there are those scores in the -2 to -3 standard deviation range to contend with. Those scores are also below the point we identified for being 2 standard deviations above the mean. So, using the +2-standard deviation value as our decision criterion, we would actually have half of the left-over 5% (5%/2 = 2.5%), plus the original 95%. This means 2 standard deviations above the mean is the 97.5 percentile (95% + 2.5% = 97.5%). The trick is to keep in mind which standard deviation range is relevant to the question being asked—is it about a range around the mean or is it about a range up to a standard deviation criterion value (see Figure 2-4)?

Figure 2-4. Computation of percentile for +2 SD.



Here is another example to consider. Imagine that policy makers want to provide services to individuals whose IQ scores are equal to or less than 1 standard deviation below the mean (this would be written as: ≤ -1 SD). This decision concerns one tail of the distribution—those at the far left (bottom) of the range, not the mirror tail on the right (top) of the range. If you wanted to determine what percentile would be covered, you would start with figuring out the total percent that fall outside of the -1 to +1 standard deviation range. That would be 100% minus the 68% that ARE in the range: 100% – 68% = 32% outside of the range. We also do not want to include those individuals in the range greater than 1 standard deviation above the mean (> +1 SD). They account for half of the number who are outside the -1 to +1 standard deviation range, which was 32%. So, we are going to subtract half of 32% (32%/2 is 16%) and we are left with only 16% of individuals falling below 1 standard deviation being eligible for the services (see Figure 2-5).

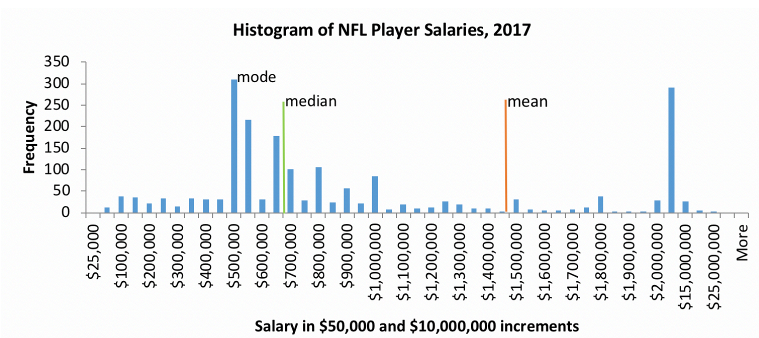
Figure 2-5. Computation of percentile for -1 SD.



In fact, this scenario became an issue in one state during the 1970s. IQ scores are presumed to distribute on a normal curve with a mean of 100 across the population, and each standard deviation defined by 26 points on the test. As a cost-saving plan, the state determined that services to persons with intellectual developmental disabilities would be offered to individuals whose IQ scores were 68 or below. Previously, the criterion value had been 70. Rather than having "cured" individuals whose scores were between 68 and 70, the state simply ceased providing services to them—several thousand individuals no longer met criteria.

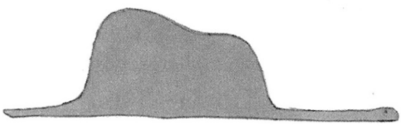
Next, let's consider the real-world example of NFL player salaries mentioned earlier. Looking at the histogram, it does not seem to be very normally distributed. First, the mean does not seem to be in the middle, the bars to the right and left of the mean are not mirror images, and the median is very much lower than the mean—they are not even close to equal (see Figure 2-6). These are ways in which distribution on a variable fails to be normal.

Figure 2-6. Histogram of NFL player salaries.



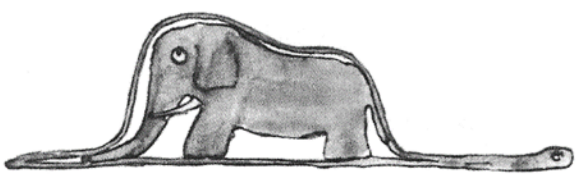
Bimodal Curve. Another way in which distributions are non-normal is when we see a bimodal curve. The normal, or "bell" curve, has only one peak or hump in the data. Sometimes, we get a curve that has two humps. If there had been more players in the 400,000 to 600,000 range we might have seen a second "hump" after the first one at around 500,000 to 600,000.

This scenario is "scary" in statistical terms, because it means that our assumption of normal distribution, the basis of many statistical analyses, has been violated. If you ever read *The Little Prince* (de Saint-Exupéry, 1943), you might understand the scary picture better. The Little Prince shows his scary picture (Drawing Number one) to the adults, who say a hat is not scary at all.



© de Saint-Exupéry

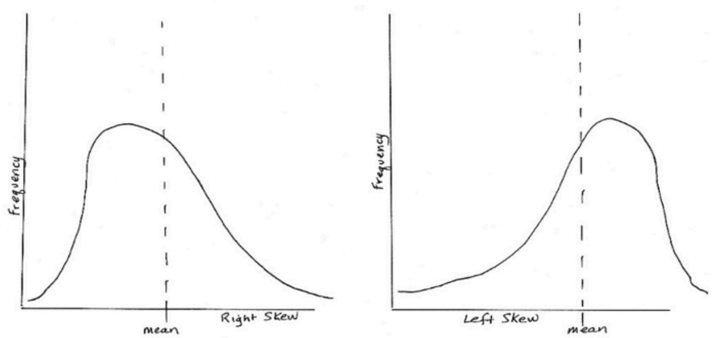
But, the picture is scary because it is not a picture of a hat, it is a picture of a snake who ate an elephant.



© de Saint-Exupéry

To statisticians, a bimodal curve is a scary picture—it has two "humps." In other words, there exists a second peak in the data and this needs to be explored. As an example for explaining the bimodal curve in The Little Prince's Drawing Number One, perhaps the curve is showing salaries where the left peak (lower salaries) is for women and the right peak (higher salaries) is for men; or, it could be about salary differences for two different raceial groups; or, it could depict an inequity or disparity in the frequency for some other variable.

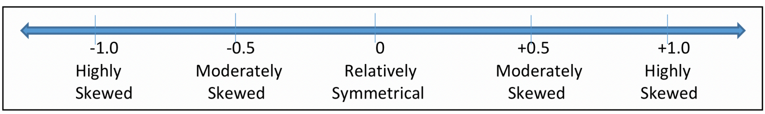
Skew and Kurtosis. Two other ways that a variable might be distributed in a non-normal way is that it is skewed to one side or the other in relation to the mean, and/or the peak of the bell is either too tall or too flat to represent those probabilities we mentioned earlier (e.g., 68% being within one standard deviation of the mean). The first of these is called ***skew*** and the second is called ***kurtosis.*** A curve with a lot of skew fails to look normal because the curve is no longer symmetrical on both sides of the mean. This is how skewed curves might look:



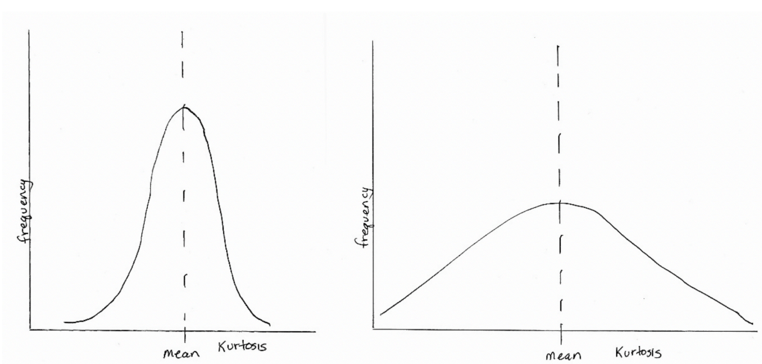
Univariate statistics often report a value for skewness. While the computer can do the calculations, the trick lies in knowing how to interpret the values. Here are some rules of thumb for interpreting skewness values:

- If the value for skewness is less than -1.0 or greater than +1.0, far from zero, the distribution is highly skewed (asymmetrical or non-normally distributed).
- If the value for skewness is between -1.0 and -0.5, or it is between +.5 and +1.0, then the distribution is considered to be moderately skewed (asymmetrical).
- If the value for skewness is between -0.5 and +0.5 (including zero), the distribution is considered relatively symmetrical.

This diagram might help you visualize what these words mean:



Kurtosis, on the other hand, has to do with how much of the distribution is sitting out in the tail or tails instead of being in the center "peak" area compared to a normal distribution. On one hand, there may be too few cases in the tails (clustered too much in the "peak"); on the other hand, there may be too many cases in the tails (spread away from the "peak"). This is how curves with kurtosis might look:



Kurtosis values are a bit more complicated to interpret than skewness values. When we learn about statistical tests of significance we can revisit the issue. In general, kurtosis is considered to be problematic when the value falls outside of the range from -7.0 to +7.0 (for social work, social science, and behavioral research). Some scholars argue for applying criteria with a tighter range, from -2.0 to +2.0.

The website called Seeking Alpha (α, https://seekingalpha.com/article/2100123-invest-like-taleb-why-skewness-matters) explains the importance of skew and kurtosis by using an example from Nassim Taleb (2012):

> *Suppose you placed your grandmother in a room with an average temperature of 70 degrees. For the first hour, the temperature will be -10. For the second hour, it will be 140 degrees. In this case, it looks as though you will end up with no grandmother, a funeral, and possibly an inheritance.*

(Note: for our purposes, degrees Celsius were translated into degrees Fahrenheit to make the point from the original example clear.) The website goes on to state:

> *While this example is rather extreme, it does show one important principle that is highly relevant…It is dispersion around the mean that matters, not necessarily the mean itself.*

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=429

—

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-2.5 Understanding Skew and Kurtosis

Understanding Graphical Representations: Categorical Variables

There exist at least 3 different ways of presenting the very same information about categorical variables: as text (words) only, as a table of information, or in graphs/figures. We can compare these approaches using a real-world data example: suicide rates by age group, as reported by the Centers for Disease Control and Prevention (CDC, 2017). A "rate" is similar to a frequency: rather than being an actual count (frequency), the frequencies are converted into population proportions. So, we often see a rate described as some number per 1,000 (or per 10,000, or per 100,000) of the population, instead of seeing the actual number of individuals.

Text: During 2016, middle-aged adults between the ages of 45 and 54 represented the group with the highest suicide rate (19.7 per 100,000 persons in the population). This was closely followed by individuals aged 85 or older (19.0 per 100,000), adults aged 55 to 64 (18.7), and older adults in the 75 to 84 age range (18.2 per 100,000). Next were adults in the 35 to 44 age group (17.4 per 100,000), adults in the 65 to 74 age range (16.9 per 100,000), and adults aged 25 to 34 (16.5 per 100,000). The suicide rate among adolescents and young/emerging adults aged 15 to 24 was somewhat lower (13.2 per 100,000), and the very lowest rate occurred among persons under the age of 15 (1.1 per 100,000). The rate for all ages combined was 13.42 per 100,000 persons in the population (age-adjusted).

Table: Table 2-6 presents the same information described in the text above, but much more succinctly.

Table 2-6. Suicide rate per 100,000 by age group during 2016.

Age Group	Rate per 100,000
Under 15	1.1
15 – 24	13.2
25 – 34	16.5
35 – 44	17.4
45 – 54	19.7
55 – 64	18.7
65 – 74	16.9
75 – 84	18.2
85 or older	19.0
overall age-adjusted rate	13.42

Bar Charts: The same information is presented once again in Figure 2-6, but graphically this time, where you can visually compare the height of the bars representing each age group.

Figure 2-6. Suicide rate bar chart.

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=429

Chapter Summary

The emphasis in this chapter surrounded univariate descriptive statistics. You learned about describing frequency and percentage results. Then, you learned about the "Central Tendency" statistics most commonly used in social work, social, and behavioral research: mean, median, and mode. You also learned about the statistics concerned with how data are distributed: variance and standard deviation. Related to the issue of data distribution, you learned about normal distribution (and histograms), as well as the nature of skew and kurtosis in relation to non-normal distribution. Finally, you witnessed the power of graphs and table to present univariate statistics information compared to written text. In the next chapter, we begin to explore bivariate analyses.

Module 4 Chapter 3: Univariate Inferential Statistical Analysis

In this chapter we introduce***inferential statistics*** strategies, beginning with univariate analyses and expanding to bivariate analyses. Why these statistical approaches are called "inferential" may seem, on the surface, to be a simple question. Simple or not, it is an important question. The word "inferential" is used because investigators and statisticians use data about samples to make inferences about the population-as-a-whole. Working with sample data is necessary when it is not feasible to measure or observe every single individual in the population. The idea with inferential statistics is that investigators can estimate the degree of confidence they can reasonably place in their conclusions about a population, conclusions based on data from a sample.
In this chapter you will learn about:

- the importance of distinguishing between populations and samples;
- confidence intervals for univariate population parameters based on sample statistics;
- degrees of freedom;
- distribution of the *t*-statistic.

Populations and Samples

Our prior discussions about study designs, variables, measurement, and study participants (Module 3) addressed how investigators gather data from observations about a sample drawn from a population. Depending on the nature of the design and sampling strategies applied, data about the sample will be either more or less representative of what is going on for the whole population. To understand the logic of hypothesis testing with statistics, it is important to understand the distinction between what we know about the sample and what we are guessing about the population based on the sample. This brings us to understanding the difference between the statistics we explored in Chapter 2 (mean, median, mode, variance, and standard deviation) and their corresponding "guessed" values in the population: the difference between ***sample statistics*** and ***population parameters***.

Consider the example from Chapter 2 about student absenteeism. We calculated the mean number of days absent for the entire classroom to be 6.93 days. This is the population parameter for the entire classroom—it was feasible to measure the absent days for everyone in the class. Now consider the possibility that an investigator wished to use that class as a sample to represent absenteeism for the entire school or school district without having to measure absenteeism for everyone in the entire school or district population. In this case, the class is no longer a population, it is a sample. Data about the sample are used to make inferences about the population from which the sample is drawn.

Statisticians like to use different symbols for population parameters and sample statistics, just to keep straight which one they are working with at any time; the symbols used in statistical computation differ somewhat from those used in APA-style reporting of statistics, as well. For example, sample mean is M in APA-style reporting and $\bar{X}$ in statistical formulas and computations. Here are some commonly used symbols compared:

description	**sample statistic**	**pronunciation**	**population parameter**	**pronunciation**
mean	$\bar{X}$ (or M)	x-bar	μ	mu
standard deviation	s (or sd)	s	σ	sigma
variance	s^2	s-squared	σ^2	sigma squared

It would be useful to have a way to decide how well that sample's statistics estimate the larger population's parameters. Sampling from the population always introduces the possibility that the sample data are not a perfect representation of the population. In other words, investigators need to know the degree of error encountered in using the sample statistics to estimate the population parameters, and what degree of confidence can be placed on the population estimates based on the sample statistics. This leads to an exploration of the ***confidence interval (CI)*** for univariate statistics.

Confidence Interval (CI) for Univariate Statistics

Although no one can ever know the true population parameter for the mean on a specific variable without measuring every member of the population, a range of values can be specified within which one can be reasonably confident the actual population parameter lies (though not 100% certain). This range of "confidence" values can be calculated if certain information about the sample is known, and with awareness about certain assumptions being made. These include:

- knowing the sample mean ($\bar{X}$, or M), since this will be used as an estimate for the population mean;
- knowing the sample standard deviation (designated as s, or *sd*);
- deciding the level of confidence desired (commonly, 95% confidence is used)
- an assumption that the sample was randomly drawn from a population normally distributed on the variable of interest;
- a formula for calculating the range of values for the confidence interval (more about computing the *t*-statistic below);
- information about the distribution of *t*-values to assess the calculated values (more about this below).

t-Statistics. Based on probability theory (beyond the scope of this course), statisticians developed an understanding of what they called the *t*-distribution. The *t*-distribution is very similar to a normal curve, but the curve differs slightly depending on the number of data points (the sample size, or N). The *t*-distribution is relevant when sample sizes are relatively small; with an infinite number of observations, or data points (N=∞, or N=infinity), the *t*-distribution becomes identical to the normal distribution. This is another way of saying that the entire possible population is normally distributed for that variable (one of our assumptions in the list above). Statisticians use a set of values from a standard table to evaluate the ***t-statistic*** computed from the formula that we will examine soon. However, to use the standard table of *t*-values, the number of degrees of freedom must be known.

Degrees of Freedom (*df*). As previously noted, *t*-values are dependent on the number of observations in a sample (n). The goal in the present situation is to estimate the mean for a population based on a sample's mean. The number of ***degrees of freedom*** involved in calculating that estimate concerns the number of values that are free to vary within the calculation. Starting with the sample mean as our estimate of the population mean, once we know the sample mean and all of the values except one, we know what that last value would be—it is the only possible value left once the other values have been identified or locked in. Therefore, the only degrees of freedom we have would be represented as (N-1) degrees of freedom (abbreviated as *df*=N-1).

A simplified example comes from working Suduku-type puzzles. The goal in this simplified Suduku example is to fill in the blank cells with numbers 1, 2, or 3 so that no number appears twice in the same row or column and each number is used once. There is only one possible answer for the remaining empty cell that meets the requirements of the puzzle: the number 2. When the puzzle was entirely empty, many degrees of freedom for that cell existed—any of the allowed numbers could be placed in that empty cell. Once the values were known for the other cells, however, only one possibility remained, leaving no degrees of freedom. Thus, that cell started out with the number of degrees of freedom for the entire range of possible values (3), but once the other 8 values were filled in, there were zero degrees of freedom remaining—only one answer would meet all criteria.

1		3
2	3	1
3	1	2

In a different example, imagine we are trying to achieve a value of 5 by adding together 3 non-negative, non-decimal (whole) numbers; we are going to use exactly three numbers to add up to that value of 5. Once we know any two values from 0 – 5, we automatically know the last, third number. In other words, with three possible values to add together, our degrees of freedom are limited to zero once those other two values are known.

0	+	1	+	4	=	5
0	+	2	+	3	=	5
0	+	3	+	2	=	5
0	+	4	+	1	=	5
0	+	5	+	0	=	5
1	+	1	+	3	=	5
1	+	2	+	2	=	5
1	+	3	+	1	=	5
1	+	4	+	0	=	5
2	+	2	+	1	=	5
2	+	3	+	0	=	5

Starting out, before any values are filled in, for any one cell and if we were going to add together three non-negative numbers to get 5, our degrees of freedom would be:

$$df = (N\text{-}1) \text{ or } df = (3\text{-}1), \text{ which is } df = 2.$$

Table of *t*-distribution. Now we are ready to find the comparison criterion *t*-value using the standard table mentioned earlier. Table 3-1 present an excerpt of such a table, with only the values for a two-tailed *t*-distribution at 95% confidence levels. You would read this table by going down the 95% confidence column to the row where the number of degrees of freedom for the sample are presented on the left. The value in that cell would be the *t*-value against which the computed *t*-statistic would be compared to draw inferential conclusions.

Table 3-1 presents the values for a two-tailed *t*-distribution at 95% confidence, depending on the number of observations minus 1 (degrees of freedom). As you can see, the greater the sample size, the greater the number of degrees of freedom, and the greater the number of degrees of freedom, the closer to the population normal distribution value of 1.960 we come. At *df*=1, we are very far away from that value (12.71 is much greater than 1.960), while at *df*=30, we are much closer (2.042 is close to 1.960). A full *t*-table would also include values for other levels of confidence, such as the 90% and 99% levels.

Table 3-1. Values at 95% confidence, for two-tailed *t*-distribution, by *df*.

df	*t*-value		*df*	*t*-value		*df*	*t*-value		*df*	*t*-value
1	12.71		11	2.201		21	2.080		40	2.021
2	4.303		12	2.179		22	2.074		50	2.009
3	3.182		13	2.160		23	2.069		60	2.000
4	2.776		14	2.145		24	2.064		80	1.990
5	2.571		15	2.131		25	2.060		100	1.984
6	2.447		16	2.120		26	2.056		120	1.980
7	2.365		17	2.110		27	2.052		∞	1.960
8	2.306		18	2.101		28	2.048			
9	2.262		19	2.093		29	2.045			
10	2.228		20	2.086		30	2.042			

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=430

Calculating a Confidence Interval for the Mean.

This work has led up to being able to calculate the confidence interval for the population parameter mean from the sample statistic mean. At this point, we simply need to plug the proper values into the appropriate formula and compute the answer. In application, we make two computations using the formula: once to identify the top/highest value in the confidence interval's range and again to identify the bottom/lowest value in the confidence interval's range. As a result, we will have identified the range of values based on our sample statistics where, with 95% confidence, the true (unknowable) population mean lies. We will have inferred the population mean from the sample statistics.

Consider what the formula asks us to do. To compute the 95% confidence interval ($CI_{95\%}$) highest value, we start with the sample mean ($\bar{X}$), then add a computed value that takes into consideration our sample size and the degree of variation observed in the sample values: that computed value is calculated from our *t*-value (located in the table using our degrees of freedom) multiplied with our standard deviation (*sd*) which has been divided by the square root of our sample size ($\sqrt{n}$)–see the formula written below. What that division does is change the overall standard deviation figure into an average degree of variation around the mean—this refers to the formula we used in Chapter 2 to calculate variation and standard deviation in the first place.

Finally, we compute our lowest value in the confidence interval range by doing the same thing but subtracting the computed value instead of adding it. Since we already know the values for everything to the right of $\bar{X}$ in the formula, we do not have to recalculate it—this time we simply subtract it from $\bar{X}$ instead of adding.

In mathematical terms, this set of steps looks like this:

$$\underset{\text{(highest value)}}{CI_{95\%}} = \bar{X} + [(t_{\text{for 95\% confidence}}) * (sd/\sqrt{n})] \quad \text{and} \quad \underset{\text{(lowest value)}}{CI_{95\%}} = \bar{X} - [(t_{\text{for 95\% confidence}}) * (sd/\sqrt{n})]$$

Working an Example. This looks a lot more complicated than it really is, so let's work a simple example. Imagine that the 28 students in our school absenteeism study were drawn randomly from the population of students in a particular school district, and that we safely can assume normal distribution of the days absent across the population of students in an entire school district. From our earlier work, we know the following information:

- $\bar{X}$=6.93 days absent
- *sd*=10.452
- n=28

- *df*=(28-1)=27
- comparison criterion *t*-value=2.052 (using 27 *df* at 95% confidence fromTable 3-1)
- square root of 28 = 5.2915 (using a square root computation program)

The highest value in our 95% confidence interval for the population mean would be:

$CI_{95\%}$ = 6.93 + [(2.052) * (10.452/√28)]

= 6.93 + [(2.052) * (10.452/5.2915)

= 6.93 + [(2.052) * (1.9752)]

= 6.93 + [4.053]

= 10.98 with rounding.

The lowest value in our 95% confidence interval for the population mean would be:

$CI_{95\%}$ = 6.93 – [4.053]

= 2.88 with rounding.

We would report this as: $CI_{95\%}$= (2.88, 10.98). This is interpreted as the school district's population mean days absent falling between 2.88 and 10.98 days, based on 95% confidence and the sample statistics computed for these 28 students.

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-3.1 Computing Confidence Interval for the Mean (CI 95%)

Chapter Summary

In this chapter, you learned important distinctions between univariate sample statistics and population parameters—that we estimate population parameters using sample statistics. You also learned about strategies for assessing the adequacy of our estimates and our confidence in the how well the sample statistics represent true values for the population. The essential information for developing confidence intervals was presented, including how statisticians select the comparison criterion *t*-value and degrees of freedom for computing and evaluation results from the statistical formula. A more advanced statistics course could help you better understand the nature of the *t*-distribution (and how this relates to normal distribution), the rationale behind degrees of freedom, and the assumptions we mentioned in relation to computing confidence intervals. The next chapter examines how some of these concepts apply to bivariate statistical analyses.

Module 4 Chapter 4: Principles Underlying Inferential Statistical Analysis

In addition to using inferential statistics to estimate population parameters from sample data, inferential statistics allow investigators to answer questions about relationships that might exist between two or more variables of interest (bivariate analyses). Inferential statistics help test initial hypotheses about "guessed at" relationships between study variables. In other words, investigators infer answers to quantitative research questions about populations from what was observed in sample data. In prior modules, you learned about study design options that maximize the probability of drawing accurate conclusions about the population based on sample data. Now we examine how to use inferential statistics to test bivariate hypotheses. To understand our level of confidence in the conclusions drawn from these statistical analyses, we first need to explore the role played by probability in inferential statistics.

In this chapter, you will learn about:

- probability principles and the role of probability in inferential statistics
- principles of hypothesis testing and the "null" hypothesis
- Type I and Type II error.

Understanding Probability Principles

You think about and calculate ***probability*** often during the course of daily living—what are the chances of being late to class if I stop for coffee on the way, of getting a parking ticket if I wait another half hour to fill the meter, of getting sick from eating my lunch without washing my hands first? While we generally guess at these kinds of probability in daily living, statistics offers tools for estimating probabilities of quantitative events. This is often taught in terms of the probability of a coin toss being heads or tails, or the probability of randomly selecting a green M&Ms® candy from a full bag.

Let's see what happens when we consider the probability of drawing a blue card from a deck of 100 Uno® cards–Uno® game card decks have equal numbers of Blue, Yellow, Red, and Green cards: 25 each (ignoring the un-numbered action and wild cards). In drawing one card from the deck, we have 100 possible outcomes–our draw could be any of the 100 cards. You may intuitively see that we have a 1 in 4 probability of blue being drawn (25%).

The following formula captures the probability of drawing a blue card (P_{blue}) in this scenario:

P_{event} = # of times event is possible / # of total possible

In other words:

P_{blue} = # of blue cards/# of total possible, or P_{blue} = 25/100 = ¼ = 0.25 = 25%

Back to Basics: Adding Fractions

To add fractions together, you simply add the numerators (top numbers) and then use the denominator (bottom numbers) if the denominators

are the same:
1/3 + 1/3 = 2/3.

But, if the denominators are not the same, you will need to convert them to a common denominator that can be kept intact.
Example: 1/3 + 2/5
Step 1. Multiply each fraction by 1 (any number multiplied by 1 is still that number). The value you select to equal 1 will be a fraction where the numerator and denominator are the same: 5/5 and 3/3 both equal 1. This is how you use that knowledge:
1/3 x 5/5 = 5/15
and 2/5 x 3/3 =6/15

Now that the

It is only slightly more complicated to determine the probability of a blue card being drawn later in the game, when the shuffled deck has been drawn down in play: for example, only 8 cards remain in the draw pile: 3 blue, 2 red, 2 green, and 1 yellow card. Now what is the probability of drawing a blue card? Again, we apply the probability formula, plugging in the different values.

P_{event} = # of times event is possible / # of total possible

P_{blue} = # of blue cards / # of total possible
P_{blue} = 3/8 = 0.375 = 37.5%

It is up to you to decide if you want to risk the outcome of the game on this chance of drawing a blue card!

Probabilities have at least three interesting characteristics. It is helpful to consider each of these three in a bit more detail.

Probabilities exist on a continuum. The continuum of probabilities is from 0 to 1.0—from absolutely no chance to a 100% guarantee.

denominators are the same, you can add the results:
5/15 + 6/15 = 11/15 (which divided is 0.7333)

You can check this by converting your initial fractions to decimals and adding them:
1/3 = 0.3333 and 1/5 =0.4, 0.333 + 0.4 =0.733

Probability differs from chance. This rule is at the heart of what is called the gamblers fallacy. Assuming that two possible outcomes have equal chances of happening (heads/tails on a coin, or having a girl versus a boy baby), it is possible to calculate the probability of any particular sequence occurring. Imagine that a couple has 3 daughters and are expecting their 4th child. They want to know whether they can expect the next child to be a boy. The answer is: the chance of having a boy remains the same—whether this is their first baby or their 10th baby. The chance remains 1 in 2, or 50%.

The probability question is a bit different than they expressed, however. They may not really be asking the probability of having a boy, they are actually asking the probability of having a family composition that is girl, girl, girl, boy. In that case, we can multiply the chance of each individual event happening to get the probability of the total combination happening:

1/2 girl x 1/2 girl x 1/2 girl x 1/2 boy

Back to Basics: Multiplying Fractions

To multiply fractions, first multiply the numerators (top numbers), then multiply the denominators (bottom numbers). You may need to resolve the resulting fraction to a simpler number or do the fraction division.

Example. Multiply 1/3 x 2/5
1 x 2 = 2 (numerators) and
3 x 5 = 15 (denominators), so
1/3 x 2/5 = 2/15 (which divided = 0.1333)

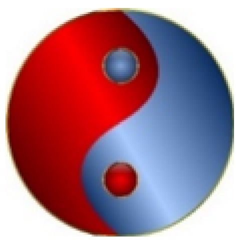
The result is 1 x 1 x 1 x 1 =1 for the numerator, divided by 2 x 2 x 2 x 2 = 16 for the denominator. In other words, the probability of this combination happening in this order of births is 1/16, or 0.625 or 6.25%. Remember there are many other possible combinations that the family could have had if the order does not matter, including no boys, one boy, two boys, three boys, or four boys among the girls, and each possible combination having the same probability of occurring in specified order (1 in 16):

The gambler's fallacy confuses the chances of the next child being a boy (1 in 2, or 50%) with the probability of a whole sequence/pattern occurring (girl, girl, girl, boy being 1 in 16, or 6.25%).

1st	2nd	3rd	4th
girl	girl	girl	girl
girl	girl	girl	boy
girl	girl	boy	girl
girl	boy	girl	girl
boy	girl	girl	girl
girl	boy	girl	boy
girl	girl	boy	boy
girl	boy	boy	girl
boy	boy	girl	girl
boy	girl	girl	boy
boy	girl	boy	girl
girl	boy	boy	boy
boy	girl	boy	boy
boy	boy	girl	boy
boy	boy	boy	boy
boy	boy	boy	boy

Remember: this only applies if the chance of the event happening is random—if there is a biological reason why boys are more or less likely than girls, the entire formula for probability will need to be revised. It is called the gambler's fallacy because people mistakenly bet on certain outcomes, confusing chance with probability. A real-life example presented in an online article about the gambler's fallacy (www.thecalculatorite.com/articles/finance/the-gamblers-fallacy.php) described a spectacular event that occurred in August of 1913 at a Monte Carlo casino: the roulette wheel turned up the color black 29 times in a row, the combined probability of which was 1 in 136,823,184! The article's authors stated, "After the wheel came up black the tenth time, patrons began placing ever larger bets on red, on the false logic that black could not possibly come up again." Obviously, it did, 19 more times. The gamblers' logic was false because the chances of black coming up was always 18 out of 37 times on this type of roulette wheel (48.6%). Therefore, each of those 29 times had an 18/37 chance of being black, regardless of what happened in prior spins. Do you see how this chance of an event (18/37) differed from the probability of the whole sequence occurring (1 in 136,823,184)?

Inverse probability. When there exist only mutually exclusive events, the probability of a specific event happening and the probability of that specific event NOT happening add up to 1. Thus, the probability that it will happen is the mathematical inverse of the probability that it will not happen. In other words, subtracting the probability that it will happen from 1.0 gives you the probability that it will not happen; and, vice versa, subtracting the probability that it will not happen from 1.0 gives you the probability that it will happen.



As an example, let's go back to our probability of drawing a blue card from the whittled down play deck with only 8 cards left. We computed the probability of a blue card (P_{blue}) as being 0.375. Therefore, we can apply knowledge of inverse probability if we want to know the probability of drawing something other than a blue card ($P_{notblue}$ is the combined probability of drawing red, green, or yellow):

$P_{notblue} = 1.0 - P_{blue}$

$P_{notblue} = 1.0 - 0.375 = 0.625$

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=431

Probability Principles Applied in Statistics

Probability is about the likelihood of an outcome. Remember that with inferential statistics we can never know with 100% certainty that our sample data led us to the correct conclusion about the population. This is where probability comes into play in the process of engaging in inferential statistical analyses. The following scenario helps make the case:

- Imagine you draw a random sample from the population for your study, accurately measure the variable of interest, and compute your sample statistics (for example, the mean and standard deviation for an interval variable).
- Now, imagine that you "put back" those sampled participants and draw a new sample. With random selection, any of the previous participants and all the other possible participants in

the population still have an equal chance of being drawn (we learned about this in Module 3, chapter 6 about random sampling).

- You repeat your measurement process and compute your new sample statistics. Because there are different participants this time, and there exists variability between individual participants, your sampling statistics for the different sampling groups will most likely be somewhat different—maybe a little different, maybe a lot different.

Back to Basics: Normal Distribution

Remember that one defining characteristic of a normal curve is that 95% of the values fall within 2 standard deviations of the mean (above or below). This is the probability of a value falling in that range (95%). This leaves a 5% probability of the

Which computed sample value is more "right" in describing the full population? The answer is: neither or both, and there is no way of knowing the population's true value without measuring the entire population (we learned about this in Module 4, Chapter 3). So, imagine your next step is to keep drawing samples, measuring, and computing the sample statistics for each sample drawn. Eventually, for a variable that is distributed normally in the population-as-a-whole, the sample statistic values that you computed over hundreds of trials (samples) will be approximately normally distributed—you learned about normal distribution in Chapter 4 of this module.

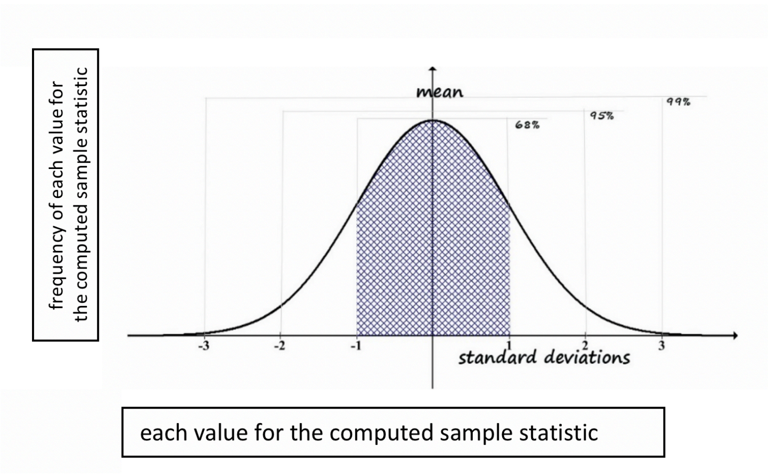
For example, imagine that you were computing the mean on a variable "number of hours each week spent working on issues related to someone's alcohol misuse" for each sample of social workers drawn 1,000 times—some social workers spent virtually none of their time in this type of activity while other spent almost all their time this way. Your first sample's mean number of hours was 15, your second sample's mean was 12, your third sample's mean was 22, and so on until you have 1,000 mean values for your drawn samples. If this variable is normally distributed across the population of all social workers and you drew each of your samples randomly from the population of social workers, those mean values would come close to looking like a normal "bell" curve. At this point:

> values falling outside of that range—4% at 2-3 standard deviations away from the mean (99% – 95%=4%), and 1% at more than 3 standard deviations away (100% – 99% = 1%). Furthermore, 5% is the inverse of 95%, since 1-.05=.95 and 1-.95=.05.

- You could compute the mean-of-means for the whole set of 1,000 samples. This sample mean-of-means would be a reasonable estimate for the population mean (population parameter).
- Knowing the mean-of-means value, you could compute the standard deviation for the mean number of hours across your 1,000 samples. The sum of squared distances for each sample's mean from the mean-of-means divided by 999 would be the variance, and the square root of the variance is the standard deviation (you learned about variance and standard deviation in Chapter 2 of this module).

Figure 4-1 is the same normal curve introduced earlier, this time with the notation that the *x*-axis (horizontal) represents the computed values for the 1,000 sample statistics, and the y-axis (vertical) represents the frequency each computed value for the sample statistic was the result. Note that our unit of analysis is no longer each individual social worker's hours—that was the unit of analysis in each of our 1,000 samples. Now the unit of analysis is the mean for each of those 1,000 samples.

Figure 4-1. Normal distribution around the mean



In reality, no investigator would set out to repeat a research study this many times. However, the investigator who completed the study the first time wants to know: is my sample statistic falling within the 95% range in relation to the true population mean, or does it lie in one of the two tails: the 2.5% below or 2.5% above two standard deviations away, adding up to the remaining 5%? This should sound familiar: it represents the 95% confidence interval (CI) for the sample statistic!

Using what we learned about the inverse probability relationship, selecting a 95% confidence level means accepting a 5% error rate (100% – 95% = 5%): or, in decimal terms, 1.0 – .95 = .05. In statistics, this is often referred to as applying a decision rule of “alpha” equals .05 (5%)–symbolized as $\boldsymbol{\alpha}$**=.05**. By definition, there exists an inverse relationship between the probability value (***p-value***) and the value of α. In other words, both of the following statements are true:

$$\alpha = (1 - p)$$

and

$$p = (1 - \alpha).$$

The social work, social science, and behavioral research convention is to adopt a 95% probability figure as its decision rule concerning conclusions drawn from sample data. In other words, we generally accept a 5% probability of being wrong in our conclusions. This is written as: α=.05.

Research Questions and the Null Hypothesis

Statistical decision rules apply to hypotheses being tested statistically, hypotheses drawn from the research questions (you learned about this question/hypothesis relationship in Module 2, Chapter 1). For example, an investigator wanted to know if readiness to change scores differ for men entering batter treatment programs voluntarily versus through court order. The investigator’s research question might be:

> *Is there a relationship between “program entry status” and “readiness to change” scores?*

Research questions often need to be rephrased in terms of statistically testable hypotheses. For example, this investigator’s research question needs to be rephrased as a statistical hypothesis:

> *A statistically significant relationship exists between the*

variables "program entry status" and "readiness to change" scores.

The investigator would know if the hypothesis is true is by determining whether the difference in readiness to change scores for the two sample groups is meaningfully different from 0—zero difference indicating that entry status is not meaningfully related to readiness to change scores.

The purpose of statistical analysis is to draw conclusions about populations based on data from a sample, making it very important to carefully phrase the hypothesis to be tested statistically. Two ways of phrasing research hypotheses are relevant here: they are the inverse of each other. One is called the ***null hypothesis***, the other is called the ***alternate hypothesis***. Let's start with the alternative hypothesis first because it is more intuitively easy to understand.

Alternate Hypothesis. An investigator develops hypotheses about possible relationships between variables. In the Safe at Home example, investigators hypothesized that a difference in readiness to change scores exists between men entering batterer treatment programs voluntarily compared to men entering the programs under court order. This research hypothesis needs to be re-worded into statistical terms before applying statistical analyses to evaluate its accuracy. If there exists a meaningful difference between the two groups, the mean readiness score for the voluntary entry group would be different from the mean for the court-ordered group—in other words, subtracting one mean from the other would give a result that is not zero. The alternative hypothesis, symbolized as H_a,would be phrased as:

H_a: *The difference in mean readiness scores for the population of men entering batterer treatment voluntarily compared to* *men entering under court order is significantly different from zero.*

You might be wondering why this re-phrasing of the research hypothesis is called the alternate hypothesis. It is the alternate (or inverse) to the null hypothesis, which is the one actually tested statistically. At first, this may seem like confused logic–why test the inverse of what we really want to know? The next sections explain this logic.

Null Hypothesis. A statistically tested hypothesis is called the null hypothesis. The null hypothesis (typically symbolized as H_0) is stated in terms of there being no statistically significant relationship between two variables among the population–in the Safe At Home example, program entry status and readiness to change scores. While the alternate hypothesis (H_a) was expressed in terms of the difference in mean readiness scores for the two groups in the population being different from zero, the null hypothesis specifies that the difference between readiness scores for the two groups is zero. The null hypothesis is the inverse of the alternate hypothesis and would be presented like this:

H_0: *The difference in mean readiness scores for the population of men entering batterer treatment voluntarily compared to men entering under court order is 0.*

Then, if a statistically significant difference in the sample means for the two groups is observed, the investigator rejects the null hypothesis. In other words, based on sample data, the investigator concludes support for the alternative hypothesis as it relates to the population.

When an observed difference is found to be non-statistically significant, the investigator fails to reject the null hypothesis. In other words, the true difference in the population could include zero.

Failing to reject the null hypothesis, however, is not the same as determining that the alternative hypothesis was wrong! Why not? Because the investigator can only conclude failure to detect a significant difference—it is possible that no difference truly exists, but it also possible that a difference exists despite not being detected for some reason. This point is so important that it warrants repeating:

> *Failure to reject the null hypothesis does not mean that the alternate hypothesis was wrong!*

Remember that inferential statistics are about what (unknown) differences might exist in the population based on what is known about the sample. Just because a difference was not observed this time, with this sample, does not mean that no difference really exists in the population. This conversation between a 6-year old and her father demonstrates the point:

> Dad: *Why are you afraid to go to bed in your own room?*
>
> Daughter: *I am afraid the zombies will come and get me.*
>
> Dad: *You don't need to be afraid. There is no such thing as zombies.*
>
> Daughter: *How do you know?*
>
> Dad: *Because no one has ever seen a real zombie.*
>
> Daughter: *That doesn't mean there aren't any—maybe no one looked in the right places!*

The 6-year old is logically more correct than her father—just because something was not observed does not mean that it does not exist. Failing to reject the null hypothesis means that the investigator recognizes that chance could be responsible for the outcome observed in the sample. Rejecting the null hypothesis means that the investigator is confident that a difference observed in the sample accurately reflects a true difference in the population, a difference not due to chance.

No Double-Barreled Hypotheses. Only one statistical hypothesis can be tested at a time. You may recall from our discussion about measurement and survey/interview questions (Module 3, chapter 4), the concept of a "double-barreled" question. The same applies in terms of testing the null hypothesis. If there were two parts to the hypothesis, it would be impossible to interpret the outcome: either or both could be rejected, and either or both could fail to be rejected. An investigator would not know whether to draw conclusions in terms of one or both parts of the hypothesis. Therefore, each needs to be phrased as a separate, unique null hypothesis.

For example, the following is not an analyzable hypothesis:

> H_0: *There is no statistically significant difference in either the*

readiness scores or number of weeks since the intimate partner violence incident for clients who enter the program voluntarily or by court order.

Instead, this should be presented as:

> *First H_0: There is no statistically significant difference in readiness scores for clients who enter the program voluntarily or by court order.*

> *Second H_0: There is no statistically significant difference in number of weeks since the intimate partner violence incident for clients who enter the program voluntarily or by court order.*

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=431

Statistical Significance in Hypothesis Testing

Notice that these hypotheses mention a "statistically significant" difference. What if the difference is not EXACTLY equal to 0, if there exists a tiny difference? Does that mean there is a meaningful relationship between the variables? Investigators want to know if an observed difference in the sample is MEANINGFUL before concluding that such a relationship really exists in the population. This is where we consider the meaning of statistically significant results—the observed difference in the sample is sufficiently different from zero to have statistical significant meaning (leading to the conclusion about rejecting or failing to reject the null hypothesis about the population). Statisticians work with probability logic in their statistical analyses of sample data.

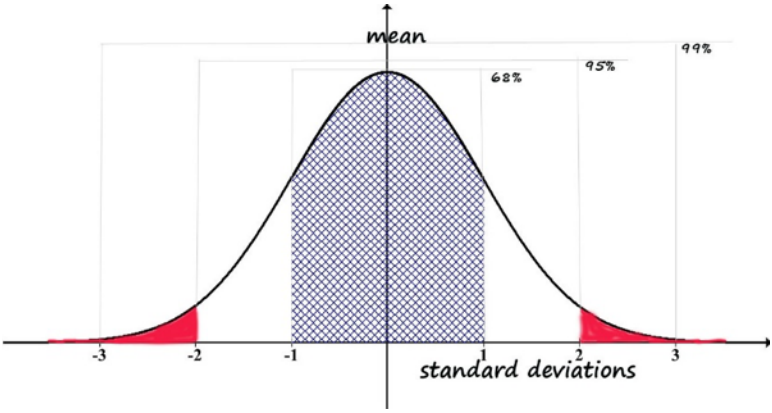
Thinking again about the Safe At Home data, the investigator knows the categorical status of each participant (n=75 voluntary, n=328 non-voluntary) and each participant's overall readiness to change score (higher scores indicate greater readiness to change one's behavior). Before we get to the actual statistics used to answer this question, we need to think about the question in terms of probability for drawing the right conclusion about the population of men entering treatment based on the sample statistics. The investigators in our Safe At Home example need a decision rule to help decide whether an observed difference for the two groups of clients (voluntary and court-ordered) is meaningful. This leads to examining the principles of hypothesis testing—remember we are statistically testing the null hypothesis: if a statistically significant result for the difference in scores for the two groups is observed, the null hypothesis of no difference (difference = 0) will be rejected.

Testing the Null Hypothesis. As previously noted, the decision rule convention in social work, social science, and behavioral research is to use α=.05 as the criterion for what is a significant result. This means that:

- if the computed significance level for a test statistic (the probability or p-value) is .05 or less, the decision is to reject the null hypothesis;

- if the computed significance level for a test statistic (probability or p-value) is greater than .05, the decision is failure to reject the null hypothesis.

Here is why. Inferential statistics draws on probability principles for help in interpreting data from samples. In any one single study, chance plays a role in the adequacy of the data for representing the population. For example, if investigators observe a difference in mean readiness to change scores between the sampled clients who entered batterer treatment voluntarily and sampled clients who were court-ordered into batterer treatment, they want to know if that observed difference happened just by chance or if it is reflecting a true difference in the population. By chance, with a normal distribution of readiness scores, 95% of the time an observed difference would fall within 2 standard deviations of the true mean. Therefore, any observed differences within that 2-standard deviation range could be the result of chance—so, investigators cannot safely conclude that a meaningful difference exists (i.e., they fail to reject the null hypothesis). However, if the observed difference is far enough out in the tails of the distribution (2.5% above or 2.5% below two standard deviations), investigators are comfortable with rejecting the null hypothesis of no difference and concluding that the observed difference reflects the population's reality.



This normal curve demonstrates our use of the α=.05 decision rule—we are willing to be wrong 5% (.05) of the time by concluding that a meaningful difference exists when none really does. In other words, an observed difference far enough out in one of these tails is considered statistically significant. So, we state that, if the probability that the observed difference happened by chance is less than 5% (α=.05), we are willing to reject the null hypothesis of no difference. The statistical approaches presented in Chapter 5 show how different test statistics are computed and how the probability is determined for these decisions. There is nothing magical about .05, it is simply a common convention in social work, social, and behavioral science research: sometimes investigators apply more stringent criteria (α=.01) or less stringent criteria (α=.10). When more stringent α=.01 decision criteria are applied, being 99% certain, it is harder to reject the null hypothesis: the observed values would have to be even further out in the tails of our distribution, out past three standard deviations. When less stringent α=.10 decision criteria are applied, being 90% certain, it is easier to reject the null hypothesis: the observed values do not have to be quite as far out in the tails of our distribution.

Turning back to our Safe At Home example, let's see what the investigators learned. The null hypothesis being tested was:

H_0: *the difference in mean readiness scores between men entering treatment voluntarily compared to men entering under court order is equal to 0.*

The investigator's sample statistics (using SPSS statistical analysis software) showed the following:

Group Statistics

	referred by cj personnel?	N	Mean	Std. Deviation	Std. Error Mean
pretest readiness to change score	yes	328	4.5644	1.25107	.06908
	no	75	4.9022	.97984	.11314

This table shows the mean value on readiness to change scores for the 328 men court-ordered to treatment was 4.56 and the mean for the 75 who entered the program voluntarily was 4.90 (with rounding). Clearly the difference between these two groups is not exactly zero: there exists a 0.34-point difference between the two group means (4.90 – 4.56 = 0.34). But is this 0.34-point difference observed in the sample sufficiently far from zero for investigators to reject the null hypothesis of there being no difference and to conclude that a population difference exists? This is another way of asking if the difference from zero is statistically significant.

The investigators in this case could use a *t*-test analysis approach for comparing the two group means (more about this later). The observed 0.34 difference in the two group means was significant at the p=0.029 level (probability=.03 with rounding) in this analysis. How does this translate into a conclusion about the results?

- The investigators adopted the tradition of using α=.05 criterion as a decision rule—accepting a 5% chance of drawing the wrong conclusion, being 95% confident of their conclusion.
- The computed statistical significance for the difference between the groups was 0.03 (rounding the p=0.029 level).
- The computed significance level (α=.03) is less than the accepted α=.05 criterion value, so the decision is to reject the null

hypothesis of no difference (concluding that a difference exists).

- Additionally, the confidence interval for the sample *t*-statistic did not include the value 0, since both the lower and upper range values were negative numbers/less than zero: $CI_{95\%}$= (-.641, -.034). Thus, the investigators are 95% certain that the real population difference does not include zero as a possibility and rejected the null hypothesis of no difference: the conclusion is that a statistically significant association between program entry status (voluntary versus court order) and readiness to change scores exists in the population of men entering batterer treatment programs.
- Looking at the observed sample means shows that individuals entering the programs voluntarily had higher readiness to change scores than did individuals court-ordered to enter the programs (M=4.90 and 4.56 respectively). This could be reported as the voluntary entry group has significantly higher mean readiness to change scores than the group entering treatment through court order.

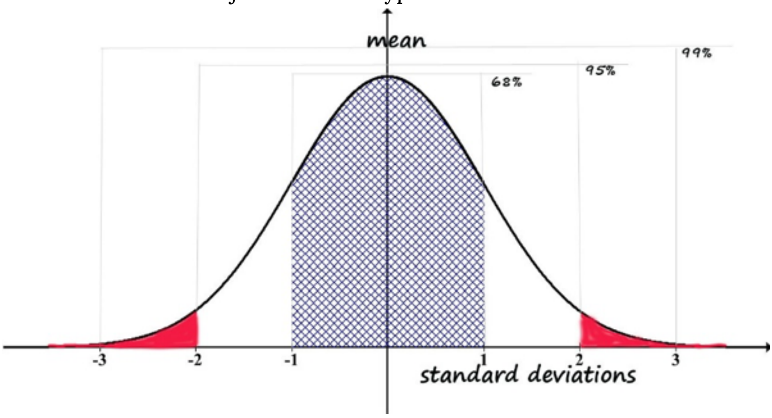
One-tailed and Two-tailed Hypotheses. Up until this point, we have worked with testing hypotheses about a "difference." The difference could be in either direction, either group being more or less than the other. For example, an alternative hypothesis (H_a) could be:

> H_a: A *significant <u>difference</u> exists in the readiness to change scores between men voluntarily entering batterer treatment and men court-ordered to treatment.*

In this case, the null hypothesis (H_0) being statistically tested would be:

> H_0: No *significant difference exists in the readiness to change scores between men voluntarily entering batterer treatment and men court-ordered to treatment (the difference is zero.)*

This null hypothesis could be rejected in either of two scenarios: either the voluntary entry group has higher readiness to change scores OR the voluntary entry group has lower readiness to change scores compared to the court-ordered group. In statistical terms, this means that we would be working with the probability of the result being either in the 2.5% tail above the mean or in the 2.5% tail below the mean (if we apply the α=.05 decision rule). This is what is meant by a ***two-tailed test***—the null hypothesis would be rejected by a result landing in either of the two tails (values falling in either of the red areas of the curve below); values landing anywhere else would result in failure to reject the null hypothesis.



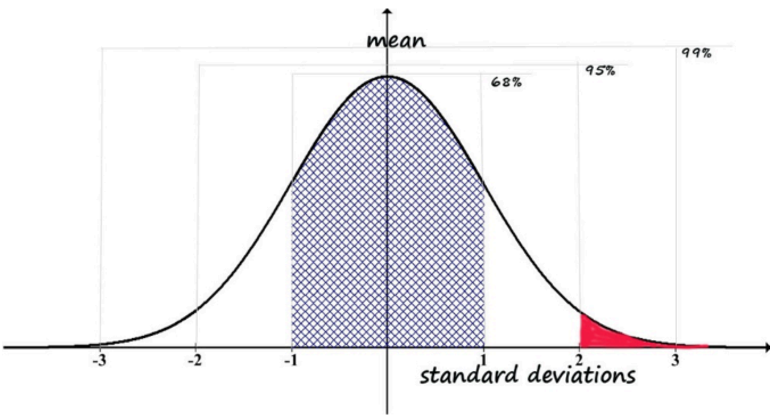
However, there exists another possible alternative hypothesis (H_a), depending on what the investigators think about the research question:

> H_a: *the voluntary entry group's readiness scores will be significantly greater than the court-ordered group's scores.*

This changes the null hypothesis (H_0) to the following:

> H_0: *the voluntary entry group's readiness scores will not be significantly greater than the court-ordered group's scores.*

Now there is only one tail that can lead to rejecting the null hypothesis: the resulting statistic would have to be in the 2.5% at the top tail of the distribution to reject. If it is in the 95% middle or in the bottom 2.5%, the investigator would fail to reject this null hypothesis. Since there is only one way to reject this time, not two, this is called a one-tail test (rejecting the null hypothesis for values falling in the red area of the curve, failing to reject the null hypothesis for values falling anywhere else).



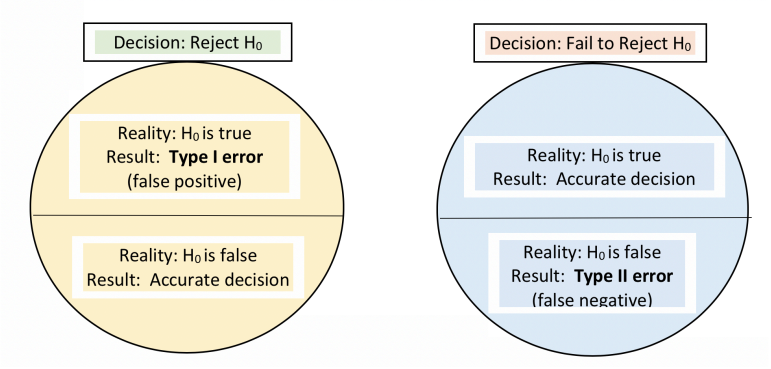
In summary, if the hypothesis is directional, the decision rule needs to be based on one-tailed test criteria; if the hypothesis is bi-directional (difference could be either way), the decision rule should be based on two-tailed test criteria.

Type I and Type II Error

While the investigators in our Safe At Home example would be happy to share their results and conclusions with the world, it is important to understand the possibility that their conclusions could have been drawn in error. Remember: being 95% confident means accepting a

5% probability of being wrong (α=.05 criterion). In any situation, 4 possible combinations of decisions and realities exist: two represent accurate decisions and two represent mistaken decisions. The mistake possibilities include a ***Type I error*** and a ***Type II error*** (see Figure 4-2 adapted from Begun, Berger, & Otto-Salaj, 2018, p. 13) depending on whether the investigator rejects or fails to reject the null hypothesis.

Figure 4-2. Type I and Type II errors for the null hypothesis (H_0=no difference)



Type I Error. The risk of making a Type I error concerns the risk of rejecting the null hypothesis of no difference when in reality there exists no significant difference in the population. For example, our Safe At Home investigator rejected the null hypothesis of no difference in readiness to change scores between the two groups. A Type I error would occur if, in fact, there is no difference in readiness to change within the population of men entering batterer treatment programs.

Sometimes, investigators decide to apply a looser criterion: α=.10 means a 10% chance of being wrong, or 90% confidence. This would increase the probability of making a Type I error since it is "easier" to reject at this level. There are few instances in social work, social science, and behavioral research where 99% certainty is preferable to 95% certainty (α=.01 rather than α=.05). The reason being that we also want to reduce the probability of a Type II error.

Type II Error. The risk of making a Type II error concerns the risk of failing to reject the null hypothesis when, in reality, a significant difference in the population does exist. For example, if an investigator uses the α=.05 decision criterion and computes a significance value p=.06 from the sample data, the null hypothesis would not be rejected. In reality, however, a difference may truly exist in the population. This is problematic with intervention research conducted with too few study participants; with small samples it is harder to reject the null hypothesis. A very promising intervention may get "scrapped" as a result of a Type II error. Thus, study design is an important strategy for reducing the probability of making a Type II error.

Trochim (2005, p. 204) explains Type I and Type II error this way:

> *Type I Error: Finding a Relationship When There Is Not One (or Seeing Things That Aren't There)*
>
> *Type II Error: Finding No Relationship When There Is One (or Missing the Needle in the Haystack).*

Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=431

Chapter Summary

This chapter laid the groundwork for understanding inferential statistical approaches presented in Chapter 5. In this chapter you learned basic principles of probability and how these probability principles play a role in the logic and assumptions underlying inferential statistics. As a result, you should now have a basic understanding of the decision rules that are commonly applied in social work, social science, and behavioral statistics—the role of p and α, for example. You learned about the logic behind the null hypothesis (H_0) and how it relates to a research question or "alternative" hypothesis (H_a),as well as what it means to reject or fail to reject the null hypothesis. You also learned about the distinction and meaning of Type I and Type II errors. Now you are prepared with a base for understanding specific statistical analyses.

Module 4 Chapter 5: Introduction to 5 Statistical Analysis Approaches

Prior modules and chapters laid the groundwork for understanding the application of inferential statistics approaches commonly used to evaluate research hypotheses based on quantitative data. Inferential statistics may answer either univariate or bivariate hypotheses.

In this chapter, you will learn about:

- Specific statistical approaches for testing hypotheses (parametric tests of significance)
        - one-samplet-test
        - independent samples *t*-test
    - one-way analysis of variance (Anova) test
    - chi-square test of independence
    - correlation test
- Approaches to consider when data are not normally distributed (non-parametric options)

What These Tests Have in Common

The five statistical analysis approaches presented here have several things in common. First, they each are based on a set of assumptions and/or have specific requirements related to the type of data analyzed. Second, with the exception of the chi-square test of independence, they rely on assumptions related to normal distribution in the population on at least one variable. Third, they all assume that the sample was randomly (properly) drawn from the population, and therefore are representative of the population. These last two factors characterize these are ***parametric tests*** (non-parametric tests are briefly discussed at the end of this chapter). For each of the statistical approaches, the task involves computing a test statistic based on the sample data, then comparing that computed statistic to a criterion value based on the appropriate distribution (t-distribution, F-distribution, χ^2 distribution, or correlations). This comparison leads the investigator to either reject or fail to reject the null hypothesis.

One Sample *t*-Test

One of the simplest statistical tests to begin with allows investigators to draw conclusions based on a single group's data about whether the population mean for the group is different from some known or expected standard mean value. For example, the standard IQ test has a mean value of 100—the test is designed so that the population is approximately normally distributed for IQ around this mean. Imagine that an investigator has concerns that a specific group in the population—perhaps children living in rental households—has a mean IQ significantly different from that standard mean of 100. In theory, this could happen from children being exposed to environmental toxins, such as lead in drinking/cooking water, and ingesting particles contaminated by lead-based paint or playground soil contaminated from heavy traffic where lead was once in the gasoline of cars driving in the area.

Test Purpose: Assess the probability that the population subgroup's mean IQ differs from the standard population mean of 100.

Test H_0: The group's mean IQ is not significantly different from 100.

Type of Data: IQ scores are interval data.

Test Assumptions and Other Test Requirements: The univariate ***one-sample t-test*** is based on the following assumptions and requirements.

- *Type of Variable:* The scale of measurement is either ordinal or continuous (interval).
- *Normal Distribution:* The dependent variable is normally distributed in the population, or a sufficiently large sample size was drawn to allow approximation of the normal distribution.

- *Independent Observations*: Individuals in the sample are independent of each other—random selection indicates that the chances of one being sampled are independent of the chances for any other being sampled.

The test statistic: The *t*-test analysis is based on the *t*-distribution. The *t*-statistic is computed by first subtracting from the sample mean ($\bar{X}$) the comparison value called "delta" (delta is indicated by the symbol Δ, which in our example has Δ = 100), then dividing this resulting value by the value computed as the sample's standard deviation (s) multiplied by a fraction with 1 divided by the square root of the sample size. The formula looks like this:

$$t_{(df)} = (\bar{X} - \Delta)/[(s) * 1/\sqrt{n})]$$

Back to Basics: Dividng by a Fraction

Dividing by a fraction is the same as multiplying by its inverse. For example, dividing by ½ is the same as multiplying by 2.

In our formula, $(sd)/\sqrt{n}$ is

It should remind you somewhat of the formula we saw back in Chapter 3 for the 95% confidence interval using the t-value—we are using the same components of sample mean ($\bar{X}$), t-value, standard deviation (s), and the square root of the sample size ($\sqrt{n}$):

$CI_{95\%} = \bar{X} + [(t_{\text{for 95\% confidence}}) * (s/\sqrt{n})]$ (computes highest value)

$CI_{95\%} = \bar{X} - [(t_{\text{for 95\% confidence}}) * (s/\sqrt{n})]$ (computes lowest value

Completing the example:In our hypothetical example, imagine the following: our sample mean ($\bar{X}$)for IQ was 92, the standard deviation (s) was

21.00, and the sample size was 49 children. This means that our degrees of freedom (*df*) is 49 -1, or 48 (because *df*=n-1). We can consult the *t*-distribution table (from Chapter 3) to determine our decision value for *t*-values based on 48 degrees of freedom and α=.05 (95% confidence). Since our table does not show a value for 48 degrees of freedom, we estimate the criterion value as 2.01 for purposes of our example.

equivalent to (sd) * 1/√n.

Plugging these values into the formula, we obtain our *t*-value:

$t_{(48)}$= (92 -100)/ (21.00/√49)

$t_{(48)}$= (-8)/ (21.00/7)

$t_{(48)}$= (-8)/3

$t_{(48)}$= (-2.67)

Now, we take that $t_{(48)}$= -2.67 value and compare it to the criterion value from the *t*-distribution table (2.01). Since 2.67 (the "absolute value" of -2.67) is more extreme than our criterion value of 2.01, we reject the null hypothesis of no difference. In other words, we conclude that a significant difference exists in the mean IQ score for this subpopulation compared to the expected standard mean of 100.

Back to Basics: Absolute Value

On the other hand, our statistical computations using the computer might have generated a *p*-value for this test. Instead of using the *t*-distribution table at 95% confidence to determine a criterion value for comparing the

In many instances, we are concerned with how far from zero a value lies. Thus, we work with figures representing the absolute value. For example, -4 and +4 are each 4 points away from 0. Thus, the absolute value of -4 is 4, as the absolute value of +4 also 4.

With comparing a test statistic to its criterion value, we generally work with the statistic's absolute value.

t-statistic we calculated, the p-value computation will lead to the same conclusion. In this example, imagine that the computer informed us that the probability of a chance difference (*p*-value) was .024: p=.024 in the output. Using the α=.05 decision criterion (95% confidence), we see that .024 is less than .05. Thus, we reject the null hypothesis of no difference, just as we did using the *t*-statistic compared to the criterion value.

Note: we used a two-tailed test in this example. If the investigator hypothesized that the subgroup has a lower IQ than the mean, ***one-tailed test*** criteria would need to be applied.

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-5.1 One-Sample t-Test Exercise

Comparing Means for Two Groups: Independent Samples *t*-Test

The next level of complexity involves comparing the mean for two independent groups, rather than what we just saw in comparing a single group to a standard or expected value. This moves us into a bivariate statistical analysis approach: one independent and one dependent variable. The independent variable is the “grouping” variable, defining the two groups to be compared. The dependent variable is the “test” variable, the one on which the two groups are being compared.

For example, we have been working throughout the module with the question of whether there is a meaningful difference in readiness to change scores for clients entering batterer treatment voluntarily versus through court order. In this case, we have one interval variable (readiness to change score) and one categorical variable with 2 possible categories—a dichotomous variable (entry status).

Test Purpose: Assess the probability that the readiness to change score in the population of clients entering batterer treatment voluntarily differs from the change scores for the population of clients entering batterer treatment through a court order.

Test H_0: The mean readiness to change scores for the two groups are not significantly different (the difference equals zero).

Type of Data: Independent “grouping” variable (entry status) is both categorical and dichotomous; dependent variable (readiness to change scores) is interval data.

Test Assumptions and Other Test Requirements: The ***independent samples t-test*** is based on the following assumptions and requirements.

- *Type of Variable*: The scale of measurement for the dependent variable is continuous (interval).
- *Normal Distribution:* The dependent variable is normally distributed in the population, or a sufficiently large sample size

was drawn to allow approximation of the normal distribution. Note: the "rule of thumb" is that neither group should be smaller than 6, and ideally has more.

- *Independent Observations*: Individuals within each sample and between the two groups are independent of each other—random selection indicates that the chances of any one "unit" being sampled are independent of the chances for any other being sampled.
- *Homogeneity of Variance*: Variance is the same for the two groups, as indicated by equal standard deviations in the two samples.

The test statistic: The independent samples *t*-test analysis is based on the *t*-distribution. The *t*-statistic is computed by first computing the difference between the two groups' means: subtracting one mean ($\bar{X}_{courtorder}$)from the other ($\bar{X}_{voluntary}$). This computed difference is then divided by a sample standard deviation value divided by the square root of the sample size, just as we did in the one mean test. However, this is more complicated when we have two groups: we need a pooled estimate of the variance (a pooled standard deviation) since neither standard deviation alone will suffice. Furthermore, we need a pooled sample size since both groups 1 and 2 contributed to the variance estimate. Our earlier formula looked like this:

$$t_{(df)} = (\bar{X} - \Delta)/[(s) * 1/\sqrt{n})]$$

Our new formula looks like this:

$$t_{(df)} = (\bar{X}_1 - \bar{X}_2)/[(s_{pooled}) * 1/\sqrt{n_1} + 1/\sqrt{n_2}).$$

Before we can start plugging in values, we need to know how to find the pooled variance estimate—the pooled standard deviation. The pooled standard deviation is computed as a sort of weighted estimate, with each group contributing the proportion of variance that is appropriate for its sample size—the smaller group contributing less to the estimate than the larger group. Thus, s_1 is the standard deviation for group 1, and s_2 is the standard deviation for group 2. Making a series of mathematical adjustments for the equivalent aspects of multiplying and dividing fractions, we end up with a formula for the pooled standard deviation (s_p) that works like this:

- take the square root of the entire computed value where
- the first group's degrees of freedom (sample size minus one, or n_1-1) is multiplied by its standard deviation (s_1^2) and
- this is added to the second group's degrees of freedom (n_2-1)multiplied by its standard deviation (s_2^2)
- and that addition becomes the numerator of a fraction where the two sample sizes are added, and the 2 extra degrees of freedom are subtracted (n_1+ n_2-2)

In terms of a mathematical formula, it looks like this:

$s_p = \sqrt{[(n_1-1)*s_1^2 + (n_2-1)*s_2^2 / (n_1 + n_2 - 2)]}$

Then, we take this pooled standard deviation and plug it into the *t*-test formula above. The degrees of freedom for this *t*-statistic will be the (n_1+ n_2− 2) that was in the denominator of our pooled variance.

Completing the example: The first step is to compute the pooled standard deviation value, then plug it into the formula for computing the *t*-statistic. In lieu of calculating this by hand, we can let Excel or another statistical program perform the calculation. The resulting t-statistic for comparing the mean readiness to change scores for our two groups, where the degrees of freedom were 401 (*df*=401):

$$t(401) = 2.189$$

The criterion value for 95% confidence was 1.960 in the *t*-distribution table. Therefore, the decision would be to reject the null hypothesis of equal means. The mean for the voluntary group was greater than the mean for the court order group: $\bar{X}$=4.90 and $\bar{X}$=4.56 respectively, a difference in the means of .34 ($\bar{X}_1 - \bar{X}_2$ = 4.90 – 4.56 = .34).

Two other ways of drawing the same conclusion (reject the null hypothesis) are offered by the computer output. The analysis indicated a 95% confidence interval of the difference in the two group means (.34) does not include zero as a possible value since the range includes only positive values: $CI_{95\%}$ = (.64, .03). In addition, the two-tailed significance level computed for the test statistic (*t*-value) was *p*=.029. Using the criterion of *α*=.05, seeing that .029 is less than .05, the same conclusion is drawn: reject the null hypothesis of no difference between the groups. All three of these approaches came to the same conclusion because they are simply different ways of looking at the very same thing.

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-5.2 Independent Samples t-Test Exercise

Comparing Means: One-Way Analysis of Variance (Anova) Test

The ***one-way analysis of variance (Anova) test*** is another bivariate statistical analysis approach. Like the independent samples *t*-test, it is used to compare the means on an interval (continuous) variable for independent groups. Unlike the independent samples *t*-test, it is not restricted to comparing only 2 groups. For example, the one-way analysis of variance would allow an investigator to compare the means between multiple racial/ethnic groups, or between 3 treatment conditions (e.g., low, medium, high dose intervention groups, or treatment innovation, treatment as usual, no treatment control groups). In this case, we still have a bivariate statistical analysis approach: one independent and one dependent variable. The independent variable is still the "grouping" variable, defining two or more groups to be compared. The dependent variable is the "test" variable, the one on which the groups are being compared.

Two points are worthy of noting here. First, because they are mathematically related, the one-way Anova test, when conducted on a dichotomous (two-group) categorical variable yields a result identical to the independent samples t-test. The F-statistic is the square of the t-statistic in this 2-group case, and the F-distribution (as well as the criterion value) are squared, as well. Thus, the conclusion drawn and the p-value are identical.

Second, in a one-way Anova involving three or more groups, the result indicates whether a difference in mean values exists—it does not indicate where the difference lies. For example, it could be that one group differs from all the others, that all of the groups differ from each other, or some combination of differences and no differences. A significant result leads to further "post hoc" analyses to determine which pairs differ significantly.

Extending our example with readiness to change scores for men entering batterer treatment, consider the possibility that investigators wanted to explore the relationship between readiness to change scores and substance use at the time of the referring incident of intimate partner violence. The "grouping" independent variable was coded for who used substances at that time: (0) substances were not involved, (1) the client only, (2) the partner only, and (3) both client and partner. Thus, the analysis involves 4 groups being compared on the dependent variable (readiness to change battering behavior).

Data were available on both the dependent and independent variables for 391 clients. Here is what the mean and standard deviations looked like, as well as the 95% confidence interval for each group mean (see Table 5-1).

Table 5-1. Results of descriptive analyses for group means.

Descriptives

pretest readiness to change score

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
substance not involved	212	4.5785	1.22717	.08428	4.4124	4.7446	.14	7.62
me only	60	4.7234	1.07567	.13887	4.4455	5.0013	2.55	7.29
partner only	22	4.0736	1.29558	.27622	3.4992	4.6480	1.43	7.07
both	97	4.8483	1.16909	.11870	4.6127	5.0839	2.40	7.57
Total	391	4.6393	1.20359	.06087	4.5196	4.7589	.14	7.62

You can see the number of individuals in each group under the column labeled "N" and the mean readiness to change score for each group in the next column (labeled "Mean"). The group with the highest mean score was the group where both partners were using alcohol or other substances at the time of the incident, followed by the group where the client was the only one using at the time. The lowest readiness to change scores were among men whose partner was the only one using alcohol of other substances at the time of the incident, lower than the group where no substances were used by either partner. The 95% confidence intervals for these mean values all excluded zero as a possibility, since lower and upper bounds of the confidence interval (range) were positive values.

Test Purpose: Assess the probability that the readiness to change scores differ for the population of clients entering batterer treatment following an incident of intimate partner violence where no substances, client only substance use, partner only substance use, or both using substances was involved.

Test H_0: The mean readiness to change scores for the four groups are equal (the differences equal zero). As previously noted, this multi-group test will not tell us which groups differ from the others, only whether there exists a difference to further analyze.

Type of Data: Independent "grouping" variable (substance use status) is categorical; dependent variable (readiness to change scores) is interval data.

Test Assumptions and Other Test Requirements: The one-way analysis of variance (Anova) test is based on the following assumptions and requirements.

- *Type of Variable*: The scale of measurement for the dependent variable is continuous (interval).
- *Normal Distribution*: The dependent variable is normally distributed in the population, or a sufficiently large sample size was drawn to allow approximation of the normal distribution. Note: the "rule of thumb" is that no group should be smaller than 7, and ideally each group has more.
- *Independent Observations*: Individuals within each sample and between the different groups are independent of each other—random selection indicates that the chances of any one "unit" being sampled are independent of the chances for any other being sampled.
- *Homogeneity of Variance*: Variance is the same for each group, as indicated by equal standard deviations in the groups' samples.

The test statistic: The Anova test analysis is based on the F-distribution. The **F-statistic** is computed as a complex difference value divided by a complex value computed with standard deviations and group sample sizes. This is the same logic used in our *t*-statistic computation, just more complexly computed because there are more groups involved. The *F*-statistic is then compared to the criterion values for the *F*-distribution, the same way we did this comparison using the *t*-distribution to find the criterion *t*-statistic values. It is beyond the scope of this course to compute the *F*-statistic manually; it is sufficient to understand what it means.

One important facet of the *F*-statistic evaluation is understanding the degrees of freedom involved. In fact, with the *F*-statistic there are two different types of degrees of freedom involved: between groups and within group degrees of freedom. Between groups degrees of freedom is related to the number of groups, regardless of group size. In our example involving 4 groups, the between groups degrees of freedom is 4-1, or *df*=3. Within groups degrees of freedom is related to the number of individuals in the entire sample, but instead of being n-1, it is n minus the number computed for between groups degrees of freedom. In our example with data provided by 390 men from 4 groups, the within groups degrees of freedom is 390 – (4-1) = 387. This is written as *df*=(3,387)—this is not read as three thousand three hundred eighty seven degrees of freedom, it is read as 3 and 387 where the first number is always the between groups degrees of freedom and the second number is always the within groups degrees of freedom.

Completing the example: The next task is to apply our computed degrees of freedom to an *F*-distribution table to determine the criterion value for *df*=(3, 387) at 95% confidence (α=.05): this turns out to be a criterion value of 2.60. Then, we allow the computer to generate the comparison *F*-statistic for our data, as presented in this table (see Table 5-2).

Table 5-2. Results of one-way analysis of variance, 4 group mean readiness to change scores.

ANOVA

pretest readiness to change score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12.486	3	4.162	2.915	.034
Within Groups	552.480	387	1.428		
Total	564.966	390			

The values in the Sum of Squares and Mean Square cells relate to what would be plugged into the formula that we are not delving into in this course. The *F*-statistic computation answer would be written as $F(3, 387)=2.92$, $p=.034$ for this analysis. Therefore, the decision would be to reject the null hypothesis of equal means since the *F*-statistic value of 2.92 is greater than the criterion value of 2.60 from the *F*-distribution table.

Another way of drawing the same conclusion (reject the null hypothesis) is offered by the computer output. The analysis indicated a significance level computed for the test statistic (*F*-statistic) was $p=.034$. Using the criterion of $\alpha=.05$, seeing that .034 is less than .05, the same conclusion is drawn: reject the null hypothesis of equal means. Both of these approaches came to the same conclusion because they are simply different ways of looking at the very same thing. Post-hoc comparisons would be required to determine which of the differences in means listed in Table 5-1 are significantly different; remember that the test only told us that a statistically significant difference exists .

Interactive Excel Workbook Activities

Complete the following Workbook Activity:

- SWK 3401.4-5.3 One-Way Analysis of Variance (Anova) Exercise

Chi-square Test of Independence

The ***chi-square test*** (often written as the χ2test, pronounced "chi-square") is a bivariate statistical analysis approach. It is used to evaluate whether a significant relationship exists between two categorical (nominal) variables. Data are conceptualized in the format of a contingency table (see Table 5-3), where the categories for one of the two variables are represented as rows and categories for the other variable are represented as columns. The chi-square test is concerned with analyzing the relative proportions with which cases from the sample are sorted into the different cells of the contingency table.

For example, we can look at the Safe At Home data concerning whether there exists a significant relationship between treatment entry status (voluntary or court order) and treatment completion (completed or dropped out). The data would be entered into a 2 x 2 contingency table, where O_i is the observed number of cases in each cell, with i being 1, 2, 3, or 4 depending on which cell is being identified (see Table 5-3).

Table 5-3. Contingency table for chi-square analysis of treatment entry status and treatment completion.

		Program Completion		
		yes (completed)	no (dropped)	row total
Entry Status	court-ordered	O_1	O_2	
	voluntary	O_3	O_4	
	column total			

Test Purpose: Assess the independence of treatment entry status and treatment completion variables.

Test H_0: No association between the variables exists (i.e., the variables are independent of each other if differences in the cell proportions equal zero).

Type of Data: Two variables are categorical, with two or more possible categories each.

Test Assumptions and Other Test Requirements: The chi-square test is based on the following assumptions and requirements.

- *Type of Variable:* The scale of measurement for both variables is categorical (nominal).
- *Distribution:* Normal distribution is not assumed with categorical variables. The χ^2-distribution is used to evaluate the test statistic.
- *Independent Observations* :Individuals within each sample and between the different groups are independent of each other—random selection indicates that the chances of any one "unit" being sampled are independent of the chances for any other being sampled.
- *Estimating Expected Cell Counts*: In order to compute the χ^2-test statistic, the number of cases expected in each cell of the contingency table if the two variables are independent of each other must be calculated; this requires sufficient sample size

and a minimum of 5 expected cases per cell.

The test statistic: The chi-square test analysis is based on the $\chi 2$-distribution. The $\chi 2$-test statistic is computed in terms of the observed counts in each cell compared to what would be the expected counts if no relationship exists between the two variables. In essence, the formula for computing the $\chi 2$-test statistic is relatively simple: for each cell, compute the difference between what was observed and what was expected, square the difference, and divide the result by the expected value for that cell; then, add up the results for the individual cells into a single value. In mathematical terms, the formula looks like this:

$$\chi^2 = \sum(O_i - E_i)^2 / E_i$$

- χ^2 is the relevant test-statistic
- the $\sum$ symbol means add the values up (compute the sum)
- O_i is the observed number in the i^{th} cell (i designates each cell in the contingency table; 1 through 4 in a 2×2 table; 1 through 6 in a 2 x 3 table; 1 through 9 in a 3×3 table; and so forth)
- E_i is the expected number in the i^{th} cell (i designates each cell in the contingency table; 1 through 4 in a 2×2 table; 1 through 6 in a 2 x 3 table; 1 through 9 in a 3×3 table; and so forth)

The only thing we still need to know is how to calculate the expected (E_i) value for each of the contingency table cells. Again, this is a simple calculation based on the row, column, and sample size total values related to each cell.

E_i= (row total * column total)/sample size

Completing the Example: In the Safe At Home study example investigators wanted to know if there exists an association between program entry status and program completion. (The null hypothesis being tested is that no association exists; that the two variables are independent of each other). Table 5-4 presents the completed 2 x 2 contingency table with the observed (O_i) values are presented, as are the calculated expected (E_i)values for each cell.

Table 5-4. Data distribution for chi-square analysis of treatment entry status and treatment completion.

		Program Completion		
		yes (completed)	no (dropped)	row total
Entry Status	court-ordered	O_1 = 174 E_1 = 173.7	O_2 = 234 E_2 = 234.3	408
	voluntary	O_3 = 38 E_3 = 38.3	O_4 = 52 E_1 = 51.7	90
	column total	**212**	**286**	**498**

- O_1 is court-ordered clients completing treatment: E_1=(212 * 408)/498=173.7
- O_2 is court-order clients dropping out of treatment: E_2=(286 * 408)/498=234.3
- O_3 is voluntary clients completing treatment: E_3=(212 * 90)/498=38.3
- O_4 is voluntary clients dropping out of treatment: E_4=(286 * 90)/498=51.7

We now have all the information necessary for computing the χ^2-test statistic.

$$\chi^2 = \sum(O_i - E_i)^2/E_i$$

$$\chi^2=[(174-173.7)^2/173.7]+[(234-234.3)^2/234.3]+[(38-38.3)^2/38.3]+[52-51.7)^2/51.7]$$

$$\chi^2=[.3^2/173.7]+[-.3^2/234.3]+[-.3^2/38.3]+[.3^2/51.7]$$

$$\chi^2=[.09/173.7]+[.09/234.3]+[.09/38.3]+[.09/51.7]$$

$$\chi^2=.00052 + .00038 + .0023 + .0017$$

$$\chi^2=.0049$$

What remains is to evaluate the χ^2-test statistic which we computed as .005 (with rounding up from .0049). This value can be compared to the criterion value for χ^2-distribution based on the degrees of freedom at α=.05. In order to do so, we need to know the degrees of freedom for this analysis. In general, degrees of freedom for a chi-square test are computed as the number of rows minus 1 and the number of columns minus 1. In a mathematical formula, this would be:

$$df=(\#rows-1, \#columns-1).$$

Thus, for our 2 x 2 contingency table, the degrees of freedom would be 2-1 = 1 for each, rows and columns: we would write this as df=(1,1). The criterion value is then looked up in a chi-square distribution table. For df=(1,1) at 95% confidence (α=.05), the criterion value is 3.841. In our example, the computed .005 is less than the criterion value, therefore we fail to reject the null hypothesis. In other words, there could very well be no association between the two variables—they could be operating independently of each other.

Another way of drawing the same conclusion is to have computed the probability value (p-value) for our analysis. In this case, the statistical program computed a p-value of .941 for our χ^2-test statistic. Because .941 is greater than our decision rule of α=.05, we fail to reject the null hypothesis.

Interactive Excel Workbook Activities

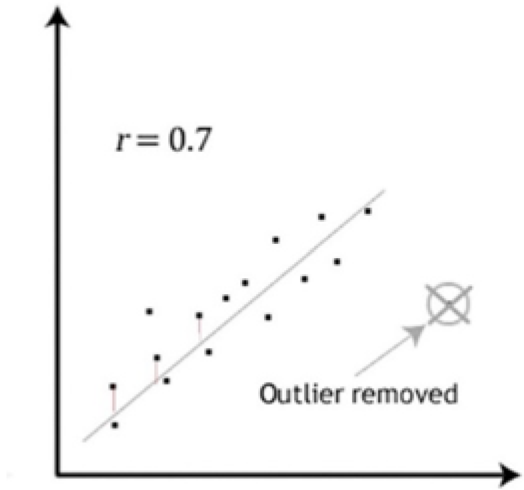
Complete the following Workbook Activity:

- SWK 3401.4-5.4 Chi-Square (χ^2) Analysis Exercise

Correlation Test

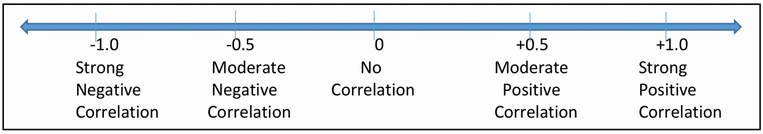
Correlation is used to analyze the strength and direction of a linear relationship between two interval (continuous) variables. Similar to the chi-square test, it assesses the probability of a linear association between the two variables, but in this case the variables are not categorical, they are interval numeric variables. The correlation analysis involves determining the single line that best fits the observed sample data points. In other words, the line which has the shortest possible total of distances each point lies from that line. Figure 5-1 shows just such a line drawn through hypothetical data and the first few distances from data points to that line are shown in red. The goal is to find the line where the total of those distances is the least possible.

Figure 5-1. Best line through data points with correlation coefficient 0.7, with one outlier removed. (Note: image adapted from https://statistics.laerd.com/spss-tutorials/pearsons-product-moment-correlation-using-spss-statistics.php)



The ***correlation coefficient*** is designated in statistics as *r*. The value for *r*communicates two important dimensions about correlation: strength of the association or relationship between the two variables and direction of the association.

Strength of Relationship. The possible range of the correlation coefficient is from -1 to +1. If two variables have no linear relationship to each other, the correlation coefficient is 0. A perfect correlation between two variables would be either -1 or +1—in other words, the absolute value of*r*is 1. Neither of these extremes usually happens, so the strength of the association is determined by how far from zero the correlation coefficient (*r*) actually lies. This diagram might help you visualize what this means:



The histograms we saw in Chapter 2 depicted univariate data—the frequency with which specific values on a single variable were represented in the sample data. Bivariate analyses are depicted using a ***scatterplot*** where each data point (dot) on the scatterplot indicates an individual participant's value on each of two variables, the *x*-variable and the*y*-variable (height and weight, for example). Figures 5-2 and 5-3 graphically depict hypothetical examples of data points with zero or no correlation (no evident line) and a perfect +1 correlation.

Figure 5-2. Hypothetical scatterplot with no correlation.

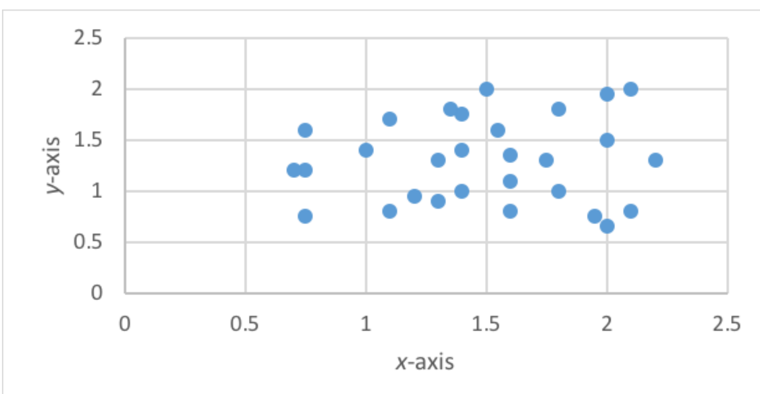
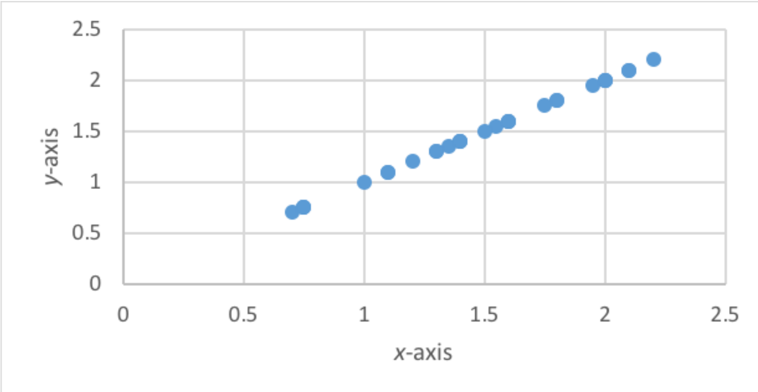
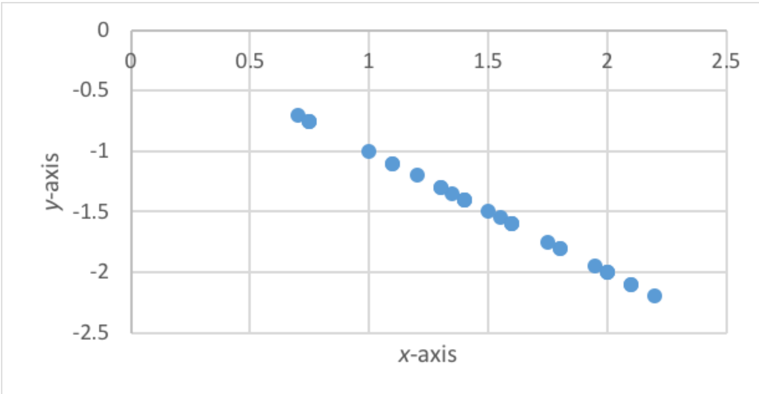


Figure 5-3. Hypothetical scatterplot with perfect +1 correlation (45^{o} angle line).



Direction of Association. There is a meaningful difference between a negative and a positive correlation coefficient. A perfect positive correlation of +1, as seen in Figure 5-2 indicates that as the values of one variable increase (*x*-axis), so do the values of the other variable (*y*-axis). On the other hand, a perfect negative correlation of -1 would indicate that as the values of one variable increase (*x*-axis), the values of the other variable decrease (*y*-axis). The perfect negative correlation is presented in Figure 5-4.

Figure 5-4. Hypothetical scatterplot with perfect -1 correlation (45^{o} angle line).



Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=436

Test Purpose: Assess the strength and direction of association between two variables.

Test H_0: The two variables are unassociated in the population (i.e., their correlation is 0).

Type of Data: Both variables are interval for a Pearson correlation; if either/both is ordinal, a Spearman correlation is preferable.

Test Assumptions and Other Test Requirements: Correlation tests require the following conditions and the following considerations.

- Each case analyzed must have complete data for both variables; if more than one correlation is computed (multiple pairs of variables), cases can be deleted from the analysis wherever this is not true (pairwise deletion) or can be deleted from all analyses (listwise deletion).
- Data on each variable should be approximately normally distributed.
- Data should be assessed for "outliers" (extreme values) which can affect the correlation analysis. Outlier cases, typically having values more than 3.25 standard deviations from the mean (above or below), should be excluded from correlation analysis.
- Distribution of the data should be in a relatively tubular shape around a straight line drawn through the data points on a scatterplot; for example, a cone shape violates this assumption of ***homoscedascity***. Homoscedascity refers to the degree of variance around the line being similar regardless of where along the x-axis the points are mapped. A cone shape would mean that low values on the $x$variable have data points closer to the line than high values on the $x$variable (a case of heteroscedascity).
- Distribution of data should be relatively linear. For example, a curved line reflects a non-linear relationship between the variables. Non-linearity violates an assumption necessary for correlation analysis.
- Statistically significant correlation coefficients are increasingly

likely as sample sizes increase. Therefore, caution must be used in interpreting relatively low correlation coefficients (below about 0.3) which are flagged as statistically significant if the sample size is large.
- Correlation simply indicates an association between two variables, it does not in any way indicate or confirm a causal relationship between those variables. Always remember that correlation does not imply causality (you have read this before in other modules). Causal relationships require different kinds of study designs and statistical analyses to confirm.

The test statistic: As previously noted, correlation analysis is based on identifying from among an infinite array of possible lines, the one line that best defines the observed data—the line where the cumulative distance from data points to that line is minimized. The formula looks more complicated than it actually is. The correlations coefficient (r) is computed from a numerator divided by a denominator, just as our previous coefficients have been. The numerator is made up of the sample size N (the number of pairs for the two variables) multiplied by the sum of each pair of values (x times y for each case, summed) minus the result of multiplying the sum of the x values times the sum of the y values. The denominator is the square root of multiplying two things together: first, the sample size N times the sum of the x^2 values minus the square of the sum of the x values, then doing the same for the y values—sample size N times the sum of the y^2 values minus the square of the sum of the y values. In a mathematical formula, it looks like this, where the final step will be dividing the numerator by the denominator:

$$r \text{ numerator} = N * \sum x * y - (\sum x)*(\sum y)$$

$$r \text{ denominator} = \text{square root of } [N*\sum x^2 - (\sum x)^2] * [N*\sum y^2 - (\sum y)^2]$$

With large numbers of cases, hand calculating this correlation coefficient is cumbersome—computer software programs can do the computation very quickly.

*Completing the example:*Consider for example that the Safe At Home project investigators wanted to assess the association between readiness to change scores and the length of time between client entry into the treatment program and the intimate partner violence incident leading to treatment entry (measured in weeks). In our example, the computed correlation coefficient was r=-.060 for N=387 clients entering batterer treatment for whom completed data were available and following removal of three outliers with extremely long time lag between the incident and program entry.

The negative sign on the correlation coefficient suggests that greater number of weeks between incident and program entry might be associated with lower readiness to change scores. This could be explained in several different causal models, so it is not acceptable to infer that one caused the other:

- More time passing leads to less impetus to change.
- Low impetus to change leads to longer lag time before entering a program.
- Some other factor or factors influence both time to program entry and readiness to change scores.

The strength of the correlation coefficient needs to be assessed. The coefficient is not far from zero (.06 is not even $1/10^{th}$ of the way to 1.0), suggesting that the association is weak. This is particularly notable since the sample size was so large (N=387): even a small correlation should be more easily detected with a large sample size. Allowing us to reject the null hypothesis is the computed two-tailed significance value where p=.241. Since .241 is greater than our decision rule of α=.05, we fail to reject the null hypothesis. In other words, we conclude that there may be no association between the two variables. In reality, this finding could be explained by some other factor or factors influencing the length of time before entering a program—events such as incarceration for the offenses could interfere with entering treatment, for example.

Interactive Excel Workbook Activities

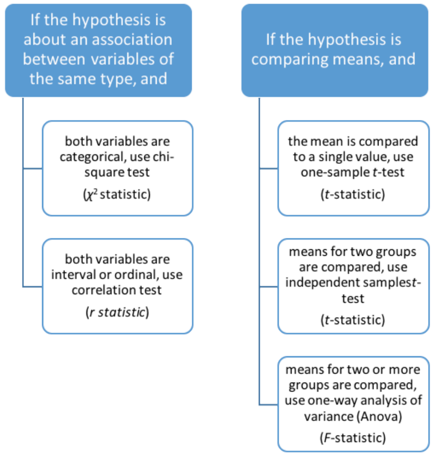
Complete the following Workbook Activity:

- SWK 3401.4-5.5 Correlation (r) Exercise

Choosing the Right Statistical Test for the Job

You have now been introduced to 5 powerful statistical tools for analyzing quantitative data. The guidelines for choosing the right tool for the job have to do with (1) the nature of the hypothesis being tested, and (2) the nature of the variables being analyzed. This flow chart might help (see Figure 5-5).

Figure 5-5. Decision flow chart for 5 statistical analytic approaches.



Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=436

With the exception of the chi-square test, all of the approaches we explored are based on assumptions of normal distribution across the population on at least one variable; an assumption assessed based on the distribution of data from the sample. In many instances, true normal distribution is not required, as long as the sample size is sufficiently large to approximate normal and extreme outliers have been addressed. However, it is important to know that there exist alternative approaches to statistical analysis when the assumption of normal distribution does not hold up. The alternatives are called ***non-parametric tests***.

For example, when a t-test is the right option but the data are non-normal in distribution, a non-parametric alternative is the ***rank sum test***—also called the ***Wilcoxon-Mann-Whitney test***. The test statistic in this case is called U instead of *t* and is compared to a different distribution table. This test has four main steps:

- Rank order all of the cases from lowest to highest on the dependent variable, regardless of which group they represent—all are in one single pool.
- Sum the ranks for one of the two groups (it does not matter which is selected)—this becomes R_1 in the formula below.
- Compute the result of [number of cases in $group_1$ * (number of cases in $group_1$+1)] divided by 2.
- The test statistic is the result of subtracting this last computation from the R_1 value.

$$U = R_1 - [n_1 * (n_1 + 1)]/2$$

The test statistic is then compared to the distribution for U and/or a *p*-value can be computed.

Similarly, the ***Kruskal-Wallis H test*** is a one-way Anova computed on ranks as a non-parametric alternative to the one-way Anova we previously learned. As in the case of the Wilcoxon-Mann-Whitney U test, the Kruskal-Wallis H test operates using the rank orders of cases rather than the actual case and mean values for the dependent variable. In this way, variance is removed as a potential problem in the computation.

Chapter Summary

In this chapter you learned about 5 different statistical analysis approaches: one-sample *t*-test, independent sample *t*-test, one-way analysis of variance (Anova), chi-square, and correlation. You learned about the principles and assumptions underlying these statistical approaches, and how to determine which might be the best choice for different scenarios. Throughout the chapter you learned how statistical probability and the application of decision criteria are applied in drawing conclusions about populations based on samples—rejecting or failing to reject the null hypothesis. You were introduced, as well, to basic principles concerning non-parametric tests that can be used when the underlying assumptions necessary for parametric inferential statistical analyses are violated.

Module 4 Summary

This information-loaded module began with an introduction to how investigators manage and analyze qualitative data. You learned a bit about transcription and coding practices and issues, and are encouraged to pursue more advanced training in these skill sets. The remainder of the module introduced a sizeable number of topics, issues, and skills related to analyzing quantitative data. One important distinction you learned about concerns univariate and bivariate analyses. In the univariate domain you learned univariate techniques for describing data representing different types of variables (categorical, dichotomous categorical, ordinal, and interval/continuous variables). These included frequency and proportion descriptive statistics, as well as central tendency (mean, median, mode) and variability (range, variance, standard deviation) statistics. You also learned how this type of information might be presented in tables and figures, and learned the basic elements of normal distribution. This module explained the probability logic underlying inferential statistics and null hypothesis testing, as well as Type I and Type II errors. You learned to craft and interpret null hypotheses, and how information about statistical criterion values are managed (distribution tables, confidence intervals, p-values, and α levels). Five basic parametric approaches for inferential statistical analysis were introduced along with their underlying assumptions and how they are interpreted: one-sample t-test, independent samples t-test, one-way analysis of variance (Anova), chi-square, and correlation analyses. Finally, you were introduced to non-parametric approaches which might be used in situations where the assumptions underlying the parametric approaches might be violated. Throughout the module, you were provided with opportunities to practice new skills using Excel. Our next (and final) course module addresses how results from social work, social science, and behavioral research might be presented to varied audiences.

Module 4 Key Terms and Definitions

***α*=.05:** convention in social work, social, and behavioral sciences for decision to reject/fail to reject the null hypothesis when analyzing quantitative data.

***alternate hypothesis*:** a research hypothesis phrased in terms of the statistical analysis to be performed on observed differences with quantitative data; often designated as Haor H1 (see null hypothesis for more information).

***a priori coding*:** qualitative data analysis utilizing a pre-determined set of coding categories (in contrast to open coding).

***bivariate analysis*:** quantitative data analysis approaches evaluating the association or relationship between two variables.

***central tendency*:** quantitative univariate statistical approaches (e.g., mean, median, mode) that serve as indicators of how data cluster around a "common" value.

***chi-square (χ2) test*:** bivariate quantitative statistical analysis for evaluating an association between two categorical variables based on the observed and expected proportions in a variable x variable contingency table; evaluated using criterion values based on the χ2-distribution and degrees of freedom.

***coding*:** means of creating categories out of themes present in qualitative data for analysis.

***confidence interval (CI)*:** range of statistical test values that would arise by chance at the determined probability level (usually 95% confidence interval in social work, social, and behavioral science).

context: descriptive information in qualitative research that aides the interpretation of coded data.

correlation: bivariate quantitative statistical analysis for evaluating an association between two interval (continuous) variables based on the "best" fitting line for the observed data points.

correlation coefficient (r): the statistical value computed in correlation analysis which helps determine statistical significance, strength, and direction of the relationship between two variables.

degrees of freedom: a component of many inferential statistics approaches based on a computation related to the number of cases in a data set and, depending on the statistical test, the number of groups being compared.

descriptive statistics: a set of univariate analysis approaches used to describe how a sample is distributed along various dimensions, particularly frequency, central tendency, and variability.

F-statistic: sample statistic computed using group means, standard of deviation, and both within and between group degrees of freedom; compared to criterion values based on the F-distribution.

field notes: supplemental information in qualitative data and analysis that provide insight into the context of data collected from participants and the thinking process of the investigator analyzing the data.

frequency: univariate quantitative statistical information concerning how a sample distributes on a (categorical) variable or some combination of categories for multiple categorical variables (subgroups) based on the count of numbers in each group.

histogram: a univariate graphing approach showing the frequency with which values appear in a sample.

homoscedascity: principle in quantitative data analysis involving data in relation to a "best" fitting line, reflecting how the data points cluster around that line; refers to the degree that these distances between data points and the line are similar.

independent samples t-test: bivariate quantitative statistical analysis for comparing mean scores (interval variable) for exactly two groups (categorical variable); evaluated using criterion values based on the *t*-distribution and degrees of freedom.

inferential statistics: set of quantitative data analysis approaches based on drawing inferences about a population based on data from a sample; these approaches have specific assumptions and limitations that need to be evaluated and addressed for accurate interpretation of the results.

Kruskal-Wallis H test: a non-parametric quantitative statistical analysis used when assumptions related to normal distribution and sample size are violated; analog to the parametric one-way analysis of variance (Anova) compares rank order of data points in groups rather than actual data values and group means.

kurtosis: one dimension for evaluating a sample's distribution on an interval (continuous) variable, specifically how it meets criteria for normal distributions regarding concentration of data in the center versus tails of the distribution curve.

mean: univariate quantitative statistic indicating one facet of central tendency (often called "average"); particularly susceptible to distortion from outliers and small sample sizes.

median: univariate quantitative statistic indicating one facet of central tendency (often called “50thpercentile”); half of cases fall below and half fall above this value.

mode: univariate quantitative statistic indicating one facet of central tendency, represents the most common value in the data (note there may be more than one mode present).

non-parametric tests: inferential quantitative statistical analyses that are not dependent on assumptions of normal distribution or large sample size (in contrast to parametric tests).

normal distribution: distribution of values on an interval (continuous) variable that falls on a curve that is (a) symmetrical around the mean, and (b) specific proportions of cases fall within 1 standard deviation, 2 standard deviations, 3 standard deviations and more from the mean (symmetrically above and below the mean); often depicted as a “bell-shaped” graph.

null hypothesis: a research hypothesis phrased in terms of no statistically significant difference based on statistical analysis to be performed with quantitative data; often designated as H0 (see alternative hypothesis for more information).

one-sample t-test: univariate quantitative statistical analysis for comparing a group’s mean score (interval variable) to a specific standard value; evaluated using criterion values based on the *t*-distribution and degrees of freedom.

one-tailed test: comparison statistic is based on probability related to one tail of the normal distribution, rather than both tails, because the hypothesis being tested is directional (more than or less than hypothesis) instead of a bidirectional “difference” hypothesis (in contrast to a two-tailed test).

one-way analysis of variance (Anova) test: bivariate quantitative statistical analysis for comparing mean scores (interval variable) for two or more groups (categorical variable); evaluated using the *F*-distribution and degrees of freedom.

open coding: qualitative data analysis utilizing coding categories that emerge out of the data being analyzed (in contrast to *a priori*coding)

outliers: data points that are exceptionally different from the majority of cases in a sample; having potential to greatly distort central tendency measures like the mean.

p-value: a probability value computed in inferential quantitative statistical analyses, used to interpret statistical test values in terms of rejecting or failing to reject the null hypothesis.

parametric tests: inferential quantitative statistical analyses that are dependent on assumptions of normal distribution and sufficiently large sample size (in contrast to non-parametric tests).

percentage: univariate quantitative statistical information concerning how a sample distributes on a (categorical) variable or some combination of categories for multiple categorical variables (subgroups) calculated as the proportion of the total in each group (out of 100%).

population parameters: values for central tendency and variability (dispersion) for an entire population; when data are not available for the entire population, sample statistics are utilized to estimate population parameters while recognizing the probability of error in these estimates.

probability: a mathematical/statistical concept concerned with computing the likelihood of an event occurring; the inverse (the likelihood of an event not occurring) may also be computed; probability forms the logic basis underlying inferential statistics.

projective test: an observational measurement technique used to reveal internal thought processes or personality traits as individuals respond to ambiguous stimuli (words, images, objects, scenarios/ situations, music, or other forms of presentation).

proportion: the segment or portion something represents of the whole (usually presented in fraction or percent format).

range: the mathematical difference between the lowest and highest values in a sample data set, or the lowest to highest possible values for a measurement tool.

rank sum test/Wilcoxon-Mann-Whitney U test: a non-parametric quantitative statistical analysis used when assumptions related to normal distribution and sample size are violated; analog to the parametric independent samples *t*-test, compares rank order of data points in groups rather than actual data values and group means.

sample statistics: values for central tendency and variability (dispersion) for a sample of cases drawn from a whole population; when data are not available for the entire population, sample statistics are utilized to estimate population parameters while recognizing the probability of error in these estimates

scatterplot: a bivariate graphing approach showing data points for the paired values of each case in a sample (x-axis and y-axis reflect the values for the two variables).

skew: one dimension for evaluating a sample's distribution on an interval (continuous) variable, specifically how it meets criteria for normal distributions regarding symmetry of data around the mean.

standard deviation: a sample statistic used to estimate population variance, computed as the square root of the sample variance.

t-statistic: sample statistic computed using group means and standard of deviation; compared to criterion values based on the t-distribution.

two-tailed test: comparison statistic is based on probability related to both tails of the normal distribution because the hypothesis being tested is bidirectional (unequal) instead of a unidirectional "greater" or "less" than hypothesis (in contrast to one-tailed test).

transcription: a common step in preparing qualitative data for analysis, constructing a verbatim written record of the information provided by each study participant.

univariate analyses: quantitative data analysis approaches evaluating data for one variable at a time (e.g., frequency, proportion, central tendency, variability/dispersion).

variance: quantitative univariate computed statistical value that serves as an indicator of how values are dispersed (vary) for either a population or a sample (see standard deviation).

verbatim: word-for-word reporting, in exactly the same words, to preserve the original meaning of what was originally said.

Wilcoxon-Mann-Whitney test (U test)/rank sum test: a non-parametric quantitative statistical analysis used when assumptions related to normal distribution and sample size are violated; analog to the parametric independent samples t-test, compares rank order of data points in groups rather than actual data values and group means.

Module 4 References

Begun, A.L., Barnhart, S.M., Gregoire, T.K., & Shepherd, E.G. (2014). If mothers had their say: Research-informed intervention design for empowering mothers to establish smoke-free homes. *Social Work in Health Care*, 53(5), 446-459.

Corcoran, J., & Secret, M. (2013). *Social work research skills workbook: A step-by-step guide to conducting agency-based research.* NY: Oxford University Press.

Creswell, J.W., & Poth, C.N. (2018).*Qualitative inquiry & research design: Choosing among five approaches, 4thed.*Thousand Oaks, CA: Sage.

Engel, R.J., & Schutt, R.K. (2018). The practice of research in social work, 3rded. Thousand Oaks, CA: Sage.

Glaser, B.G., & Strauss, A.L. (1967).*The discovery of grounded theory: Strategies for qualitative research.* New Brunswick, NJ: Aldine.

Graham, H. (no date). APA format for statistical notation and other things. Retrieved from http://users.sussex.ac.uk/~grahamh/RM1web/APA%20format%20for%20statistical%20notation%20and%20other%20things.pdf

Hinduja, S., & Patchin, J.W. (2008). Personal information of adolescents on the Internet: A quantitative content analysis of MySpace. *Journal of Adolescence*, 31(1), 125-146.

McInroy, L.B., & Craig, S.L. (2015). Transgender representation in offline and online media: LGBTQ youth perspectives. *Journal of Human Behavior in the Social Environment*, 25, 606-617.

Nadelman, L. & Begun, A.L. (1982). Effects of the newborn on the older sibling. In M. E. Lamb & B. Sutton-Smith (Eds.), *Sibling relationships: Their nature and significance across the lifespan*, (pp. 13-37). New Brunswick, NJ: Lawrence Erlbaum.

Rubin, A., & Babbie, E. (2013). *Essential research methods for social work*, *3rded.*Belmont, CA: Brooks/Cole, Cengage Learning.

Salkind, N.J. (2017). Statistics for people who (think they) hate statistics, 6thedition. Thousand Oaks, CA: Sage.

Taleb, N. (2012). *Antifragile: Things that gain from disorder.*NY: Random House.

Tracy, S.J. (2013). *Qualitative research methods*: *Collecting evidence, crafting analysis*, *communicating impact*. Malden, MA: John Wiley & Sons, Inc.

Trochim, W.M.K. (2005). *Research methods: The concise knowledge base*. Mason, OH: Atomic Doc/CENGAGE Learning.

MODULE 5

Module 5 Introduction

In our final module we learn how data and evidence about diverse populations, social phenomena, and social work problems are prepared for presentation to various audiences. As social work professionals you will be called on to help different types of audiences understand complex social work problems, social phenomena, and diverse populations. In this module, some powerful tools are explored for their communication potential. We not only emphasize the ways investigators present evidence, we also address how professional reports and presentations can help our audiences identify or recognize social justice issues and can help advocate for change. In this module we learn to construct tables, figures, graphs, and graphic materials using commonly accessed software like Word®, PowerPoint®, and Excel®. We continue to practice the American Psychological Association (APA) style guidelines for professional writing.

READING OBJECTIVES

After engaging with these reading materials and learning resources, you should be able to:

- Describe parts of an empirical manuscript or report;
- Use the American Psychological Association (APA) style guidelines for professional writing;
- Create a strong presentation for various audiences;
- Develop graphs and charts of different types for different communication purposes;
- Design an infographic communication tool;
- Summarize what was learned in this first course in the 2-course research and evidence sequence;

- Define key terms related to communicating social work evidence to varied audiences.

Module 5 Chapter 1: Parts of an Empirical Manuscript or Report

In Module 2 you learned about the different parts of an empirical article from the standpoint of critically reviewing them. The emphasis in this module is from the other side of the coin: how they are composed. The same sections are discussed: the title, abstract, introduction, methods, results, discussion, and references. However, emphasis in this module is placed on writing issues, and how professional reports might differ from empirical manuscripts. In this chapter you will learn:

- What to include in each section of a manuscript or report;
- Important writing issues of language use and avoiding (accidental) plagiarism;

What to Include and Where

In this section, writers are coached about what pieces of information belong in each section of an empirical manuscript or professional report. A great deal of this information is available in the Publication Manual of the American Psychological Association 6thedition (APA, 2009).

Title. A good title conveys to readers what the manuscript or report is about. It is also no more than about 12 words in length. A good title is fairly specific and concrete rather than vague and general. Consider these "good" examples–note that each describes the main topic, methodology, and/or population addressed.

- Understanding Latino grandparents raising grandchildren through a bioecological lens (Mendoza, Fruhauf, Bundy-Fazioli, & Weil, 2018)
- Parent, teacher, and school stakeholder perspectives on adolescent pregnancy prevention programming for Latino youth. (Johnson-Motoyama et al., 2016).
- Innovative ethics: Using animated videos when soliciting informed consent of young people for online surveys (McInroy, 2017).
- Self-Identified strengths among youth offenders charged with assault against a non-intimate family member (Mengo et al., 2017).
- Gender differences in pathways from child physical and sexual abuse to adolescent risky sexual behavior (Yoon, Voith, & Kobulsky, 2018).
- Pathways to age-friendly communities in diverse urban neighborhoods: Do social capital and social cohesion matter? (Parekh, Maleku, Fields, Adomo, Schuman & Felderhoff, 2018).
- Using social determinants of health to assess psychological distress and suicidal thoughts and behavior among youth on probation (Quinn, Liu, Kothari, Cerulli, & Thurston, *in press*).
- The role of family financial socialization and financial management skills on youth saving behavior (Kagotho, Nabunya, Ssewamala, Mwangi, & Njenga, 2017).

Notice that these report titles include details about when or where the information relates:

- 2015-2016 Ohio State's LiFE Sports annual review. Columbus, OH: College of Social Work. Ohio State University (Anderson-Butcher, Wade-Mdivanian, Davis, Newman, & Lower, 2016).
- Connecting Youth in Child Welfare to Behavioral Health Services in Franklin County: Provider Perspectives. Evaluation report

written for Franklin County Children Services (OH), (Bunger, Robertson, Hoffman, & Cao, 2015).
- Evaluation of Marathon County Community Response Program: Report to Marathon County Department of Social Services (Maguire-Jack & Campbell, 2015).

For more information, consult the Publication Manual (APA, 2009) on page 2.

Abstract or Executive Summary. Although the ***abstract*** or ***executive summary*** appears first to readers, it is one of the very last things written by the author(s). This is because it summarizes all the remaining information. For a manuscript (or professional conference presentation), an abstract allows readers to quickly review what the manuscript is about, to see if it is sufficiently relevant for them to continue reading the entire manuscript. A manuscript abstract is usually 150-200 words. Abstracts for professional conference presentation proposals vary in length, depending on the preferences of the conference organizers. Typically, they range from 200 to 500 words in length. The abstract typically follows the format of the manuscript: introduction, methods, results, and discussion. Hitting just the right balance between these sections of the abstract is important: the introduction and discussion each are usually only one or two sentences long, with the bulk of the space devoted to methods and results. In addition to the APA manual (APA, 2009), a useful resource is available on line: https://www.verywellmind.com/how-to-write-an-abstract-2794845.

On the other hand, if the written product is a professional report (to a funding agency or board, for example), then the abstract is replaced with an initial summary statement, often called an executive summary. Like the abstract, it highlights key points, especially a project's process, results, and recommendations/implications. However, it is written in a more "bullet point" style, and less like a journal abstract. The report's executive summary ideally fits on a single page. If you want to see some examples, search for "executive summary" or "technical report" samples.

For more information, consult the Publication Manual (APA, 2009) pages 25-27.

Introduction. The ***introduction*** to a manuscript is usually organized in the following manner:

- An introduction of the topic or problem addressed—its background and significance (i.e., the problem dimensions, costs, impact on other problems, impact on which populations). This is the "so what" statement—why was it worth the effort to address this problem and why should readers care.
- Summary of literature and prior research, with a critical review of what has been done, relevant theory, and important knowledge gaps.
- Research questions and the rationale for them; if a quantitative study, hypotheses are also identified here.

A strong literature review and introduction to an empirical manuscript (adapted from John Elliott, an instructor at Ohio State University):

- Is clearly organized
- Flows logically
- Uses primary literature sources
- Specifies types of studies reviewed (qualitative, quantitative, mixed methods)
- Includes recent studies, although classic studies can be included if they help set the stage
- Provides a balanced appraisal of the material presented, not leaving out sources that might disagree with an author's point of

view and discussion of strengths/weaknesses of reviewed studies is balanced
- Presents what is relevant for providing context related to the current issue/project, does not include the entire universe of possible materials
- Provides a framework for what is known and what is not known
- Leaves the reader aware of what has come before, what is needed, and why it is needed, and leads logically to the specific research questions and hypotheses.

A professional report, on the other hand, might rely less on the literature in the introduction (although the project should be informed by literature), emphasizing instead the background and context of the project's initiation and development.

For more information, consult the Publication Manual (APA, 2009) pages 27-29.

Methodology. The ***methodology*** or methods section is written with the intent of describing for readers what was done to conduct the study so that (1) it could be replicated by another investigator and (2) evaluation of the methods can be performed by a reviewer or potential consumer of the evidence. This section typically begins with a brief introductory paragraph describing the study design and/or sources of data (e.g., if it is secondary or administrative data set, where the original data came from). It should be clear to a reader what "model" is being tested in a quantitative study (e.g., univariate descriptives or bivariate relationships between specified independent and dependent variables).

Next, the study participants are described along with details about their recruitment, including participation rates and sampling strategy where appropriate. The description includes information about numbers, characteristics, how they reflect or differ from the population they represent, and how many fit into each group or category if there were quantitative analyses conducted.

The next subsection of methods describes how data were collected—the measurement tools and procedures. These are described in terms of the variables that they represent (in quantitative projects) or in terms of qualitative methods employed. This section also describes any manipulation of the data that needed to be made (e.g., creating new variables out of measured variables or transcription of qualitative data, for example).

For more information, consult the Publication Manual (APA, 2009) pages 29-32.

Results or Findings. The art of writing the ***results*** section is to clearly and succinctly present the findings in an objective manner—this is not the place where the results are analyzed, interpreted, or critiqued (that is in the discussion section). This is the place to tell readers what was observed in the data collected. It is important that all relevant results be reported—those that fail to confirm hypotheses are just as important as those that confirm. The tables, graphs, and charts that you learned about in earlier sections of this course are valuable tools for reporting details; they replace the need to describe the details in the test of the manuscript or report. This section may need to explain HOW findings were generated (what analyses were conducted), in addition to WHAT the finding might have been. In a research manuscript, the label "results" is usually applied; in a professional report, the label might either be "results" or "***findings***."

For more information, consult the Publication Manual (APA, 2009) pages 32-35. In addition, the Publication Manual provides a great amount of detailed advice concerning how to report numbers and statistics, as well as how to display results in tables and figures (pages 111-161).

Discussion or Recommendations. The final segment of an empirical manuscript is the author's ***discussion*** of the study's interpretation or analysis of the study and its results. It usually begins with a brief summary of key results and how they relate to the study aims initially presented in the introduction. Then, the results are discussed in terms of their implications for theory, future research, and/or practice. These conclusions need to be justified by the evidence presented in the study—it is not acceptable to reach far afield of what the data actually support. Included in the Discussion section is an analysis of any potential limitations of the study, such as the sample, the measurement tools, or the data quality. The Discussion section typically concludes with a statement of what the study contributes to knowledge and what more might need to be explored in the future. In a professional report, the focus of the concluding discussion is a bit different: it emphasizes specific ***recommendations*** based on the findings.

For more information, consult the Publication Manual (APA, 2009) pages 35-36.

Writing Tips

The APA (2009) Publication Manual discusses writing style for a manuscript that apply to most professional writing: write clearly, write concisely, and organize the content (see pages 61-70). One of the best ways to improve your professional writing is to read others' manuscripts with an eye to style as well as to content. Several additional tips are worth considering, as well.

Verb Tenses. For the most part, an empirical manuscript or professional report is written in the past tense. This is because the work has been completed. Even in summarizing the literature, past tense is preferable. For example, in their review of literature concerning integrative body, mind, and spirit practices in social work, the authors (Lee, Wang, Liu, Raheim, & Tebb, 2018) generated past-tense statements such as:

- "Studies were identified through electronic databases…" (p. 252)
- "The review process consisted of …" (p. 252), and
- "The review process involved two independent reviewers rating each study" (p. 253).

Here is an example of ***confused verb tense*** (see underlined verbs), along with a reasonable repair:

> *Begun, Early and Hodge (2016) reported that 82.5% of women preparing for release from jail engagedin pre-incarceration substance misuse that placed them at risk of a substance use disorder; 64% areat very high risk.*

> *Begun, Early and Hodge (2016) reported that 82.5% of women preparing for release from jail engagedin pre-incarceration substance misuse that placed them at risk of a substance use disorder; 64% wereat very high risk.*

Regardless of your choice of verb tense, it is critically important that the tense be applied consistently.

Parallel construction. The concept of ***parallel construction*** applies at two different levels. First, it applies to the different sections of a manuscript or report. For example, if you describe three research questions or hypotheses in the introduction (1, 2, and 3), then it is important to maintain that structure in that order in the methods, results, and discussion sections, as well (i.e., 1, 2, 3 throughout). That structure becomes the skeleton on which the rest of the manuscript or report is built.

Second, parallel construction also applies to sentences. For example, if a sentence begins with a singular subject (e.g., a person) then the verb and other parts of the sentence need to be in the singular, as well (e.g., was, not were). Likewise, if a sentence begins with a plural subject (e.g., adolescents), the verb and other parts of the sentence need to be plural, too (are, not is). One more example of parallel construction: notice that the last two sentences are similarly constructed—the "bridge" word came first (i.e., For example and Likewise), followed by the "if" clause which was followed by a "then" clause.

Language use. In its most recent editions, the Publication Manual emphasized the importance of using language that is respectful, free from bias, and non-stigmatizing (APA, 2009, pp. 70-77). Reviewers, editors, and audiences are increasingly sensitive to the need for using language that reflects an appreciation for diversity, particularly as related to being inclusive around gender, age, race, ethnic group, national origin, religion/spirituality, sexual orientation, and residence, as well as other social, health, functional, and personal circumstances, conditions, or diversity (Begun, 2016; Broyles et al., 2014). For example, strong efforts are made to move away from stigmatizing terms like "mental retardation" and labels that identify people by diagnosis: mentally retarded, schizophrenic, addict, and epileptic. Instead, a predominating theme is to "***put people first***" by using terms such as "persons experiencing addiction" or "individuals diagnosed with schizophrenia." In this way, we avoid making the diagnosis "be the person," and instead make it a part of the person's experience. In addition to avoiding the use of labels, the APA (2009) guidelines recommend avoiding describing groups in stigmatizing comparisons. The example they offered relates to comparing one group to a "normal" group, such as comparing lesbians to "normal women" or "the general public" is inappropriate.

This awareness of language use in social work professional writing extends beyond diagnosis. Reducing bias, stigma, and microaggression that stems from careless use of language is an important goal. One strategy recommended in the APA (2009) Publication Manual is to ask persons from the groups represented in a manuscript or report what terms are preferred and to have them review what is written for the presence of inadvertent bias. For example, staff in an Africentric program may wish to avoid the term Afrocentric, believing the latter relates to a hairstyle and the former describes a program grounded in African cultural traditions (i.e., the seven core principles of Kwanzaa being Umoja, Kujichagulia Ujima, Ujamaa, Nia, Kuumba, and Imani). Others may prefer the term Afrocentric, believing that there is no such word as Africentric. Both groups may prefer these terms to being described as an African American program. As another example, in reference to designated parking, many individuals prefer the term "accessible parking" rather than "handicapped parking" because the emphasis is on the accommodation rather than the disability.

The APA (2009) guidelines also suggest describing compared groups using parallel terms. They gave the example of unparalleled nouns: man and wife. This is considered problematic because the woman is described only in terms of her relationship to the man; preferable is the use of parallel terms, such as man and woman or husband and wife. Or, a non-gendered term like "spouse" might be appropriate.

Paraphrasing and citing. It is always important to avoid ***plagiarism***. Period. Always. Knowing how to ***paraphrase*** and cite your sources are two skills to help avoid accidental plagiarism. Here is an example of paraphrasing with the use of ***citation***. In the original article, the authors (Mengo & Black, 2015, p. 244) wrote:

> *Sexual victimization appears to have a stronger negative impact on students' academic performance than physical/verbal victimization.*

A first choice for applying this information in your own writing is to use it as a direct quote with an appropriate citation. For example:

> *Mengo and Black* (2015) *concluded that*, "*Sexual victimization appears to have a stronger negative impact on students' academic performance than physical/verbal victimization*" *(p.* 244).

Because it is a direct quotation, the page number is also cited. However, the APA (2009) publication manual discourages over-use of direct quotes. Used too often, they create a choppy, poorly integrated result. Sometimes a previous author has stated something so perfectly that a direct quote is desirable. Most of the time, you will need to convert your references to others' work by paraphrasing. Here is the example paraphrased and accompanied by the proper citation:

> *Students' academic performance is more negatively affected by sexual victimization than by victimization through acts of physical or verbal violence* (Mengo & Black, 2015).

Formatting citations. In the APA style, citations are either "in text" or "parenthetical" in nature. The first example shows how an "in text" citation is formatted. The authors' names are connected using the word "and" since you are reading it as part of sentence, and the publication date appears in parentheses (2015). The second example shows how a "parenthetical" citation is formatted. Notice here that the authors' names are combined with the "&" sign, not the work "and"' also, their names appear in the set of parentheses along with the publication date of their article. The paraphrased statement ends with a period afterthe reference citation in parentheses. Many students believe that formatting citations is the main contribution of the APA (2009) style guidelines. While this is covered (p. 169-192), there are many other important style tips and formatting guidelines presented in the manual.

References. The full ***reference*** for any citation appears in the reference list at the end of your manuscript or report. References are placed in alphabetical order by the first author; initials are used rather than full first names. Here is the reference listing for our example's citation:

Mengo, C., & Black, B.M. (2015). Violence victimization on a college campus: Impact on GPA and school dropout. *Journal of College Student Retention: Research, Theory & Practice*, *18*(2), 234-248.

If you check the reference list at the end of this module, you will see this reference is placed ahead of the next entry authored by the same first author, Mengo, Lee, Bolton, Lehmann, & Jordan (2017). Similarly, you will see that the reference for Begun, Early, and Hodge (2016) comes after the reference for Begun (2016) but before the reference for Begun and Rose (2007). This is because the second author's names are the next step in alphabetizing the references.

The format requirements for articles, books, book chapters, websites, news or magazine articles, and other sources are presented in the

APA (2009) publication manual (see pages 193-216). On line resources are also available to help with formatting references, one of the most commonly used is called Purdue Owl, distributed by Purdue University through the website https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_guilde/general _format.html.

For each item identify if the paraphrasing is appropriate for writing your own manuscript or report (True) or not (False). The paraphrasing is based on the following material published in an article by Doogan and Warren (2017, p. 1430). The article reports findings from a social network analysis into the dynamics of communication exchanges between group members participating in therapeutic community (TC) sessions as part of residential treatment while incarcerated.

> *These findings support the TC use of the community of peers as the method of treatment* (De Leon, 2000). *They suggest that the best way for TC clinical staff to foster an active unit is to encourage residents to intervene with each other. The increased likelihood of residents to correct peers after receiving a peer affirmation supports De Leon's* (2000) *claim that affirmations serve to energize residents and increase program participation. Finally, this pattern fits the definition of generalized reciprocity; residents who have received help from peers show a tendency to pass it along* (Molm, Collett & Schafer, 2007, Stanca, 2009), *perhaps because they feel an*

increase in their own positive feelings (Isen & Levin, 1972). There is evidence that generalized reciprocity increases feelings of self-competence (Alvarez & van Leeuwen, 2015). Generalized reciprocity in response to peer affirmations may therefore promote positive clinical outcomes.

An interactive or media element has been excluded from this version of the text. You can view it online here: https://ohiostate.pressbooks.pub/swk3401/?p=599

Chapter Summary

This chapter identified what content belongs in which sections of an empirical manuscript or professional report. Additionally, important tips about writing professionally were presented, with specific attention to using language free of bias and stigma. Writers were repeatedly referred to the information presented in the APA (2009) Publication Manual since the guidelines for writing in APA style are applied prevalently in social work.

Module 5 Chapter 2: Creating Strong Presentations

Presenting evidence to audiences depends a great deal on the audience. A room full of scientists will be looking for a different emphasis in the information shared than an agency board, group of practitioner colleagues, policy decision makers, or a community group like the League of Women Voters, Parent Teachers Association, or Kiwanis service club. However, all audiences want to be respected with interesting, accurate, concise, and clearly presented information. The major difference between preparing a written report and a live presentation has more to do with style than content. In this chapter, you will learn:

- Basic elements of strong presentations;
- PowerPoint® tips;
- Using images in presentations, including copyright issues and locating open source materials;
- Presentation style tips.

Elements of Strong Presentations

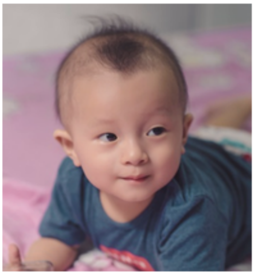
One of the most important aspects of a strong presentation is its organization—how content flows from start to finish. However, a strong presentation requires appropriate and adequate preparation up front. Here are six important tips (based on experience and tips from http://thevisualcommunicationguy.com/2015/10/30/how-to-give-a-powerful-presentation-eight-steps-to-an-awesome-speech/ and https://www.entrepreneurship.org/learning-paths/powerful-presentations).

1. **Preparation**. Effective presenters prepare for their specific audiences. At the very least, they develop an understanding of who the audience is, what the audience expects to learn from the presentation, and how the presentation fits into the rest of the day's experiences. For example, a presenter would want to know if their presentation is keeping the audience from lunch or if it is following a meal that makes people sleepy; if it is being presented at a critical, contentious time in the group's process; if are they being forced to attend or attending out of interest; if it is one of many presentations in a day or stands alone.
   Knowing the audience expectations is also important: are they looking for the science behind the key points (emphasis on the methodology) or the implications of the evidence? A strong presentation starts with understanding why the audience wants to hear what the presenter has to say instead of looking at their smartphones or laptops as they politely (or impolitely) endure the presentation.
   Timing is another important preparation point. Attention tends to wane after only a few minutes—is the presentation expected to be a short 15-minute Ted-style talk or an hour-long lecture? In either case, preparing to engage the audience at multiple points is critical for maintaining attention, interest, and learning. Having appropriate and interest-grabbing visual, auditory, and/or action-oriented aids planned into the presentation is important. Long gone are the days when audiences tolerate presentations from lecturers who "wing it"—they expect the respect of strong preparation. Realize that it can take several hours to prepare a 15-minute presentation when you have a great deal of expertise in the topic—it can take days or weeks to prepare for one where you are developing the expertise at the same time.

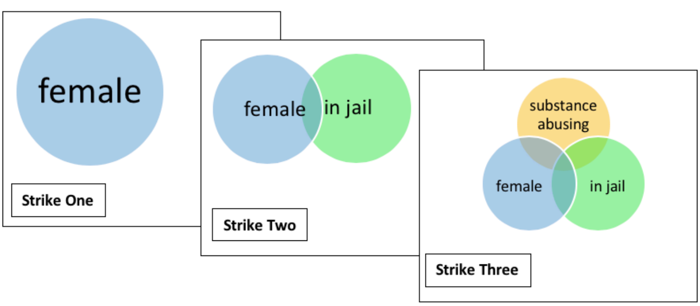
> *"It usually takes me more than three weeks to prepare a*

good impromptu speech."
–Mark Twain

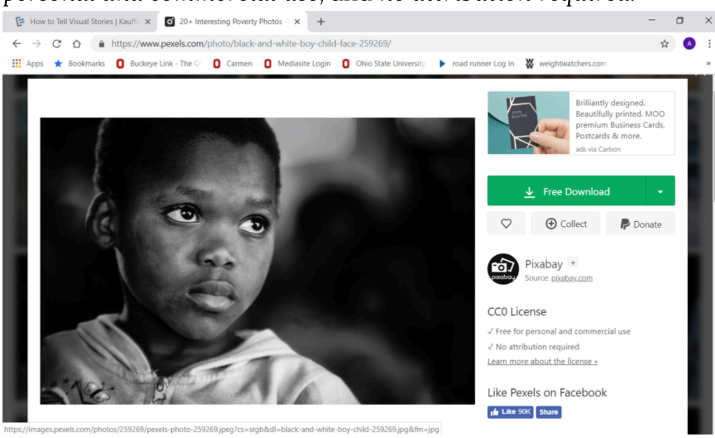
2. **A strong opening**. Audiences are most attentive in the first minute or two of a presentation. This is the time to capture attention and interest with a briefstory, anecdote, or fable to make the topic relevant; a sensitive and appropriate joke; a thought-provoking question; some "Wow!" statistics.
   For example, which lecture would you rather hear in your child development class:
   – Theories of psychomotor development during infancy, or
   – Why can't the baby ride a bike?
   The opening should be a form of "Wow!" statemen, one or two sentences—never more than three (according to Kauffman Founders School, www.entrepreneurship.org). You have learned in this course how to locate powerful, informative, "oh, wow" statistics and data—evidence-supported facts that grab attention. This is also an option for a strong opening. Or, perhaps a powerful quote—maybe from a client or study participant, maybe from history or pop culture (but be very certain about citing the person properly—there are many misquoted and miscited quotations circulating on the internet). Showing the audience an object or image is the way to capture their attention. Pro tip: if you start with a story, be sure that you end with the story's resolution.

3. **Integrate Visual and/or Auditory Elements.** According to thevisualcommunicationguy (http://thevisualcommunicationguy.com/2015/10/30/how-to-give-a-powerful-presentation-eight-steps-to-an-awesome-speech/), people recall information better when an image is attached to the message—almost twice as well as when the message is delivered in text alone. The trick with visual tools is to make them clear and simple. If the message is complex, build from a simple visual toward a complex one. Here is an example adapted from a lecture "Three Strikes: Female, In Jail, and Substance Abusing" (Begun & Rose, 2007)—the images build towards the final complex visual aid and the presenter is able to discuss what is relevant to each sphere as it is added.



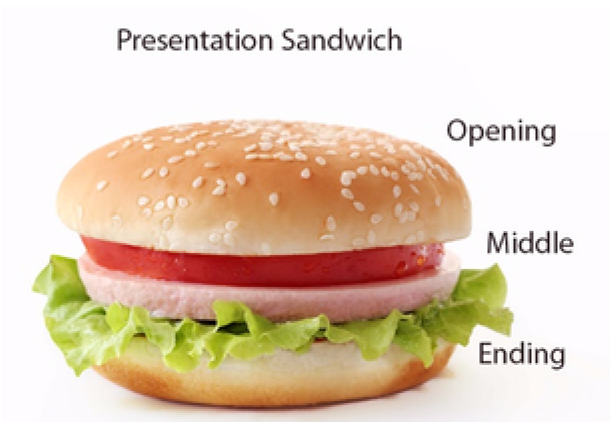
Images (pictures, charts, diagrams) may come from open access resources, or you may have to create them specifically for your message (see Chapter 3 of this module). Sources to consider include pexels.com, pixabay.com, and 123rf.com where all or a limited number of images may be downloaded for free. These are royalty-free sources where it is relatively easy to identify the copyright conditions for use of the images; capturing images from the wider internet or blogs or advertisements can be risky with regard to potential copyright infringement. Note that cartoons are usually copyright protected, as well. You want to focus your search on objects with "***Creative Commons***" status that has no restrictions or where your presentation meets the restrictions in place. In this screenshot example, see what it says under the CC0 License information on the far right: *free for personal and commercial use*, and *no attribution required.*



If you decide to incorporate audio (or video) components to your presentation, it is important to make sure the platform and facilities where you will be presenting can support the presentation. For example, ensuring sufficient connectivity if the material is being pulled from the internet and not locally embedded and that speakers and volume control are sufficient for the facility. Some agencies where you might be asked to present lack projection equipment and screens, as well.

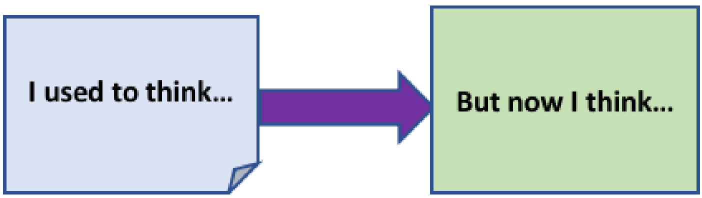
4. **Progress Logically Start to Finish.** Follow an outline of key

points and follow a consistent format for making each point. Within the first minutes, it is important to tell the audience what you are going to tell them: a statement of the presentation goal and objectives is helpful. This provides the audience with a roadmap for where you are taking them with this presentation. In the opening you tell the audience what you are going to tell them about. Next, in the middle, you tell them (the "meat" of the sandwich). Here you provide "signposts" to cue them as to where on the roadmap they are. Make clear when the topic has shifted to a new point with transition cues (new slide headings, numbered points, and transition statements). Next, see how to finish the presentation "sandwich."

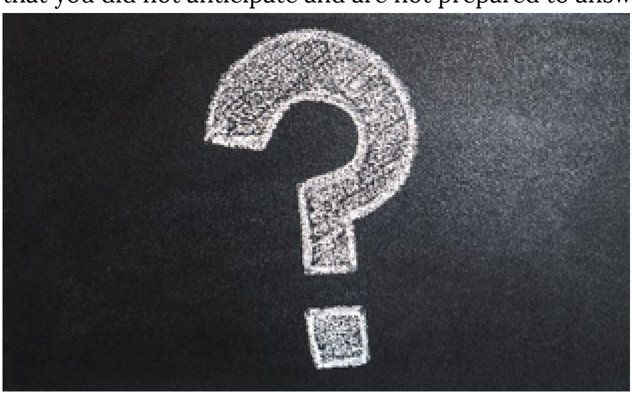


5. **Wrap up/Conclude.** In conclusion, you summarize or remind the audience of what you told them. This is not the place for new content to be introduced. If you followed a story line throughout your presentation, this is the place to deliver the "punch line" (for a joke) or the story's resolution. Or, it may be the place for a new (brief) story with a different outcome. Or, perhaps a (different) quote. It is the place to summarize the key points if your purpose was to share information, or to identify key action points if your purpose was to stimulate an action

response. You might even end with a group "evaluation" activity–a quick survey or quiz that everyone can now get correct after hearing the presentation. This is effective if you chose to open with a couple of "quiz" questions that they might not know how to best answer or where their opinions could change in the course of the presentation.

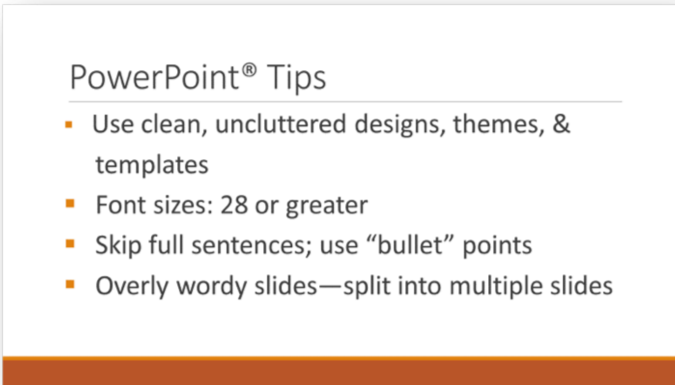


6. **Prepare for Q&A.** After a presentation concludes there may be a formal period for question and answer by the audience. Prepare for these questions–try to predict what they might ask and have a response prepared. And, have a response ready for questions that you did not anticipate and are not prepared to answer, too!



Tips for Presentation Slides

PowerPoint® is a commonly used tool for creating visual presentations. Presentations created using this software can be relatively simple or quite complexly structured. Suggestions for a strong PowerPoint® presentation are presented in this example "slide."



Create the first 3 slides of a simple PowerPoint presentation. The first should be a title slide, the second a list of either what the presentation covers (if it were a full presentation) or learning objectives for the audience to achieve from the presentation, and the third should include an image (it is the first slide of your information content). Assess your presentation slides for clarity, design, and succinct writing. Suggested topic: the most useful knowledge you have learned from this course.

Chapter Summary

In this chapter we examined the elements of strong presentations and some tools to help create and deliver strong presentations to various audiences. You were introduced to six steps for developing strong presentations, tips for using visual aids. Copyright concerns and resources for avoiding them were identified. Next, you will learn tips for creating your own visual aids (figures graphs) when just the right thing is not readily available.

Module 5 Chapter 3: Creating Figures and Graphs

Just as you learned in Module 4, information is often better understood by audience when presented in a table, figure, or graph compared to lengthy verbal descriptions. Previously, you learned how to interpret different types of graphs; in this chapter you are introduced to building them using Word®, PowerPoint®, and Excel® software. In this chapter you will learn:

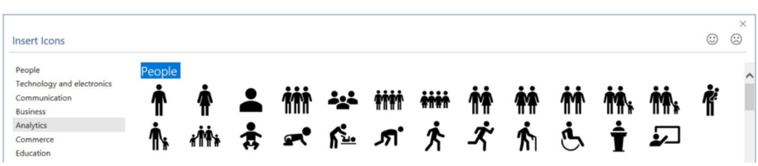
- Simple tools for creating figures;
- Tools for creating graphs.

Creating Figures

Word® and PowerPoint® have tools available for creating basic figures or diagrams. For example, the "Three Strikes" ***Venn diagram*** presented in Chapter 2 was created using the SmartArt design feature in Word®.

In the Word® top menu bar is a tab called "Insert." If you select this tab, then select "***SmartArt***" from the next menu, an array of options becomes available which are preformatted for making simple figures. It is also possible to use a combination of "Shapes" from the menu to create your own figure, but this can be more challenging, time and effort consuming, and difficult to format.

Another "Insert" option for creating your own figures is the "Icons" menu. Here is an example of the icons available under the heading of "people."

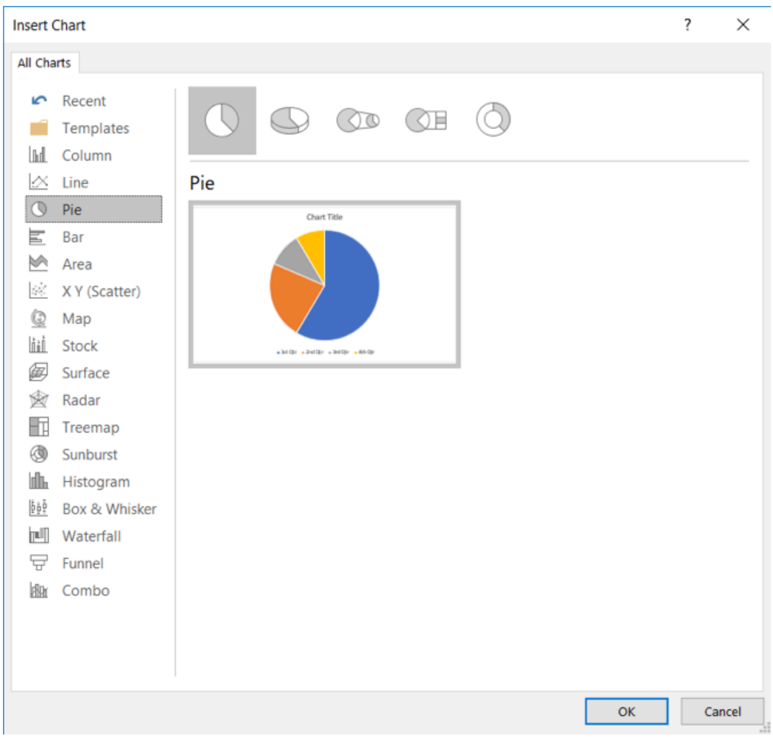


Follow the steps outlined for creating a simple figure using SmartArt. Some suggestions: a hierarchy diagram of who lives in your household (including pets), a process diagram explaining your current mood, a relationship diagram depicting your various interests, or other ideas. See what you can create on your own.

Creating Graphs

In the Word® "Insert" menu is an option called "Charts." This option provides tips and tools for creating the kinds of charts and graphs we studied in Module 4: bar charts, pie charts, scatterplots, histograms, and line charts for example. If you need charts or graphs in a PowerPoint® presentation, you can also create them the same way or you can create them in Word® and then perform a "copy & paste" operation to place them in a PowerPoint® slide. Furthermore, you can create a graph in Excel® and copy it into a Word® or PowerPoint® document, as well. The three programs are compatible in this way. In addition to these types of quantitative charts and graphs, you are introduced to one additional quantitative type (bubble charts) and one for demonstrating more qualitative types of information (word clouds).

Basic Steps. The steps for creating graph or chart are, basically, selecting "Insert" and "Charts" from the toolbar menus, select the type of graph or chart you wish to create, enter the data for the chart or graph in an Excel® spreadsheet that opens for this purpose, then fine tune the "look and feel" features of the chart or graph (labels, fonts, color scheme, proportions, and more). Here is a screen shot of the options for types of graphs or charts available in Word® and PowerPoint®.



Pie Chart

A ***pie chart*** is used to represent univariate distribution (proportion) on a categorical variable: the proportion out of 100% (the full pie) of individuals in each group on that variable. In our hypothetical example of children living in rental versus family owned homes (n=50 rental, n=128 family owned, from the “count” Excel workbook exercise in Module 4), the pie chart would be created in the following manner.

Step 1. Open the “Insert” and “Chart” menu, select “Pie” chart option and click “OK.”

Step 2. A small Excel spreadsheet box opens for you to enter the relevant information. The “Sales” column is waiting to be overwritten with your data; the 1stQtr, 2ndQtr, 3rdQtr, and 4thQtr are also place holders waiting to be overwritten with your data. Replace the number 8.2 in cell B2 with the number 50, and the cell B3 number 3.2 with 128. Then delete the rows for 3rdQtr and 4thQtr since we only had 2 categories in our variable.

Step 3. Change the place holder label “Sales” to the variable name you want to have showing in your chart (try Family Housing) and the place holder labels 1stQtr and 2ndQtr with the category names you want to have showing (try rent and own).

Step 4. Close the Excel box by clicking on the X in the upper right corner—this puts the data entry routine in the background but you will not lose your work.

Step 5. Now is the time to adjust the appearance of the pie chart. Right now, it is unclear to an audience what they are seeing. You can start by clicking on the “legend” which is the area of the chart where the color for rent and the color for own are being defined. This should open a menu to the right of the chart with four symbols.

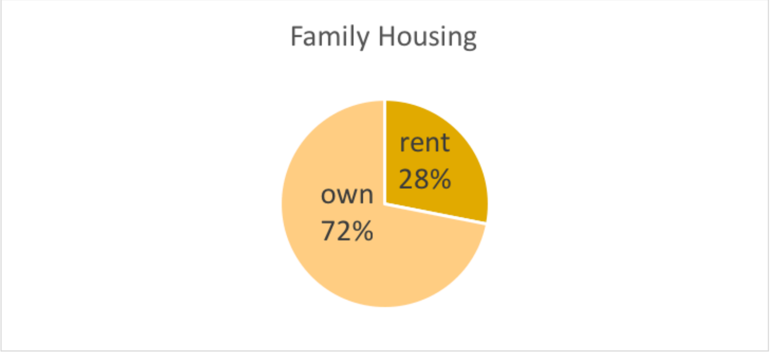
- The second symbol, a + sign in a box allows you to manipulate the labels. Select each option to see what it changes in your chart—”data labels” inserts the values of the data you entered.
- The third symbol, a paintbrush in a box allows you to manipulate the design and color scheme of your pie chart.

- The fourth symbol, that looks like a funnel in a box allows you to manipulate more about the appearance.

Additionally, if you right click on one of the pie segments, you should see a drop-down menu with one of the options being "format data labels." This allows you to select the "percentage" as what shows in the chart instead of the actual raw numbers. Selecting this should change the 50 and 128 to 28% and 72%. You could also delete the legend and check "category name" to make the chart easier to interpret. Then, you can double click the name and number in a slice of the pie chart to edit the font (different font size, bold, different color, different font) using the regular Word tools. You can also drag the box around to place it where you like.

Another option is to work with the toolbar menu running across the top of the screen once you have clicked on the chart (it will appear within a box). The Chart Tools" menu is what you want to work with in design or format. Note that there is also an option to edit the data (this will reopen that Excel box) or even to change the chart type if you decide a different type is preferable (a bar chart, for example).

Here is an example of what it might look like when you are finished–many options exist. You want to make the chart as easy to interpret as possible.



Follow the steps outlined for creating a pie chart for the children from rental or family owned homes. See what you can create on your own.

Bar Chart

A ***bar chart***, like the pie chart, is used to represent univariate distribution (frequency) on a categorical variable: the number (frequency) of individuals in each group on that variable. When the frequencies do not represent a total 100% in proportions, a bar chart is more easily interpreted than a pie chart; it is also easier to interpret a bar chart if there are more than a few categories involved–wedges of the pie might become very small and too numerous to understand easily. For example, according to the Henry J. Kaiser Family Foundation (www.kff.rg/other/state-indicator/distribution-by-age/) the U.S. population census can be reported in terms of age groups: children aged 0-18 years (24%), adults aged 19-25 years (9%), adults aged 26-34 years (12%), adults aged 35-54 years (26%), adults aged 55-64 years (14%), and those aged 65 and older (15%) as of 2017. The percentages total 100%, so it could be presented as either a pie or bar chart. If, however, the data were in the form of numbers (not percentages/proportions), a bar chart would be preferable regardless of there being more than a few categories represented,

A bar chart of this age group information would be created in the following manner.

Step 1. Open the “Insert” and “Chart” menu, select “Column” chart option and click “OK.” Note: the “bar chart” option is a sideways presentation.

Step 2. A small Excel spreadsheet box opens for you to enter the relevant information. The “Series” columns are waiting to be overwritten with your data. The series would be the 6 age groups. Categories would be relevant if you had the same series in multiple years or multiple locations, for example (1997, 2007, and 2017 perhaps). In our example, we have only the one category, so you could begin by deleting the rows for category 2 through 4 from the Excel spreadsheet.

Step 3. Enter the data into the Excel spreadsheet. Cell B2 would be the percent for children aged 0-18 (24%), C2 is 9%, D2 is 12%, E2 is 26%, F2 is 14%, and G2 is 15%. Change the place holder label “Category 1” with the variable name (try Percent U.S. Population, 2017) and the “series” place holder labels with category labels (try 0-18 yrs, 19-25 yrs, and so forth).

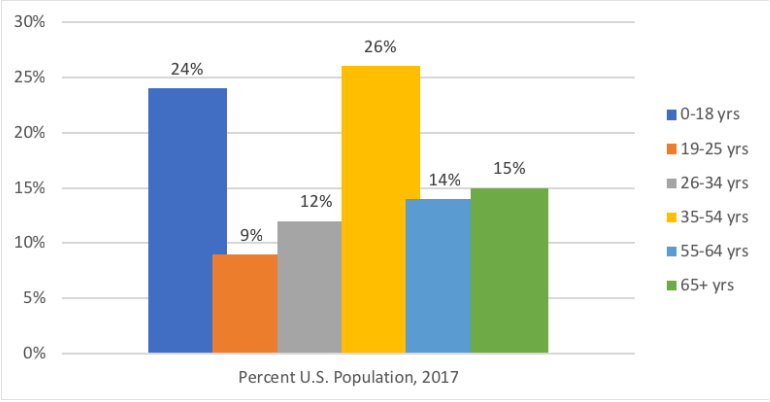
Step 4. Close the Excel box by clicking on the X in the upper right corner—this puts the data entry routine in the background but you will not lose your work.

Step 5. Now is the time to adjust the appearance of the bar chart—right now, it is a little confusing to an audience. You can start by clicking on the chart itself. This should open the “Chart Tools” menu on the upper tool bar, as well as the four symbols to the right of the chart.

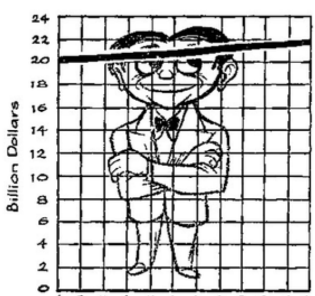
- The second symbol, a + sign in a box allows you to manipulate the labels. Select each option to see what it changes in your chart–"data labels" inserts the values of the data you entered and moving the legend to the right might help. The "format data series" option allows you to alter the spacing between the bars (columns) which can make the chart a bit more interesting.

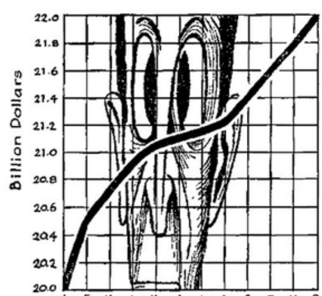
Another option is to work with the toolbar menu running across the top of the screen once you have clicked on the chart (it will appear within a box). The Chart Tools" menu is what you want to work with in design or format. Note that there is also an option to edit the data (this will reopen that Excel box) or even to change the chart type if you decide a different type is preferable (a bar chart, for example).

Here is an example of what it might look like when you are finished–many options exist. The goal is to make the chart as easy to interpret as possible.



A word of caution about bar charts: their interpretation is influenced by the ***scale*** of the y-axis. In other words, if you increase the spread of the values on the *y*-axis, it makes even small differences between columns seem impressive; if you compress the spread of values, it makes the columns seem more similar. This point is made in images from a book called *How to Lie with Statistics* (Huff & Gels, 1954); here is the authors' example demonstrating that the very same numbers (a difference of 2 billion dollars) can be interpreted as "great" or "alarming" news. Imagine the *y*-axis is about billions of dollars spent on *[insert any category of spending you like here]* during different months of a single year; the *x*-axis reflects which month the data represent. The chart on the left uses a scale where each increment represents 2 billion dollars. The movement from 20 billion to 22 billion does not look like very much; the line is only slightly angled (Huff & Gels, 1954, p. 61). The chart on the right uses a scale indicating 0.2 billion dollars for each increment (Huff & Gels, 1954, p. 63). The movement from 20 billion to 22 billion looks like a whole lot; the line is very dramatic. Yet, both charts are representing the same 2 billion dollar difference! So, the moral of this story is that scale matters in graphing.





Follow the steps outlined for creating a bar chart for the population by age group. See what you can create on your own.

Histogram

In Word® and Powerpoint®, ***histograms*** essentially are column/bar charts that include many columns–they are univariate description tools just like the bar chart. As previously explained, the histogram depicts the frequency with which each value on an interval (continuous) variable appears in the data. With enough data points, it is possible to assess how close to normal the distribution appears. Turning to the student absenteeism data with which we worked in Module 4, these are the steps involved in creating a histogram using Word.

Step 1. Open the original data file and sort the data by the variable of interest (days absent). Then, copy the column of data for that variable. This step would not be engaged if you are entering the data from "scratch" into the Excel® spreadsheet in the histogram creation routine.

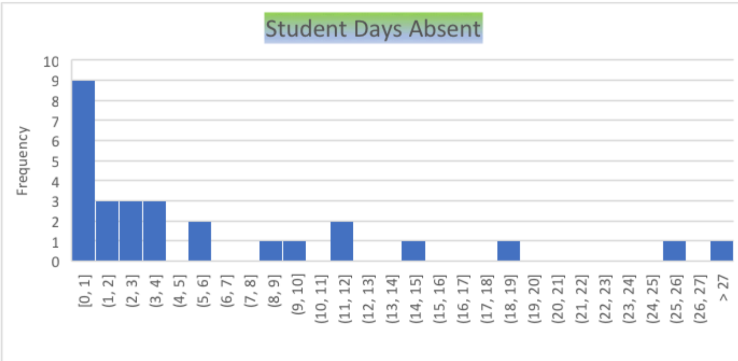
Step 2. Open the "Insert" and "Chart" tabs from the Word® menu bar.

Replace the “series” placeholder data with the data copied from your Excel® data file. Or, replace the “series” placeholder data with the data you are entering from “scratch.”

Step 3. Now you will specify the increments that you want included in each column of the histogram. The “placeholder” increments are “chunking” your data in increments of 12. However, you want a more refined picture. To change the increments, double click on the area of the graph where [0,12] (12,24] and so on appear (horizontal axis). This should open a “format axis” menu to the right of your work. In Excel® the increments are called “bins” and you want to adjust the bin width. If we change the 12.0 in the Bin width box to 1, the chart shows the frequency with which each value for number of days appears among our 28 students–since the lowest value was 0 and the highest value was 39, the result is a total of 49 bins (increments) in our chart. This seemed to be confusing because of that one outlier at 49 days absent. Checking “overflow bin” in the format axis menu and changing the default value (38) to 27 retains that one student in the picture but doesn’t spread the axis out as much: the student is in the >27 bin (all values greater than 27). Then, the format axis menu can be closed.

Step 4. The final steps relate to the chart’s appearance. Double clicking on the Chart Title placeholder allows overwriting of those words with a meaningful title. It also opens a Format Chart Title menu where color, position, and other aspects can be modified. Clicking on the chart opens the three symbols menu where axis titles may be added.

Here is an example of how a histogram of these data created in Word® might appear.



Another option is to create the histogram in Excel® and perform a copy & paste operation into the Word® document or PowerPoint® presentation slide.

Follow the steps outlined for creating a histogram for the student absenteeism data (located in the Excel file called student absenteeism start.xlsx). See what you can create on your own.

Scatterplot

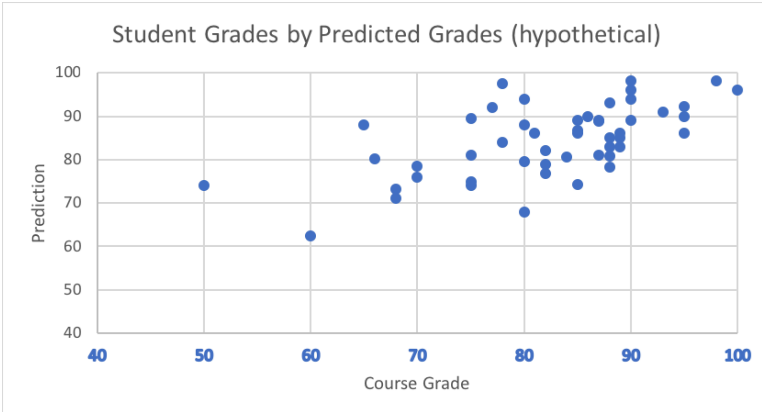
As previously discussed, a ***scatterplot*** is used to visually depict the relationship between two variables (bivariate). Using a hypothetical set of data from 50 undergraduate social work students in a research and statistics course, a scatterplot is created depicting the relationship between their predictions about their course grades generated at the start of the course and their actual final course grades.

Step 1. Open original data file (unless entering data from "scratch" into the chart development routine); copy each column and replace the "placeholder" values in the "Scatterplot" Excel® spreadsheet that opens with "Insert" and "Chart" selected from the toolbar.

Step 2. Double click on the chart to access the formatting tools–for example, the format axis tool allows you to adjust the "y" axis values ("Number" option) to remove the irrelevant decimal place; it allows you to adjust the scale of the "y" axis, as well. Not only does it allow adjusting of the scale (e.g., 10-point increments versus 20-point increments), it allows deletion of the bottom or top range where there are no values in the data–instead of starting at 0.0 (no one predicted a score this low), the scale can start at 40 instead. Instead of extending to 120 (the highest possible value was 100%), it can end at 100.

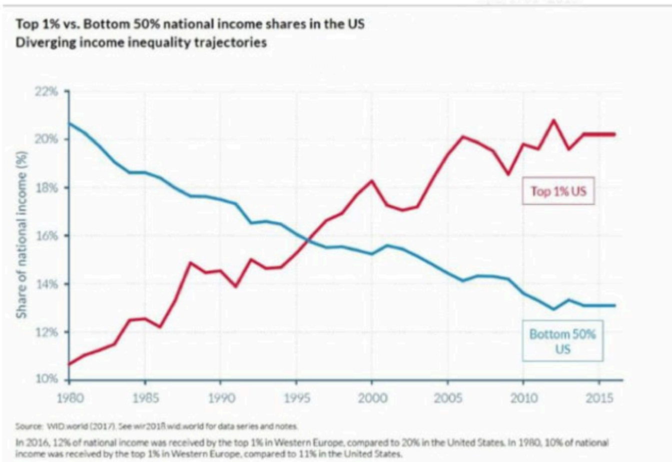
Step 3. Formatting the "x" axis allows similar functions. These formatting tools also allow insertion of labels for the "x" and "y" axis. Changing the chart title is likely to be necessary. If the axes are labelled, the "legend" can be deleted.

Here is how a scatterplot of these data might appear.



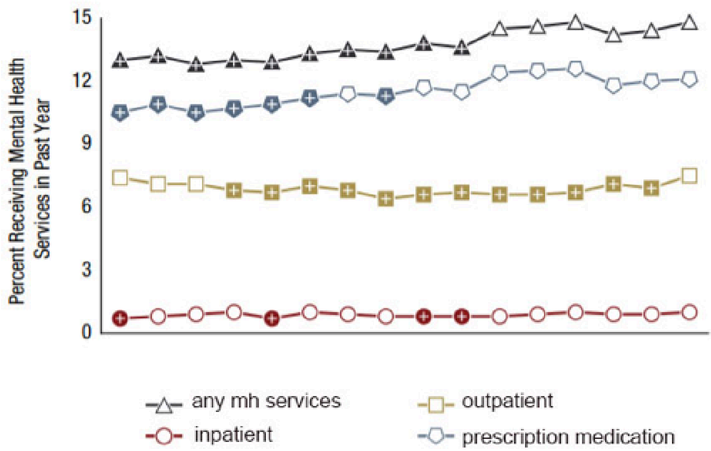
Line Graph

In many instances the point that a presenter wishes to make relates to a trend in data, for example how some feature has changed over time. Here is a modified ***line graph*** with two lines showing the trend in income shares between 1980 and 2016 (see the World Inequality Report, Alvaredo et al., 2018, p. 8). In this chart, you can quickly see that households in the bottom 50% of income level have declined in their share of income and households in the top 1% have climbed. In other words, the income inequality gap increased considerably.



These charts can be created using the same "Insert" and "Charts" menu we have been working with so far. The choice is "Line" in the menu of options. The data are inserted into the drop-down Excel® worksheet, just as in the previous chart examples. Editing the features of the chart are the same as in previous examples, as well. The income inequality line graph shows an interaction between the two lines, one for each group. Line charts can be as simple as tracking one variable at two time-points or can be very complex, showing a trend for multiple groups. Here is an example from the National Survey on Drug Use and Health (SAMHSA, 2017). This figure depicts the percent of adults receiving different forms of mental health services each year from 2008 to 2017. The legend at the bottom shows the different types of services—a separate line is presented for each. Not only can you see the trend for each type, you can see the types in relation to each other (e.g., inpatient services are less common than prescription medicine).

Figure 72. Type of Mental Health Services Received in the Past Year among Adults Aged 18 or Older: Percentages, 2002-2017



+ Difference between this estimate and the 2017 estimate is statistically significant at the .05 level.

Note: Mental health service is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health.

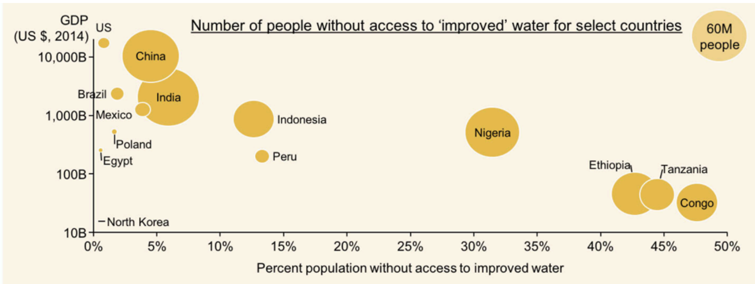
Bubble Chart

A type of chart with a different emphasis is called a ***bubble chart***. It is basically a type of scatterplot, but in a bubble chart the "dots" (bubbles) vary in size. The size of the bubbles represents an added dimension of information. As in the scatterplot, we still have an *x*-axis and *y*-axis. In a bubble chart, there is a 3rd dimension (variable) in the diagram. The 3rd dimension is represented by the size (area) of the bubble, or sometimes by colors. Bubble charts do not work if there are too many bubbles (cases) represented (see https://datavizcatalogue.com/methods/bubble_chart.html).
Microsoft hosts a site where you can learn more about how to use Excel to create a bubble chart: https://support.office.com/en-us/

article/present-your-data-in-a-bubble-chart-424d7bda-93e8-4983-9b51-c766f3e330d9

Here is an example of a bubble chart based on National Geographic data, posted on the website www.aploris.com/blog/charts/category/bubble-chart/. The *x*-axis represents the percent of population in each country who lack access to water from pipes or well protected from contamination. The *y*-axis represents the GDP (gross domestic product, indicator of national wealth) in billions of U.S. dollars for each nation in the sample. The size of each bubble reflects the number of persons affected in each nation (this is different from percent and reflects a 3rd variable).

You can see that relatively wealthy nations, like the U.S. have very small bubbles far on the left of the graph (very small percent of population affected). Four African nations have the largest percent of population without access to water from protected sources (Congo, Tanzania, Ethiopia, and Nigeria) and these are also nations in the lower quadrant which indicates they have lower GDP. Nigeria has a larger sized bubble than Tanzania, for example, indicating that the problem affects a larger number of persons in the nation, even though a larger percent of the Tanzanian population is affected. The "legend" on this chart shows the size of a bubble representing 60 million persons, allowing for comparisons on this dimension. The size of the bubble for the U.S. is very small compared to this 60 million persons bubble, whereas China and India have larger bubbles, even though their percent of affected population is near the U.S. value.



Source: National Geographic (2015 data)

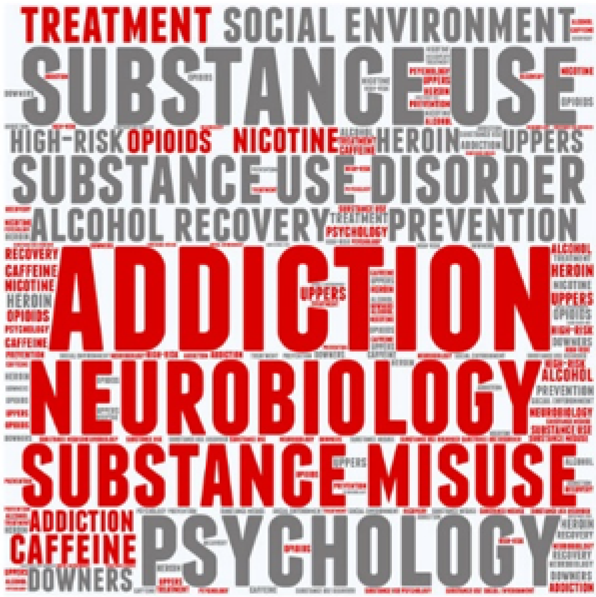
Note: 'Improved' water is water from pipes or wells that are protected from contamination

Word Cloud

You have seen them: Word Clouds. A ***word cloud*** is a graphic image composed of a cluster of words that fit together under a general theme or topic. A resource for learning to create them using Word® software with an add-in from Microsoft Office Store (created by Orpheus Technology, called Pro Word Cloud) is available at https://www.youtube.com/watch?v=my1JRX84tyc, and there exist several free software programs for creating them, as well. They also can be created in PowerPoint® with an add-in, or you could create one in Word® that you then copy and paste into a presentation slide.

A word cloud begins with a set of words. These can be entered as a list or a body of text can be copied into the program. They can be created with frequency of each word's appearance determining its positioning and prominence. Common words like "a," "the," "and," can be ignored in the options, and other features can be modified (color, orientation, and others).

Here is an example of a word cloud used as an online textbook cover for a different course.



Take a moment to complete the following activity.

An interactive or media element has been excluded from this version of the text. You can

view it online here:
https://ohiostate.pressbooks.pub/swk3401/?p=601

Chapter Summary

In this chapter you were introduced or re-introduced to 7 types of visual graphs and charts that are useful in communicating information to an audience. Not only did you refresh your skills at interpreting them yourself, you also learned how to set about creating them. These tools can be easily embedded in a written report or presentation. The next chapter describes another communication tool gaining in popularity for communicating powerful messages: the infographic.

Module 5 Chapter 4: Creating Infographics

Unlike the research manuscript, professional report, or presentation, an ***infographic*** presents data and implications in a brief but impactful format. The challenge involves making complex information easily understood by diverse audiences. Infographics can be incorporated into presentations, but are often free-standing tools. As the producers of the venngage blog suggest, "they also make dry information more engaging" (https://venngage.com/blog/create-infographics). As social work professionals, you may elect to present empirical evidence in this format, depending on the audience and purpose of the communication effort. In this chapter you will learn:

- What infographics are and when they might be useful;
- Steps in creating an effective infographic;
- Tool for developing infographics;

Infographics Described

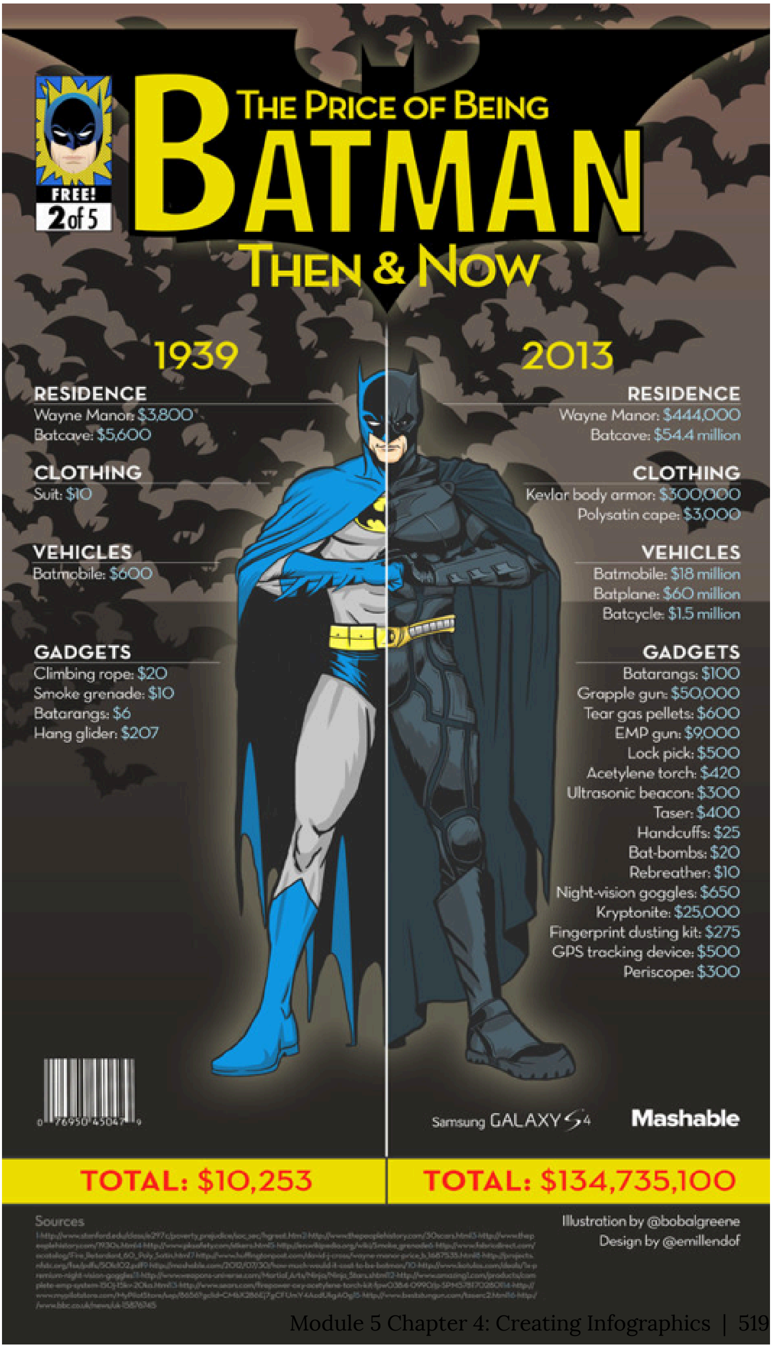
An infographic is a brief informational graphic tool (usually 1 page or computer screen shot). According to the venngage blog introduction to infographics:

> *They are often used to visualize data, step-by-step processes, timelines, research findings, organizational hierarchies, and many other types of information* (https://venngage.com/blog/create-infographics).

Infographics tend to be more engaging to audiences than lengthy text—like reading a book with pictures often is more interesting than reading a book with only words. Imagine reading Hawthorne's *The Scarlet Letter*as a high school student with pictures to make the characters and their situations come alive for you, compared to reading the text alone. This helps explain the popularity of graphic novels.

Infographics tap into a fact long recognized in psychology: images are more easily interpreted and recalled than the written word (Paivio, 1971). For example, thenextweb blog presented the example below where you can quickly see how costs (of being a superhero) exploded between 1939 and 2013. Imagine if you had to read through this information in text form?

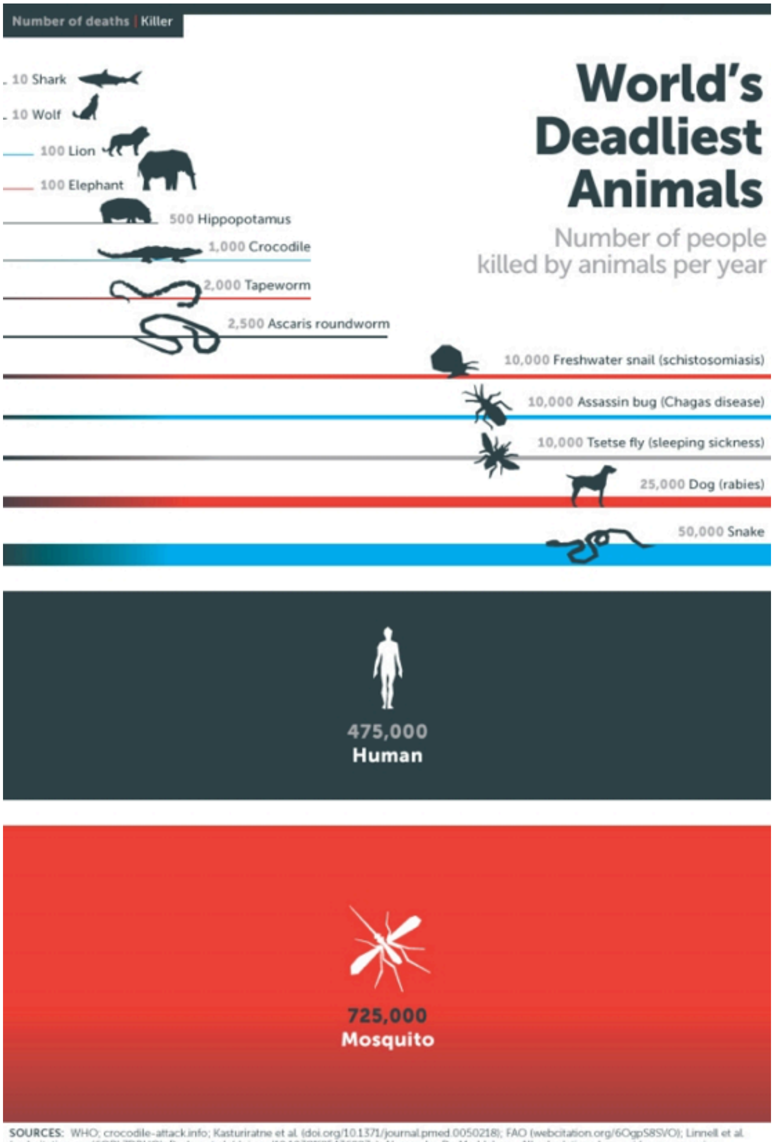
https://thenextweb.com/wp-content/blogs.dir/1/files/2013/10/batman-then-and-now.jpg


FREE!
2 of 5
THE PRICE OF BEING
BATMAN
THEN & NOW
1939
RESIDENCE
Wayne Manor: $3,800
Batcave: $5,600
CLOTHING
Suit: $10
VEHICLES
Batmobile: $600
GADGETS
Climbing rope: $20
Smoke grenade: $10
Batarangs: $6
Hang glider: $207
2013
RESIDENCE
Wayne Manor: $444,000
Batcave: $54.4 million
CLOTHING
Kevlar body armor: $300,000
Polysatin cape: $3,000
VEHICLES
Batmobile: $18 million
Batplane: $60 million
Batcycle: $1.5 million
GADGETS
Batarangs: $100
Grapple gun: $50,000
Tear gas pellets: $600
EMP gun: $9,000
Lock pick: $500
Acetylene torch: $420
Ultrasonic beacon: $300
Taser: $400
Handcuffs: $25
Bat-bombs: $20
Rebreather: $10
Night-vision goggles: $650
Kryptonite: $25,000
Fingerprint dusting kit: $275
GPS tracking device: $500
Periscope: $300
Samsung GALAXY S4
Mashable
TOTAL: $10,253
TOTAL: $134,735,100
Sources
Illustration by @bobalgreene
Design by @emillendof

Steps in Creating an Infographic

The following presents 4 steps in creating an infographic for communicating with your audiences.

Step 1. Translate your topic into a burning question that interests your audience. As an example, venngage blog (https://venngage.com/blog/create-infographics) highlights the use of an infographic by the Bill and Melinda Gates Foundation to generate interest in a topic of little interest to many people in Northern climates: malaria. They framed it in terms of the world's deadliest animals—if you are a fan of "shark week," have a fear of crocodiles, snakes, spiders, or creatures, or are concerned about interpersonal violence rates, this infographic might peak your interest.



Number of deaths | Killer
World's Deadliest Animals
Number of people killed by animals per year
10 Shark
10 Wolf
100 Lion
100 Elephant
500 Hippopotamus
1,000 Crocodile
2,000 Tapeworm
2,500 Ascaris roundworm
10,000 Freshwater snail (schistosomiasis)
10,000 Assassin bug (Chagas disease)
10,000 Tsetse fly (sleeping sickness)
25,000 Dog (rabies)
50,000 Snake
475,000
Human
725,000
Mosquito
SOURCES: WHO; crocodile-attack.info; Kasturiratne et al. (doi.org/10.1371/journal.pmed.0050218); FAO (webcitation.org/6OgpS8SVO); Linnell et al. (webcitation.org/6ORL7DBUO); Packer et al. (doi.org/10.1038%2F436927a); Alessandro De Maddalena. All calculations have wide error margins.

Step 2. Gather the information to be presented in the infographic. Most likely, you will collect more information than is needed. You will need to be selective and/or creative in how you combine different pieces of information. However, it is also critically important that you maintain truthfulness in your selection and presentation of information. You also need to cite your sources!

Step 3. Design the infographic. One feature to keep in mind as you design an infographic is the medium through which it will be distributed. If it will be presented in print, you may have the luxury of using both sides of a page; if a poster or electronic "screen" is the medium, you are limited to one side/one page. Various software packages are available, many for free, to help in the design process (e.g., Canva Infographic Maker; Easel.ly; Piktochart; Snappa; Venngage; Visme; Visualize). Most of these are free to use in their "base" form; advanced options are available with a fee. Another option is to design the tool "from scratch" using either Word® or PowerPoint® with or without design templates. Remember to be concise, precise, and accurate in what you write. Consider image and graph/chart tools previously discussed, as well.

Step 4. Produce the infographic in your selected medium (print, electronically, or both) and disseminate your information to your audience.

Spend some time in a "scavenger hunt" viewing and critiquing infographics that others have developed. Consider what seems to work and what does not for the different examples you locate. Use your observations to generate a set of guidelines for developing your own infographics (e.g., layout, color, design, content, writing style).

Chapter Summary

Infographics as a powerful communication tool were introduced. You learned a bit about why they are effective means of communicating with diverse audiences, and you learned 4 steps in their production. In addition, several free tools were identified for assisting in their design.

Module 5 Chapter 5: Course Conclusions

This concludes the reading activities for the first course in our two-course sequence, SWK 3401. In these five modules you covered a great deal of information related to the course title: *Research & Statistics for Understanding Social Work Problems and Diverse Populations*. Here is a brief synopsis of what you learned in each module.

Module 1.The Importance of Research to Social Work Practice; Research Ethics

The first course module provided an orientation to the important relationship research and evidence play in social work practice and promoting social justice, including the profession's mandates regarding the place for research in social work. This orientation analyzed different types and sources of knowledge ("ways of knowing") used by social workers in developing an understanding of social work problems, diverse populations, and social phenomena. It also provided a context for understanding the place of research and empirical evidence in the wider realm of knowledge we use. The module presented principles of critical thinking engaged by social workers, and presented key distinctions between science, pseudoscience, and opinion. Within this orientation, the philosophical roots of different approaches to evidence were identified (positivism, empiricism, and ethnoepistemology in quantitative and qualitative research), as were relationships between theory, research, and practice. Strategies for locating sources of empirical evidence were explored in this first module. Finally, a great deal of attention was directed to the ethical conduct of research and important considerations for research involving human participants.

Module 2. Identifying and Applying Evidence

Our second course module extended what was learned in the first by developing a framework for understanding different types of research questions and study specific aims. The distinction between research questions and research hypotheses was explained, as were differences between background/foreground, exploratory, descriptive, and explanatory research questions. The relationship between types of questions and scientific approach was addressed. A translational science framework explained differences between basic, intervention, and implementation science, as well as "research for and about research." These categories were developed as strong influences on all subsequent decisions made by investigators. This module analyzed the role of theory in research and identified ways of identifying and assessing empirical evidence regarding social work problems, diverse populations, and social phenomena. Furthermore, a great deal of emphasis was placed on evaluating empirical literature and distinguishing between empirical and non-empirical literature. In the course of completing Module 2, you were introduced to working with Excel spreadsheets.

Module 3. Research Approaches, Designs, and Methods

Our third course module extended from the background presented in Modules 1 and 2 to explore specific approaches, designs, and methodologies used in qualitative, quantitative, and mixed methods research. The relationship between research questions and research approaches was elaborated in this module with an overview and details for each approach (qualitative, quantitative, mixed methods). Six major traditions in qualitative research were explored: narrative, phenomenological, grounded theory, ethnographic, participatory action, and content analysis research. Design issues of generalizability, internal and external validity, cross-sectional and longitudinal study were examined with regard to quantitative research traditions. Distinctions between descriptive, correlational, and experimental research were identified, as was the distinction between correlation and causality. Delving deeper into quantitative methods, a great deal of attention was directed to recognizing different types of variables and their implications for study design, measurement, and analysis (demographic, independent, dependent, categorical/nominal, ordinal, interval/continuous). Measurement issues in quantitative science were addressed, including measurement validity and reliability, as well as cultural competence in measurement approaches. Different strategies for collecting qualitative and quantitative data were explored: naturalistic observation, Photovoice, artifact/content analysis, administrative and secondary data, geographic information systems (GIS), key informant information, surveys, clinical screening or assessment tools, interviews, focus groups, journaling, ecological momentary assessment (EMA), concept maps, and social network analysis. The final methodology chapter was devoted to study participants. Distinctions were made between samples and populations, and key issues related to the recruitment and retention of study participants were explored. Not only were matters of sample size and diversity addressed, strategies for participant selection were analyzed (random selection as probability sampling, convenience and snowball selection as non-probability sampling). In the course of completing Module 3, your introduction to working with Excel spreadsheets and data files was extended.

Module 4. Understanding Descriptive and Inferential Statistics

The fourth module in this course was devoted to approaches used in analyzing data collected from the methods described in Module 3—qualitative, quantitative, and mixed methods analyses. In the chapter concerning qualitative analysis, you were introduced to data preparation, field notes, coding, coding confirmation, and cross-checking coding decisions. The subsequent chapters introduced quantitative analysis strategies and issues. These included distinguishing between univariate and bivariate analyses, use of inferential statistics, the role and interpretation of a null hypothesis, Type I and Type II error, and different types of analyses related to the nature of the variables and the research questions. You learned to conduct and interpret univariate descriptive analyses involving the frequency, proportion, mean, median, mode, variance, standard deviation, probability, and normal distribution of data. In addition, you learned to conduct and interpret five types of bivariate analyses: one-sample *t*-test, independent samples*t*-test, one-way analysis of variance (Anova), chi-square, and correlation analysis. The idea behind non-parametric analysis and two examples were also developed in this module. In the course of studying this content, you learned to conduct different types of analyses using Excel spreadsheets and data files.

Module 5. Presenting Evidence

By now you may recognize that the course has been structured, loosely, along the structure of an empirical report: an introduction (Module 1 and Module 2), Methods (Module 3), Results (Module 4), and Discussion (Module 5). Module 5 itself concerned ways that social workers communicate with various types of audiences about empirical evidence and scientific data concerning social work problems, diverse populations, and social phenomena. The module discussed important aspects of empirical manuscripts and professional reports, features of strong presentations, creating figures (graphs and charts), and creating infographics. With this concluding chapter of Module 5, you are now well-prepared to succeed in meeting the goals of the second course in our two-course sequence. In SWK 3402 your skills and knowledge will be applied to understanding social work interventions. Together the content in these two courses prepares you to engage effectively with evidence in social practice, regardless of the nature of practice in which you engage.

What do you believe was the single best thing you learned about in this course? Consider sending your answer to your instructor and/or including it in your course evaluation comments.

Module 5 Key Terms and Definitions

abstract: a brief, comprehensive survey summarizing contents of an empirical manuscript or professional report, presented at the front of a manuscript.

bar chart: a figure used in graphically reporting univariate statistics for a categorical variable (frequency count for each category represented in a bar or column).

bubble chart: a form of scatterplot chart (two variables) with a third variable reflected in the size of the data points (dots or bubbles).

citation: identification of the source of information used in report, manuscript, or presentation, allows audiences to evaluate the information and/or locate the original source for themselves.

confused verb tense: a problematic writing practice where two or more different verb tenses are used in the same sentence, clause, or paragraph.

Creative Commons: a public copyright license whereby the creator of a freely distributed work specifies how that work can be shared, used, or modified by others.

discussion: final section in the body of an empirical manuscript or professional report where results are briefly summarized, conclusions and interpretations are presented, strengths and limitations are analyzed, and implications of the results are presented.

findings: like study results, the findings might be presented in a professional report.

histogram: a figure used to graphically represent the frequency with which each value for a variable is represented in the data; reflection of data distribution for a variable.

infographic: a clear, concise visual tool for demonstrating a key point of information, data, or knowledge.

initial or executive summary: similar to an abstract, a summary of the contents of a professional report presented at the front of the report.

introduction: opening section of an empirical manuscript or professional report that introduces the topic/problem, surveys and critiques existing literature, and states study aims/research questions/hypotheses.

line graph: a figure used to graphically represent the relationship between two variables (the "x" and "y" axes), often used to depict trends over time (time as the "x" axis) or other demonstrations of change.

methodology: detailed description in an empirical manuscript or professional report that describes how a study or project was conducted and sources/methods of data collection; includes descriptions of study approach/design, study participants, and measurement tools/procedures.

parallel construction: a writing technique where two or more sentences, clauses, phrases, paragraphs, or sections have the same grammatical structure.

paraphrase: using different words to reflect the meaning of someone else's ideas, writing, or verbalized speech, usually to either clarify or summarize what was originally stated.

pie chart: a figure used to graphically represent the univariate statistic of proportions for a categorical variable (proportion for each category represented in a "slice" or "wedge" of the pie).

plagiarism: using someone else's words or ideas without properly identifying them as belonging to that person rather than being one's own.

put people first: a "use of language" practice using descriptors that begin with the person rather than a condition, circumstances, or label.

recommendations: part of the final discussion in a professional report where conclusions are offered in terms of advisable actions or next steps based on the project findings.

reference: full information about a resource, usually one cited in a manuscript or report, providing sufficient detail that readers can evaluate the source and/or locate the source for themselves.

results: section of an empirical manuscript or professional report where data analysis procedures and their outcomes are described.

scale: the units used in a graph, chart, or other figure where the distance between points has systematic meaning in numerical terms.

scatterplot: a figure used to graphically represent the relationship between two variables (the "x" and "y" axes), often used to depict the strength and direction of their correlation or association.

SmartArt: a tool available in Word® and PowerPoint® with ready-made templates for creating figures.

Venn diagram: a type of figure depicting the existence or extent/degree of overlap between different groups, constructs, themes, ideas, or events where what is shared in common is represented in the area of overlap; usually constructed from circles, ovals, or spheres, but may be formed from other geometric shapes.

Module 5 References

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2018). *World inequality report: Executive summary*. https://wir2018.wid.world/files/download/wir2018-summary-english.pdf

Anderson-Butcher, D., Wade-Mdivanian, R., Davis, J., Newman, T., & Lower, L. (2016). *2015-2016 Ohio State's LiFE Sports annual review*. Columbus, OH: Ohio State University College of Social Work.

Begun, A.L. (2016). Considering the language we use: Well worth the effort. *Journal of Social Work Practice in the Addictions*, 16(3), 332-336.

Begun, A.L., Early, T.J., & Hodge, A. (2016). Mental health and substance abuse service engagement by men and women during community reentry following incarceration. *Administration & Policy in Mental Health & Mental Health Services Research*, 43(2), 207-218.

Begun, A. L., & Rose, S.J. (2007).*Three strikes: Substance abusing, female, and in jail.*University of Wisconsin-Milwaukee Morris Fromkin Memorial Lecture, Milwaukee, WI.

Broyles, L.M., Binswanger, I.A., Jenkins, J.A., Finnell, D.S., Faseru, B., Cavaiola, A., … Gordon, A.J. (2014). Confronting inadvertent stigma and pejorative language in addiction scholarship: A recognition and response. *Substance Abuse*, 35, 217-221.

Bunger, AC, Robertson, H, Hoffman, J, Cao, Y. (June 2015). *Connecting youth in child welfare to behavioral health services in Franklin County: Provider perspectives. Evaluation report written for Franklin County Children Services* (OH). Columbus, OH.

Doogan, N.J., & Warren, K.L. (2017). Saving my life: Dynamics of peer and staff corrections among therapeutic community residents. *Substance Use & Misuse*, 52(11), 1429-1438.

Huff, D., & Gels, I. (1954). *How to lie with statistics.*NY: W.W. Norton & Co.

Johnson-Motoyama, M., Moses, M., Kann, T. K., Mariscal, E. S., Levy, M., Navarro, C., & Fite, P. (2016). Parent, teacher, and school stakeholder perspectives on adolescent pregnancy prevention programming for Latino youth. *Journal of Primary Prevention*, 37(6), 513-525.

Kagotho, N., Nabunya, P., Ssewamala, F., Mwangi, E. N., & Njenga, G. (2017). The role of family financial socialization and financial management skills on youth saving behavior.*Journal of Adolescence*, 59, 134- 138.

Lee, M.Y., Wang, X., Liu, C., Raheim, S., & Tebb, S. (2018). Outcome literature review of integrative body-mind-spirit practices for mental health conditions. *Social Work Research*, 42(3), 251-266.

Maguire-Jack, K., & Campbell, T. (2015). Evaluation of Marathon County Community Response Program: Report to Marathon County Department of Social Services. Columbus, OH: Ohio State University College of Social Work.

McInroy, L.B. (2017). Innovative ethics: Using animated videos when soliciting informed consent of young people for online surveys. *Social Work Research*, *41*(2), 121-128.

Mendoza, A.N., Fruhauf, C.A., Bundy-Fazioli, K., & Weil, J. (2018). Understanding Latino grandparents raising grandchildren through a bioecological lens. *International Journal of Aging and Human Development*, 86(3), 281-305.

Mengo, C., & Black, B.M. (2015). Violence victimization on a college campus: Impact on GPA and school dropout. *Journal of College Student Retention: Research, Theory & Practice*, 18(2), 234-248.

Mengo, C.W., Lee, J., Bolton, K., Lehmann, P., & Jordan, C. (2017). Self-identified strengths among youth offenders charged with assault against a non-intimate family member. *Journal of Human Behavior in the Social Environment*, 27(8), 881-893.

Paivio, A. (1971). *Imagery and verbal processes*. NY: Holt, Rinehart, & Winston.

Parekh, R., Maleku, A., Fields, N., Adomo, G., Schuman, D., & Felderhoff, B. (2018). Pathways to age-friendly communities in diverse urban neighborhoods: Do social capital and social cohesion matter?*Journal of Gerontological Social Work*, 61(5), 492-512.

Quinn, C.R., Liu, C., Kothari, C., Cerulli, C., & Thurston, S. (*in press*). Using social determinants of health to assess psychological distress and suicidal thoughts and behavior among youth on probation. *Adolescent Psychiatry*, 7.

Substance Abuse and Mental Health Services Administration (SAMHSA). (2018). Key substance use and mental health indicators in the United States: Results from the 2017 national Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. Retrieved from www.samhsa.gov/data/sites/default/files/cbhsq-reports/NSDUHFFR2017/NSDUHFFR2017.pdf).

Yoon, S., Voith, L.A., & Kobulsky, J.M. (2018). Gender differences in pathways from child physical and sexual abuse to adolescent risky sexual behavior. *Journal of Adolescence*, 64, 89-97.