# PSYC 6190 Lecture Notes

Moses Rivera, Ph.D.

# Table of contents

# Preface

This is the collection of Lecture Notes written by Moses Rivera, Ph.D., for the PSYC 6190 course: *Research Methods*.

You can download the entire collection in the upper-left of any page.

All items in the *References* section of each chapter are available for you to read in our shared Zotero Group Library.

# 1 Introducing: Research Methods

## 1.1 Overview of this course

Via this course, I'm giving you a beginner's toolbox for you to do research. Like any beginner toolbox, it doesn't include *everything* you'll ever need, but it's enough to get you going.

It can take years to master research methods. In fact, you could spend a lifetime learning about research methods. Some experts dedicate their entire careers to creating new research methods (e.g., new statistical methods). One semester isn't enough for you to master all of the typical research methods that are used by **industrial-**, **work-**, and **organizational-psychologists** (**IWO Psychologists**). However, in this course, I will give you a tour of the research methods landscape.

Instead of asking you to read a research-methods book that contains many sections that are irrelevant for IWO Psychologists, or outdated, or poorly-explained, I will give you lecture notes (you are reading a lecture notes file right now). You can download a PDF version of each lecture notes file. Also, I will give you supplemental reference materials that you can use during your career whenever the need arises. I recommend you keep all of these resources handy (e.g., in your cloud storage, or your computer, and/or your Zotero library).

## 1.2 What are research methods?

The online *Oxford Dictionary of English* (2025) defines *research* as:

> the systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions.

In a nutshell, research is *investigation*, also known as *inquiry*. Anyone who does research is typically trying to arrive at a conclusion, such as making a decision or solving a problem.

In your personal life, you've already done a lot of research. For example, when you needed to buy something and you compare the features of different options, you were doing research.

Obviously, some research methods are useful for some specified purpose, whereas other research methods aren't so useful. This course focuses on the ***scientific* research methods** that IWO Psychologists typically use. In a subsequent lecture, I will explain what makes a research method *scientific*.

For now, let's just say a research method is *scientific* if it is systematic, logical, and tied to observable reality.

## 1.3 What do IWO Psychologists use research methods for?

IWO Psychologists use scientific research methods to understand things and to have a real impact on real workers. The emphasis on scientific research methods is what typically distinguishes IWO Psychology versus Human Resources (HR). Most HR degree programs haven't focused on training their students to acquire scientific research methods, but that's starting to change as HR personnel are becoming more aware about the usefulness of scientific research methods.

IWO Psychologists can use scientific research methods for many purposes. To give you an idea, here are some examples:

- To understand the people in this specific team and/or company/organization.
  - e.g., their attitudes, beliefs, abilities, habits, etc.
- To understand any people in teams/organizations in general.

- To evaluate the validity of a hiring procedure (e.g., is our hiring procedure effective?).
- To evaluate the impact that a training & development system will likely have, or the impact it already had on some workers.
    - To evaluate the effectiveness of a leadership-development program.
- To help design new systems and interventions that will be very likely to achieve their intended purposes (e.g., increase worker wellbeing, performance, company revenue).

I could go on and on, but you get the point.

## 1.4 The design of this course

This course was designed in collaboration with several professors—especially I–O Psychology professors—with several important considerations.

First, we consulted SIOP's *Guidelines for Education and Training in Industrial-Organizational Psychology* (2016, available in our Zotero Group Library). Those *Guidelines* describe the topics that IWO psychologists generally consider important to include in any educational sequence for students who are studying to work in IWO psychology.

Next, we (the professors) pooled all of our collective knowledge, together with several popular textbooks about research methods, and we designed a curriculum (i.e., a sequence of topics) for this course. This course's curriculum is fairly standard for any first-semester graduate student studying any behavioral science (including psychology), with a few important distinctions:

**SIOP** is the acronym of the **Society for Industrial and Organizational Psychology**, which is Division 14 of the **APA** (**American Psychological Association**). SIOP is the primary professional association for IWO psychology in the USA.

- The latter part of this course includes methods about **causal inference** (i.e., inferences about how one thing causes another thing), particularly some methods that haven't yet permeated across the landscape of IWO psychology. Such methods were developed by many scholars—especially from the mid-1990s to the present day—but it's fair to say **Judea Pearl** was and still is a primer driver. When you ask your colleagues about such methods, don't be surprised if they have no clue what you're talking about.

- The early parts of this course will explicitly identify several habits and assumptions that are prevalent among researchers in IWO psychology. By labeling those assumptions and tendencies, we seek to give you an awareness and self-awareness that should result in better research.

- Lastly, this course includes a brief introduction to **qualitative research methods**, which are often completely absent from research methods courses. Such methods are becoming increasingly recognized for their usefulness in the scientific endeavor.

**Judea Pearl** is a professor of computer science at the University of California, Los Angeles.

Aside from that, the remainder of this course is very typical. Here is a birds-eye view of the curriculum:

- Introduction to the fundamental concepts about scientific research methods in general, and how they relate to IWO psychology.

- Ethics in psychological research. Many research methods courses insert this near the end of the course, but we're putting it front-and-center near the beginning, because of its fundamental importance to all research.

- Overview of qualitative and quantitative research methods.

- Literature-review methods. This is how you figure out what has already been researched and concluded in relation to your topic(s) of interest.

- Building a theory, model, and hypotheses. This is very typical in IWO psychology, which is why this course will emphasize these topics.

- Sampling, measurement, and psychometrics. This degree program includes a required course on psychometrics, so we will just introduce you to the main topics in this course.

For now, let's just say **qualitative research methods** are contrasted against **quantative research methods**. The latter focus on numbers, whereas the former focuses on other qualities that aren't numeric.

- A brief introduction to statistical methods. This degree program includes a required course on statistical methods, so we will just give you a brief introduction in this course.
- Classic and cutting-edge methods for causal inference.
- How to disseminate the results of research.

## 1.5 Conclusion

Now that you've finished reading this page, you should go back to Canvas to see what the next steps are in this course.

# 2 IWO Psychology as a Science

## 2.1 Science

The online *Oxford Dictionary of English* (2025) has a good definition for *science*:

> the systematic study of the structure and behavior of the physical and natural world through observation, experimentation, and the testing of theories against the evidence obtained:
> "*the world of science and technology*" · "*science laboratories*"
>
> - a particular area of science:
>   "*veterinary science*" · "*the agricultural sciences*"
> - a systematically organized body of knowledge on a particular subject:
>   "*the science of criminology*"

As you can see, the word **science** can refer to any and all of these:

- any **(scientific) process** for studying something (e.g., *"IWO psychologists use science to study people in the context of work"*);
- any **topic** that is studying via a scientific method (e.g., *behavioral sciences*);
- any **collection of scientifically-derived knowledge** on a topic (e.g., *"the science of personality indicates that people are complicated…"*).

What makes science so special is its epistemology. A good definition of **epistemology** is:

> The theory of knowledge and understanding, esp. with regard to its methods, validity, and scope, and the distinction between justified belief and opinion; (as a count noun) a particular theory of knowledge and understanding.
>
> (Oxford English Dictionary, 2025)

In a nutshell, there are many different epistemologies. For example, some people rely on others' *authority* to figure out what's true. Others may rely on *consensus* to figure out what's true. Others may rely on *intuition* to figure out what's true. And still others may rely on *logic* to figure out what's true. However, scientific methods emphasize **systematic empirical observations** as an important basis to figure out what's true. Scientists still use logic, consensus, intuition, and (sometimes) authority to figure out what's true, but the feature that distinguishes science is systematic empirical observations.

An observations is **empirical** if it is based on and/or verifiable by experiences, rather than only theory or logic. Therefore, *systematic empirical observations* are repeatable, reasonable, logical, and they're based on lived experiences.

Despite rigorous research methods, any scientifically-derived knowledge is almost always tentative and subject to revision based on new empirical results. In other words, scientifically derived-knowledge is still often *fallible*. Thus, researchers—including those in the *behavioral sciences* such as IWO psychology—tend to "hedge" their conclusions by using language such as: "the results *suggest* that…", "the results *may indicate* that…" "the results *provide support* for…".

## 2.2 IWO Psychology as a science

For each topic that is scientifically studied (e.g., biology, physics, psychology, etc.) the scientists who study those topics use special tools and methods—and the same is true for IWO psychology. To help contextualize the scientific approaches that are typically used in IWO psychology, I'll now introduce a couple more -*ologies*: *axiology* and *praxiology*. **Axiology** is:

> The study of the nature of value and valuation, and of the kinds of things that are valuable.
>
> - A particular theory of axiology
>
> (Oxford Dictionary of English, 2025)

When we talk about the axiology of IWO psychology, we're talking about the IWO psychologists' typical principles or beliefs about what they generally consider to be important. In other words, we're talking about the *values* of IWO psychologists.

Relatedly, the **praxiology** (also spelled *praxeology*) of IWO psychology is the set of *actions* that IWO psychologists generally *do* in their praxis.

Obviously, the axiology and praxiology of IWO psychology evolves over time, and it can never be described perfectly because it is a generalization about all IWO psychologists. However, in the following sections, we can try to describe its historical development and current state.

**Praxis** is also known as *practice*, but it doesn't mean *beginner training.* Rather, praxis is like the word *practice* in *medical practice*, *clinical practice*, or *IWO psychology practitioner.* Scholars often use the word *praxis* instead of *practice*, to disambiguate between those two difference senses of *practice.* So when we say *praxis*, we're referring to the application of some skill or ability, as a practitioner would.

### 2.2.1 The axiology and praxiology of IWO psychology

There are several excellent summaries of the historical development of IWO psychology in the USA as a field of science and praxis. Consider reading Salas et al. (2017) and Porter and Schneider (2014). For the enthusiast, Cascio and Aguinis (2008) and Katzell and Austin (1992) are other good (though older) resources.

In a nutshell, IWO psychology in the USA has historically been dominated by an axiology that values quantitative rigor and a balance between science and praxis—hence the classic term: the **scientist–practitioner** approach in IWO psychology. Historically, the praxiology of IWO psychologists has prominently featured the study of people situated in real work-contexts (though laboratory studies have also played an important role). Such studies typically involve a hypothesis and a quantitative analysis of data in an attempt to evaluate the truth-value of the hypothesis via falsificationism, and hopefully benefit the company and workers being studied and/or other workers more broadly. All of this is quite standard among many scientific disciplines.

My goal in mentioning all of that is to present to you a vocabulary that can enable you to think about and talk about the way IWO psychologists do their work today. Equipped with such language, you can have greater acuity in your analyses of the work of other IWO psychologists, and you can have greater self-awareness of your tendencies and preferences in doing IWO psychology work.

*Stance* can be a good word to describe an overall orientation or approach for inquiry. For example, someone could have a critical stance, or a practical stance, or a combination of several stances. For an excellent introduction to the typical stances that researchers adopt, see Craig (2009). For brevity, here are some of the most commonly discussed stances:

- **Empirical-explanatory stance**: This stance is probably what you were taught in science class in grade school. In this stance, researchers typically seek to explain phenomena in an objective manner that is devoid of judgment about whether those phenomena are "good or bad", and they typically do that by creating testable hypotheses and analyzing empirical data.
- **Practical stance**: The main distinguishing features of the practical stance are its axiology and praxiology. It seeks to not only understand and explain phenomena, but also to take action to improve the human condition by solving concrete problems. One popular method is known as *participatory action research*, in which the intended beneficiaries of some solution are involved as partners in co-creating those solutions with the researchers. The practical stance can accommodate a wide variety of epistemologies (e.g., quantitative, qualitative, deductive, inductive, etc.).
- **Critical stance**: The main distinguishing features of the critical stance are its axiology and praxiology. A core aim of the critical stance is to emancipate society from conditions that perpetuate injustice (e.g., unjust power relations)—in other words, justice is a common feature of their axiology. Persons operating in the tradition of the

critical stance tend to be politically active (i.e., that's a common feature of their praxiology). The critical stance can accommodate a wide variety of epistemologies (e.g., quantitative, qualitative, deductive, inductive, etc.).

- **Interpretive stance**: According to this stance, it is important that the significance of human action be interpreted from the meaning that humans attribute to their own actions. Thus, the focus isn't so much on general cause-mechanism theories that are devoid of rich context. A prominent feature of its axiology is the intrinsic worth of understanding the culture and history of the persons being studied.

Traditionally, the field of IWO psychology has operated with a blend of the empirical-explanatory stance (to understand what's going on) and the practical stance (to help *do* something good about what's going on). However, this was primarily due to the culture of the researchers that dominated the field of IWO psychology in its early days in the early-to-mid 1900s. Often, their goal was to understand humans in the context of work, so that their performance could be enhanced (typically for the benefit of the company). Compared to those early days, today there are many more IWO psychologists that skillfully adopt a critical stance and/or interpretive stance, and other stances. Likewise, many IWO psychologists today have a much broader set of topics they focus on, including the wellbeing of workers for its own sake (not simply as a stepping stone toward greater profits), and the social responsibilities of companies.

I encourage you to explore these and other stances for inquiry, and discover which stance(s) resonate with you, and have confidence in the power of being flexible to adopt different stances when they are warranted.

### 2.2.2 The epistemology of IWO psychology

Most likely, when you were in grade school, you were taught in science class to specify a hypothesis and then test it via an empirical study. That approach is known as **hypothetico-deductivism**. *Deductivism* comes from the word *deductive*, which is often used to describe a type of reasoning known as

**deductive reasoning**: any process of reasoning in which an inference is made about a particular instance of something, based on general principles that transcend that instance. In contrast, **inductive reasoning** is any process of reasoning in which an inference is made about general principles, based on particular instances. Deductive and inductive reasoning is perhaps easiest to explain via examples, which I've provided in Table 2.1.

Table 2.1: Examples of deductive and inductive reasoning

| Deductive reasoning | Inductive reasoning |
| --- | --- |
| We're fairly certain that flexible scheduling reduces turnover across the organization; therefore, any department that switches to flexible schedules will likely see a reduction in voluntary turnover. | In three departments that switched to flexible schedules, voluntary turnover fell noticeably; therefore, flexible scheduling likely reduces turnover across the organization. |
| We're fairly certain that structured interviews produce performance-ratings that are more accurate at our company; therefore, for the next hiring season, structured interviews should result in higher-accuracy supervisor-rated performance than unstructured ones. | During this hiring season, structured interviews resulted in higher-accuracy supervisor-rated performance than the unstructured ones; thus, structured interviews likely have better predictive validity in our company. |

Hypothetico-deductivism is the research approach in which the first step is to specify a theory about how stuff works or what stuff is like. Then, from that theory, we use deductive reasoning to create an empirically-testable proposition, which we call a *hypothesis*, to test our theory. Then, we analyze some empirically-derived data to test the hypothesis, as a way to test our theory. Technically speaking, there is often some inductive reasoning involved when you use your analyses to make an inference about whether your hypothesis and/or theory is true—

but we still call the entire approach *hypothetico-deductivism* because of the focal step of deductively creating a hypothesis from a theory.

Another important way to think about the epistemology of IWO psychology is in terms of *confirmatory* versus *exploratory* modes of research. In **confirmatory research**, we seek to test a theory, to figure out whether it is true or not. In **exploratory research**, we seek to identify patterns in the data, without focusing on testing any specified theory. Both confirmatory and exploratory modes of research are useful, although you may notice a lot of IWO psychology research has traditionally been confirmatory (probably because of how those researchers were taught to do research).

In subsequent lectures in this course, I will guide you through a more formal definition of *theory*.

#### 2.2.2.1 Falsificationism

Scientists typically seek to make universal statements that are true. A **universal statement** is any expression that says a specified property is true for *all elements* in a specified set. For example, each of the statements in Table 2.2 is a universal statement because they each make a claim about all elements in a specified set.

A **set** is any collection of things (e.g., a set of tangible objects, a set of emotions, a set of abstract ideas, etc.). The things inside a set are called the **elements** of that set.

Table 2.2: Examples of universal statements

| Universal statement | The (implied) universal set |
|---|---|
| Dogs are better than cats. | All dogs and all cats. |
| My dogs are better than my cats. | All my dogs and all my cats |
| Workers who are at or higher than 5/10 on the conscientiousness scale have higher salaries than those who are below 5/10. | All workers. |
| Engineers at Company ABC who are at or higher than 5/10 on the conscientiousness scale have higher salaries than those who are below 5/10. | All engineers at Company ABC. |

As you can see from the first two examples in Table 2.2, not all universal statements are useful, nor true. Scientists typically seek to make universal statements that are true, and some scientists seek to make universal statements that are true and useful at the same time.

The theories and hypotheses of IWO psychologists are typically universal statements. We test universal statements by trying to find systematic empirical evidence that disproves them. Why? It's because universal statements are often difficult or impossible to *prove true*, but they can be quite easy to *disprove* (i.e., prove them false). The process of trying to disprove a statement is called *falsification*, and **falsificationism** is the epistemological belief that the best way to evaluate the truthiness (or, more formally, the *truth-value*) of a universal statement is to seek and evaluate evidence that seems to disprove it.

Specifically, most of the universal statements that scientists seek to make are about sets that contain such a large number of elements that it would be practically and/or literally impossible to observe every element of those sets to verify the truth of any universal statement about those sets. For example, the third statement in Table 2.2 is about *all workers*. To *prove* that the universal statement is true for all workers, we would need to empirically observe the conscientiousness and salaries of all workers. Every worker that has ever lived in the entire world. Obviously, that's impossible. However, we can very easily *disprove* the universal statement if we find at least one worker with conscientiousness below 5/10 whose salary is higher than another worker's whose conscientiousness is at or above 5/10. Obviously, this example is a bit simple and therefore a bit silly, but this powerful principle is still true for universal statements that are much more realistic and useful, such as this one: *"A worker's conscientiousness tends to correlate moderately with their salary"*. If that statement is truly seeking to make a claim about all workers everywhere, a good way to evaluate its truth-value is to try to disprove it.

Obviously, if we try our best to falsify a universal statement, but all of those attempts fail to disprove it (i.e., the universal statement appears to be true via all of those tests), we haven't proven the statement true (unless we were somehow able to test all elements in the universal set). Therefore, when all of our tests fail to disprove a universal statement, all of that evidence merely lends more support and/or credibility to the belief that the statement is actually true. Thus, when we specify a hypothesis in the form of a universal statement, and our best empirical methods failed to falsify the hypothesis, we don't say the theory was proven true. Rather, we say things like: "the results *suggest* that the theory is true…", or "the results *indicate* the theory *may* be true…", or "the results *provide support* for the theory…".

## 2.3 A brief introduction to the types of scholarly works in IWO psychology

In subsequent lectures in this course, we will dive into understanding the different types of scholarly works that IWO psychologists produce. At this point in this course, it would be useful for you to have a basic familiarity with these types:

- **empirical study**: any study that involves the analysis of empirical data.
- **theory paper**: any written work that focuses on theories, rather than analyzing any empirical data. For example, it may describe, compare, and/or critique one or more theories.
- **lab study**: any study that involves the analysis of empirical data that was collected in a laboratory setting (e.g., a IWO psychologists research lab at a university).
- **field study**: any study that involves the analysis of empirical data that was collected about people behaving in their natural settings, such as a real workplace, rather than in a laboratory. Even if a laboratory is decorated to look like a natural setting, it's not considered a field study.
- **literature-review study**: any study that focuses on analyzing multiple other published studies, typically to summarize and synthesize the current state of the science on some topic, but without doing a formal analysis of the aggregated data from those other studies. Typically, a

literature-review study will culminate in a *literature-review paper*.

- **meta-analysis study**: similar to a literature review study, except a meta-analysis must involve a systematic approach for identifying the other studies that will be included or excluded from the meta-analysis, and the meta-analysis must also involve a systematic analysis of the aggregated data across all of the studies that were included in the meta-analysis. Most often, meta-analyses are quantitative, but they can also be qualitative, and/or mixed-methods.

## 2.4 A brief note on pre-registration

**Pre-registration** is a process or result of producing publicly verifiable time-stamped evidence related to a research endeavor (e.g., "*she is pre-registering her hypothesis*", "*I'm reading his pre-registered plan of analysis*"). The most commonly pre-registered items are hypotheses and plans for how the data will be analyzed (but you could theoretically pre-register anything you want, such as a written statement of your underlying theories and/or your expected implications of your study).

One of the primary purposes of pre-registration is to prevent some types of **HARKing**: **H**ypothesizing **A**fter the **R**esults are **K**nown. HARKing is the act of creating a hypothesis after you've already analyzed the data and seen the results about that hypothesis. For example, suppose I collect some data about worker motivation and salary, and then I analyze the data and find that motivation is positively correlated with salary in my dataset. At that point, if I hypothesize that motivation is positively correlated with salary, I've just HARKed.

Scholarly discussions about HARKing distinguish between **SHARKing**—**S**ecretly HARKing—versus **THARKing**—**T**ransparently HARKing (Hollenbeck & Wright, 2017). HARKing isn't necessarily a problem. THARKing is definitely not a problem, and in many cases it is arguably *unethical* not to THARK (Hollenbeck & Wright, 2017). But SHARKing is definitely unethical. For a scholarly analysis of the various problems caused by SHARKing, see Aguinis and Murphy (2019), Hollenbeck and Wright (2017), and Kerr (1998).

Pre-registering your hypotheses and other aspects of your research plan is a great way to establish evidence that you're being honest and transparent. That's important for increasing other people's trust in science—especially for gaining the trust of the general population that might be skeptical about science. Many scholarly journals are increasingly requiring pre-registered study materials for papers that wish to be published in those journals.

In a subsequent lecture in this course, I will describe a popular website for pre-registering your studies, called the Open Science Framework https://osf.io.

> **A deeper dive if you're curious…**
>
> Almost all theories and their subsequent hypotheses are based on our observations of the world around us. Therefore, you might be thinking: "*aren't all hypotheses ultimately created after we already know some results about the world around us?*" This brings up an important distinction. Let's suppose you analyze a dataset containing data about two variables: worker motivation and salary. If you analyze the data and find that motivation is positively correlated with salary—before you ever made any hypothesis about their correlation—you would be SHARKing if you pretend that you hypothesized a positive correlation before analyzing the data. However, if you see the correlation is positive in the first dataset and then you hypothesize that the positive correlation will occur in a new dataset from a different group of workers, and then you collect that new data and find the correlation is also positive in that dataset, then that's good science! The difference is we used a different set of data to test our hypothesis that the correlation would be positive among other workers.

## 2.5 Suggested Readings

The only readings I suggest on the topics in this lecture are the ones I've already cited in these lecture notes, which you can see in the *References* section immediately below. All of those are available in our shared Zotero Group Library for this course.

# 3 Ethics in Psychological Research

## 3.1 Introduction

These lecture notes on ethics in psychological research focus primarily on the APA's ethics code (2017). Although the APA does strive to have an international presence, other countries may have specific laws and/or regulations about the praxis of psychological research. Therefore, if you want to work outside of the USA, you may need to supplement your knowledge beyond these lecture notes.

Also, the final semester of this degree program includes a course that will focus on ethics in IWO psychology. At that point, we will do a deeper dive into ethics related to psychological research and praxis. For now, we are merely introducing you to fundamental ideas that would be good for you to know now. As you likely know by now, repeated exposure is very beneficial for your long-term retention of information.

## 3.2 A brief USA history on ethics in psychological research

Since before recorded history, researchers have been doing research on humans—also known as *research on human subjects* or *human subjects research*.

If you look at the history of ethics in psychological research in the USA, you can see a clear trend toward greater and greater respect for the wellbeing of the research participants, but it's taken many years for the professional ethical standards of psychologists to get to where they are today.

As you might already be familiar with, there are many examples of early research in the 1960s and 1970s that would not be allowed by today's professional ethical standards for psychologists. Notable examples include the *Tuskegee Syphilis Study* from the 1930s thru 1970s, Stanley Milgram's electric shock studies of the 1960s, and Philip Zimbardo's *Stanford Prison Experiment* in 1971. Surprisingly, these studies occurred even though the first version of the APA's Ethics Code had already been published in 1953.

Most students of psychology learn about a historically significant milestone known as the Belmont Report. The **Belmont Report** (1979) was a document that was officially published in 1979 by the *U.S. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research*. It was created as a response to controversial and ethical problems about the Tuskegee Syphilis Study. One of the major problems with the Tuskegee Syphilis Study was that the researchers were studying the effects of syphilis on adults, but the researchers weren't informing the participants there was already a widely available and standard treatment for syphilis since 1947 (which was the antibiotic named *penicillin*). Thus, the researchers allowed the participants and their families to suffer with syphilis without proper treatment.

Although that research was conducted by biomedical professionals (not psychologists), I'm mentioning all of this because it was a major historic event that eventually caused the production of the Belmont Report which was one of the first widely-known reports that established the need for a formal ethical system to guide the manner of conducting a research study. Importantly, the Belmont Report highlighted three ethical principles:

1. respect for persons;
2. beneficence;
3. justice.

Although the first edition of the APA's ethics code was published in 1953, it didn't receive nearly as much attention as the Belmont Report. However, today arguably every researcher of psychology is familiar with many of the ethical standards in the APA ethics code.

## 3.3 The APA Ethics Code

The current version of the APA ethics code (2017) is formally titled the *Ethical Principles of Psychologists and Code of Conduct*—but many people shorten it to the *APA ethics code*. It has been updated over the years, and the most recent major revision was in 2002, followed by minor amendments in 2010 and 2016.

The APA ethics code applies to anyone who identifies as a professional in psychology, by virtue of the activities the person performs. The APA ethics code doesn't explicitly define who is a psychologist, but rather it is *implicitly* defined via examples of the types of activities that a professional in psychology typically performs. Also, the *Introduction* section of the APA ethics code states: "*Membership in the APA commits members and student affiliates to comply with the standards of the APA Ethics*".

The APA ethics code formally recognizes five **General Principles**:

1. beneficence and nonmaleficence
2. fidelity and responsibility
3. integrity
4. justice
5. respect for people's rights and dignity

The APA ethics code is organized into 10 sections that contain the enforceable obligations for psychologists. You will learn about those 10 sections in a later course in this degree program, but for now we will introduce you some of the important concepts in *Section 8: Research and Publication*.

### 3.3.1 Informed Consent to Research

According to the APA ethics code, psychologists who wish to do any of a variety of activities—such as assessments, consulting, counseling, research, and/or therapy—must receive informed consent from prospective participants and/or clients (though there are a few exceptions where informed consent isn't required, such as when a law or court order requires the psychologists to do something). **Informed consent** is *permission* that the participant gives to the psychologist, after the participant is well-informed about what they're consenting to.

According to standard 8.02(a) of the APA ethics code (2017), if the psychologist wants to recruit a potential participant for a research study, the psychologist must provide all of the following information to the potential participant so that the potential participant can give informed consent:

1. the purpose of the research, expected duration, and procedures;
2. their right to decline to participate and to withdraw from the research after participation has begun;
3. the foreseeable consequences of declining or withdrawing;
4. reasonably foreseeable factors that may be expected to influence their willingness to participate, such as potential risks, discomfort, or adverse effects;
5. any prospective research benefits;
6. limits of confidentiality;
7. incentives for participation;
8. whom to contact for questions about the research and the research participant's rights.

Additionally, during the informed consent process, the psychologist must provide the opportunity for the potential participant to ask questions and receive answers.

### 3.3.2 Deception in research

In the context of human-subjects research, **deception** is the process or result of misleading a research participant, especially for a purpose that is useful for the goals of the study.

According to standard 8.07 of the APA ethics code (2017), psychologists are allowed to use deception in a research study, but only when the situation meets ALL of these criteria:

- the use of the deception is justified by the study's significant prospective scientific, educational, or applied value, AND effective nondeceptive alternative procedures are not feasible, AND the study is reasonably expected to NOT cause physical pain or severe emotional distress.

Psychologists are never allowed to use deception in a research study if the research activity "is reasonably expected to cause physical pain or severe emotional distress."

If a psychologist does use deception in a research study, they must eventually reveal the deception to the participants who were deceived—this is known as **debriefing**. Specifically, "they explain any deception that is an integral feature of the design and conduct of an experiment to participants as early as is feasible, preferably at the conclusion of their participation, but no later than at the conclusion of the data collection" (American Psychological Association, 2017, standard 8.07).

### 3.3.3 Informed Consent for Recording Video, Audio, and/or Photos

According to standard 8.03 of the APA ethics code (2017), prior to recording audio or video of a research participant (e.g., for data-collection) the psychologist must obtain the informed consent from the participant (or a legally authorized person) to record them. However, the following two situations are allowed to be exceptions:

- The data-collection is targeted only toward *naturalistic observations in public places*, AND the researcher expects the recordings will only be used in a manner that *doesn't cause the people to be personally identified nor harmed*.
- The research design includes *deception*, AND the consent for the use of the recording is obtained during debriefing.

## 3.4 IRB in research on human-subjects

*IRB* is an acronym for **Institutional Review Board**. An IRB is a committee that reviews, approves, and monitors research. One of the main goals of IRBs is to protect the rights, welfare, and privacy of human-subjects in research. In the USA, it's probably safe to assume every university has an IRB that is part of that university. There are also IRBs that are part of government agencies, and also commercially private IRBs that provide services for researchers who aren't affiliated with any institution that already has an internal IRB.

If you are a student at a university and you wish to do research on human-subjects, you almost always need to get approval from your university's IRB first (many universities grant exceptions, such as when the research is for a class project led by an instructor). The approval process involves sending to the IRB a report that describes your proposed research (often, researchers call that report *"the IRB"* or *"the IRB paperwork"*, e.g., *"Did you do the IRB?"*).

There are many nuanced rules about what types of research can be quickly approved by your university's IRB, but that is beyond the scope of this course. For now, I just want you to know what an IRB is and why they exist.

## 3.5 Suggested Readings

If you want to learn more, the APA ethics code is actually a short read. If you're mostly curious about its ethical standards about research, you could skip to *Section 8: Research and Publication*. You can find the APA ethics code easily on the APA's website, and I've also placed a copy of the PDF version into our Zotero Group Library.

# 4 Qualitative versus Quantitative Methods, and Mixed Methods

## 4.1 Introduction

In Section 2.2.1, I briefly mentioned that the historical preference among IWO psychologists for quantitative methods (rather than qualitative methods) likely stemmed from the prevailing culture of influential persons in the domain of IWO psychologists during its early days. Over the decades there's been increasing interest and appreciation for the usefulness of qualitative methods and mixed methods. There isn't enough time to go in-depth into qualitative and/or mixed methods research in a single semester together with quantitative methods, but I aim to at least provide you with a sufficient introduction to such methods. The interested reader who wants to learn more may consult the Suggested Readings section of this chapter.

## 4.2 What are quantitative, qualitative, and/or mixed methods?

If you search for definitions for *quantitative research method* and/or *qualitative research method*, you will likely soon realize there is no universal consensus about how to define nor distinguish between those two. Yet, many researchers feel they have an intuitive understanding about how those two are different from each other. Interestingly, a definition of *mixed methods* research is perhaps easier to arrive at: it is research that combines qualitative and quantitative methods. Now, all we need to do is gain some clarity on what we mean when we say *qualitative* or *quantitative* methods.

Ultimately, the best place to start is with the meaning of *quantitative* and *qualitative*. Consider the following definitions from the online *Oxford Dictionary of English* (2025):

**quantitative** (adj.)

> "relating to, measuring, or measured by the quantity of something rather than its quality."

**qualitative** (adj.)

> "relating to, measuring, or measured by the quality of something rather than its quantity."

Now let's try to unpack the meaning of *quantitative research method*, versus *qualitative research method*. A *method* is a process, and a *process* is merely a sequence of actions or changes across time. Consider a scenario in which I give you instructions for how to calculate the sales tax for an apple (based on the sale price and the tax rate), and another scenario in which I give you instructions for how to mix yellow and red paint to create orange. In both scenarios, you go through a process. In both scenarios, you use your hands to either use a calculator or write some numbers on paper or a computer, or you use your hands to maneuver the paint brush. In both scenarios, you use your eyes and your mind to keep track of what you're doing, and what the next step should be. In both scenarios, you reach an end-point.

However, an important difference between the two scenarios is the *type of information* you were operating on and producing. In the first scenario, you were combining numbers to produce another number, also known as *numeric information*, also known as **quantitative data**. In the second scenario, you were combining colors to produce another color, which is a type of *qualitative information*, also known as **qualitative data** (e.g., your eyes would notice if there was too much yellow, and your mind would conclude you need to add more red, and then re-evaluate to figure out whether you've arrived at orangeness).

It is unfortunate that scholars typically use the term *qualitative* to mean *non-numeric*. Perhaps a more straightforward meaning would be, *"of or relating to any quality"*. In the latter sense, a *numeric quality* is still a quality (e.g., the size of a house in squared feet), and we could therefore map *numeric data* and *non-numeric data* onto *quantitative* and *qualitative research*, respectively. However, we are unlikely to change overnight how scholars around the world use these

Thus, one useful perspective to help define *quantitative* versus *qualitative research methods* is to consider whether the data being analyzed and/or produced are *numeric* or *non-numeric*. However, that's not the only important consideration in defining those terms. Often, each qualitative method and/or qualitative method is bundled with an epistemological stance, an axiological stance, and a praxiological stance. For example, when scholars describe *ethnography* as a qualitative method, they don't merely mention the fact that it involves observations and interviews. Rather, they emphasize the attitudes and mindset that the researcher must have, as well as the appropriate manners with which to interact with any participants. All of that is part of the *ethos* of ethnographic research. Ethnographic researchers typically maintain an open mindset towards exploring phenomena they didn't initially anticipate, and they often seek to understand by adopting the perspective of the persons they are studying. Such an orientation is often lacking in many quantitative methods.

Although there is no single universally agreed-upon definition nor typology of qualitative versus quantitative methods, I hope I've at least given you a conceptual framework to help you understand what scholars typically mean. With that, I will now turn to describing some of the most common types of qualitative methods you may encounter in IWO psychology.

**ethos** (n.): "the characteristic spirit of a culture, era, or community as manifested in its beliefs and aspirations." (Oxford Dictionary of English, 2025)

## 4.3 Some of the most common types of qualitative methods

A complete list of all types of qualitative methods is likely impossible, because there are so many ways to categorize the methods. For example, some researchers consider *ethnography* as a topic to be studied (i.e., the study of ethnic groups and cultures), whereas other researchers consider *ethnography* as a set of research *methods*. Here's a different example: many researchers use *quantitative* methods to study one individual person, and they call that a *case study*. However, other researchers also use the noun-phrase *case study* as a type of *qualitative* method. As you can see, there is no universal consensus on what these words mean, but you'll eventually figure out what the people in your area of the world mean when they use those words.

Still, here I will briefly describe some common types of qualitative methods. Just keep in mind other scholars may adopt slightly different definitions for some of the labels. What's important here is that you understand the underlying concepts in each type of method.

**Narrative research** focuses on stories. Those stories can be factual, fictional, or somewhere in the between (e.g., where we're not sure whether the story is fictional). Often, the stories are about the lives of one or more persons, but the stories can also be about non-human things, such as animals, spirits, or other things in the natural or supernatural world. For a brief introduction into narrative research methods, see Clandinin and Caine (2008).

**Case study research** is an intensive investigation about a bounded "case" (e.g., person, team, organization, event) in its real-world context. It draws on multiple sources of evidence (e.g., documents, interviews, observations, quantitative data) to produce a rich description or explanation. The cases may be single or multiple, and the findings may be generalized into a broader theory. Of course, case studies can also be purely quantitative, but here we are describing qualitative case studies. For a brief introduction into qualitative case study research, see Blatter (2008).

**Phenomenology** (also known as phenomenological research) seeks to understand one or more *lived phenomena* across multiple persons, to distill its essential structures or meanings. Examples of *lived phenomena* are: what it is like to fall in love; how students acquire a self-identity as emerging experts in their domain; or what it is like be taking a research methods course. Data typically come from rich first-person accounts, and the analysis typically seeks to identify themes that characterize the lived phenomena being studied. For a brief introduction into phenomenology, see Adams and van Manen (2008).

**Ethnography** (also known as ethnographic research) is characterized as the prolonged, naturalistic study of a culture or group to understand shared meanings, norms, and other behaviors from the insiders' perspectives (i.e., an *emic* perspective). It often involves participant observation, fieldnotes, interviews, and artifacts to generate a rich description. For a brief introduction into ethnography, see Fetterman (2008).

**Action research**: The key feature of action research is its emphasis on partnering with community members to conduct research that will lead to action that produces good changes for the benefit of that community. Though IWO psychologists often partner with local organizations to do research, such partnerships often lack the depth with which archetypal action-research tends to involve those partners in the planning and execution of the research. For a brief introduction into action research methods, see Somekh (2008).

This is a good time to introduce the *emic* versus *etic* distinction. An emic (/ i m k/) perspective is an insider's first-hand view about some concept (often this is the research participant), whereas an etic (/ t k/) perspective is a researcher's view as an outsider about the same concept. The goal of many qualitative research methods is to combine an emic and etic perspective.

**Grounded theory research** culminates in producing a theory (i.e., the grounded theory), based on an iterative process of data-collection and data-analysis. It typically starts with analyzing qualitative data from an *emic* perspective, then collecting and analyzing more qualitative data until very little new emic perspectives are obtained (this point is called *data-saturation*). At that point, the researcher analyzes the data via an *etic* perspective, using their outsider concepts and terminology. Thematic patterns in the data are aggregated into broader themes. Finally, the researcher identifies relationships among the themes, thus culminating in a theory. That theory is called a *grounded* theory because it is based on (i.e., *grounded in*) the data that was rigorously analyzed. For a brief introduction into grounded theory methods, See Charmaz and Bryant (2008).

## 4.4 Mixed Methods

As I mentioned before, **mixed methods** research makes use of both qualitative and quantitative methods within either a single study or a program of studies (also known as a sequence of studies). As quantitatively-trained researchers become increasingly aware of the utility of qualitative methods, mixed-methods research has become quite popular.

Many IWO psychology scholars believe qualitative research is best suited for inquiries into topics for which we're not confident

A **program of research** is a logically organized sequence of research studies that build upon each other. This mustn't be confused with a *computer program* for doing research.

that we know what questions to ask in the first place (i.e., a *nascent* area of inquiry), whereas quantitative research is best suited for inquiries into highly specific topics that we believe we already have much knowledge about (i.e. a *mature* area of inquiry). For a good explanation of that conceptualization from IWO psychology scholars, see Edmondson and McManus (2007). Although that may be a useful rule-of-thumb, one can make very valid arguments about why a qualitative approach is useful regardless of how nascent or mature the area of inquiry is.

## 4.5 A note about the remainder of this course

For better or worse, this chapter is the most in-depth we will go into qualitative research methods in this course. For the interested reader, I've provided many supplemental references that go into great depth on many qualitative research methods. However, the current praxis of IWO psychology focuses heavily on quantitative methods, and employers of graduates with a degree in IWO psychology often seek employees with strong quantitative skills. Thus, for the sake of your immediate employability and acclimation into the world of IWO psychology, the remainder of this course will lean heavily on quantitative research methods (though I attempt to mention qualitative methods throughout). Fortunately, you have the option to learn about qualitative methods on your own with the reference material I've cited here and also listed in the Suggested Readings section below. One day, standard Research Methods courses for IWO psychologists may have much greater balance in their inclusion of qualitative and quantitative methods. Unfortunately, that day is not today.

## 4.6 Suggested Readings

Due to copyrights, I cannot give you full-text versions of entire books without prior permission from the copyright holder. However, I can recommend the following books for anyone interested in learning much more about qualitative methods:

- Creswell, J. W., & Poth, C. N. (2024). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (5th ed.). SAGE Publications, Inc.
    - Creswell and Poth present a great conceptual typology for qualitative methods. Just be sure to keep in mind there are other ways of categorizing those methods. There isn't one single "correct" way to categorize qualitative methods.
- Willig, C., & Rogers, W. S. (Eds.). (2017). *The SAGE Handbook of Qualitative Research in Psychology* (2nd ed.). SAGE Publications Ltd. https://doi.org/10.4135/9781526405555
    - This is a large handbook that you can use like an encyclopedia, to give you a quick but thorough summary of many of the most important topics in qualitative research methods.
- Given, L. M. (Ed.). (2008). *The SAGE Encyclopedia of Qualitative Research Methods* (Vols. 1–2). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909
    - This is an encyclopedia, which provides very brief but insightful introductions to concepts. For topics I'm less familiar with, I'll often reach for an encyclopedia before I reach for a handbook.

# 5 Literature-review Methods

## 5.1 What is a scholarly literature-review?

In a nutshell, doing a scholarly literature-review is like doing a sophisticated search of the internet. The word scholarly literature-review has two meanings, and both are important. One on the one hand, it can mean a *review of scholarly literature*. On the other hand, it can also mean a *scholarly review of literature*. In the former sense, *scholarly* describes the *types of literature* you're reviewing. In the latter sense, *scholarly* describes *your process* for reviewing literature. A piece of literature is generally regarded as *scholarly* if it is produced by and for scholars and it cites references that experts typically consider as valid evidence related to the author's claims. A process for reviewing literature is loosely regarded as scholarly if the process is rigorous and well-organized. In these lecture notes, I will describe where you can find scholarly works, and how you can keep your search process rigorous and well-organized.

## 5.2 Why is it important to do a scholarly literature-review?

In Section 2.3, I briefly mentioned a few types of scholarly works that are typically produced by IWO psychologists. One of those is a *literature-review study*, which often culminates in a *literature-review paper* (for examples of literature-review papers, see: Legood et al., 2023; and Newman et al., 2017). However, on a smaller scale, arguably all scholarly projects involve at least some amount of literature-review. Broadly speaking, a **literature-review** is any process or result of searching for and discovering prior scholarly works related to your topic of inquiry. At a minimum, most published scholarly works provide their audience some amount of relevant literature-review that helps the audience understand the current study. Thus, even if you aren't aiming to publish a literature-review *paper*, the process of conducting a literature-review is still important for your own education and for the education of your audience. This is true even if you want to conduct cutting-edge research on anything that is generally considered to be a revolutionary topic, because even such topics are contextualized and informed by the prior history related to such topics.

One of the main purposes for doing a literature-review is to avoid "reinventing the wheel". At the risk of sounding overly reverent toward our elders, allow me to provide you an illustration. Decades before you were born, many scholars spent their entire careers thinking about many of the same questions you're now wondering about. They wondered about the best ways to promote worker wellbeing, reduce stress, enhance performance, foster team cohesion, change company culture, etc., etc. Now imagine how much you've learned in the past 5 or 10 years of your life, and then imagine how much you'll learn after a full career of researching IWO psychology topics! Those scholars, who came and went before you were born, published their thoughts and insights that they cultivated across their entire careers. Sure, some of it may be outdated in some ways, but many (arguably *most*) of the IWO psychology research from prior decades is still highly relevant today. Although human *culture* may seem to change rapidly, human *nature* changes very slowly at the pace of genetic changes in our collective DNA. Thus, by conducting a literature-review, we can reap in a matter of *hours* the knowledge of thousands of scholars who dedicated their *careers* to studying the same questions we're interested in today. That is an extremely efficient way to learn a lot in a little bit of time.

Indeed, conducting a thorough literature-review has been and continues to be a feature of the praxiology of IWO psychologists. For better or worse, scholarly works produced by IWO psychologists have been criticized for overly lengthy literature-review sections that go into arguably too much detail about the past. Of course, there must be a balance. By knowing your target audience—what they likely know and don't know—you can have a better idea of how much detail to include in the literature-review sections of your scholarly works.

> Don't be upset if others have already started researching your idea.
>
> Often, beginner researchers will lament, "I can't research this topic because it's already been done". However, such thinking is sorely misguided! Even though other researchers examined phenomena related to your topic of inquiry, there is often so much more to be learned about that topic! Throughout the remainder of this course, I'm sure your eyes will be opened to all of the different types of questions that can be asked about any topic.

### 5.2.1 Types of scholarly works you can search for

In Section 2.3, I briefly mentioned a few types of scholarly works that are typically produced by IWO psychologists. Specifically, I mentioned: empirical studies, theory papers, lab studies, field studies, literature-review studies, and meta-analysis studies. Each of those types of studies typically culminates in one or more discrete *products* (also known as an *output*), which could be a written article, or a video, podcast, etc. In this chapter, I want to focus on *written* scholarly works, which are products that consist primarily of *written language*, because that is by far the most prevalent format available in scholarly databases (in subsequent lectures we will discuss other formats that heavily involve audio, video, and in-person events).

The very generic word that scholars use to refer to written works is *paper* (e.g. *"I was reading a paper about…"*, *"I'm working on a paper about…"*). Using more formal language, all written works start as **manuscripts**, which are *unpublished* written works. A manuscript might be unfinished, or it might be completely written. As long as it's unpublished, we can call it a manuscript. When a manuscript becomes published, we call it a **publication**, which is any published work (in this case, a published *written* work).

Among scholars, almost any publication can be categorized as either a book or an article.

I'm sure you're already familiar with books, but scholars distinguish between two important subcategories: *authored books* versus *edited books*. An **authored book** is written entirely or almost entirely by a single author or a small team of authors, such that the entire team gets credit for writing the entire book. In an **edited book**, one or more persons known as the *editors* of that book will create a plan for what they want the book to be about, and then they will invite several other authors—who are considered experts in the various topics related to the book—to write one or more specific chapters for the book. Typically, the editors are experienced scholars who are knowledgeable about the topics in the book, and they provide editorial feedback to the chapter authors' chapter manuscripts for the book before a final version is reached. Each chapter may be written by one or more authors, and each author typically only contributes one chapter (though multiple chapter contributions isn't unheard of).

In the context of published written works, you may have some sense of what we mean by *articles* (e.g., *"I was reading an article about..."*). Among scholars, the word *article* almost always means a **peer-reviewed journal article**, which is an article that was reviewed by other scholars—who are considered experts in the topics of the article—before it is published in a scholarly journal. A **scholarly journal** is essentially a magazine for scholarly articles (and recall earlier in this chapter I mentioned that any written work is generally regarded as *scholarly* if it is produced by and for scholars and it cites references that experts typically consider as valid evidence related to the author's claims). Before the internet, journals were published in print, and that's why each new "print" from a journal is referenced via its *volume* and *issue* number (just like magazines). The **volume** typically corresponds to the number of years the journal has existed up to that point, and the **issue** is the number of times a new instance of the journal has been published that year. Modern journals are increasingly getting rid of *issue* numbers, because journals are almost exclusively accessed online without the need for any physical prints.

In modern edited books, the book will be assigned a DOI, and each individual chapter will also be assigned a separate DOI. That way, the authors of each chapter can keep track of how many people are reading/citing/downloading their specific chapters. See

In Chapter 12, I provide more detail about the peer-review process.

In this modern era in which scholarly works are almost always produced digitally, you can think of any publication as a type of **digital object**. To help keep track of digital objects, technicians invented the **Digital Object Identifier**, or **DOI** (pronounced: / di .o a /, like *dee-oh-eye*) which is a globally unique alphanumeric identifier for a specified digital object (e.g., an article, dataset, picture, audio file, book, book chapter, preprint). They are convenient because they are *persistent*, meaning the DOI for an object won't change, even if the website URL for the object is updated. That way, searching for and finding a digital object via its DOI is super simple.

### 5.2.2 Where to find scholarly works

An experienced researcher can find high-quality scholarly works from a simple internet search (e.g., Google search), because such researchers are skilled at figuring out whether any research article is rigorous and transparent in its methods. However, less-experienced researchers might not know how to differentiate a high-quality scholarly peer-reviewed journal article, versus an article that *looks like* a peer-reviewed journal article but is actually published in an outlet that enforces very low-quality standards for the scientific validity of the contents of its articles. To help you avoid getting misguided by poor-quality research articles, you can play it safe by sticking to publication outlets that have a long reputation of enforcing high standards for their content.

#### 5.2.2.1 A word of caution about Google Scholar

In a moment, I will give you a brief list of good sources to get you started. But first, let me say something about **Google Scholar**. Google Scholar has done a lot of good, especially by simplifying access to many scholarly works that would otherwise be difficult to find. However, less-experienced researchers can be easily duped on Google Scholar, because it contains anything and everything that is even remotely classified as scholarly. In other words, it's very easy to find low-quality research on Google Scholar. Therefore, unless you already have a list of high-quality journals you want to search for on Google Scholar, I would advise caution before you assume any papers on Google Scholar are high-quality. With time, you will learn how to identify a good paper (e.g., by evaluating the contents of the paper, the history of its authors, the history of the editors of the journal that published the paper, etc.), and then you can search Google Scholar with confidence.

#### 5.2.2.2 APA Databases

If I had to recommend one database that is essentially guaranteed to only include high-quality publication outlets, it would be the **APA PsycArticles®** database—for which most universities pay a subscription on behalf of their students. The APA PsycArticles® database contains the scholarly peer-reviewed journals that are published by the APA Journals™ publisher, as well as other high-quality journals that are affiliated with the APA.

You may have also heard of the **APA PsycInfo®** database, which would be my second recommended database for new researchers—maybe even tied in first place with APA PsycArticles®. The APA PsycInfo® database contains many more journals than the APA PsycArticles® database (most of which aren't affiliated with APA at all), but APA PsycInfo® doesn't necessarily include the full text of the articles for you to read (often, it only includes the short summary paragraph known as the *abstract*), unless you have a separate subscription to those third-party providers for those articles. In contrast, APA PsycArticles® provides you with the full text for all articles included in the database.

> **Databases versus hosts and/or platforms...**
>
> Now might be a good time to explain to you the difference between a database, versus a host and/or platform. A database is a collection of data, and databases are typically updated to include the newest data into it. In contrast, a host and/or platform is merely a service that gives you access to a database. APA built a website interface—which it

calls APA PsycNet®—as a platform where you can access all of APA's various databases, including APA PsycArticles® and APA PsycInfo®. Similarly, you may have also heard of something known as EBSCOhost, which is another online platform for accessing many databases, including APA's databases, depending on which databases you are subscribed to via EBSCOhost. Therefore, when describing where we will search for scholarly sources, we mustn't say "I will search EBSCOhost", because the databases you are subscribed to in EBSCOhost (perhaps via your university) may be very different than the ones I'm subscribed to, even though we're both using the EBSCOhost platform. Instead, we must name the actual databases we are searching.

#### 5.2.2.3 Web of Science™

Another good database for you to be aware of is the **Web of Science Core Collection**™, which is typically accessed via the **Web of Science**™ platform. Many researchers often refer to the Web of Science Core Collection™ as simply the *Web of Science database*—often abbreviated as *WoS*—but that is technically incorrect and confusing because the Web of Science™ platform can also be used to access other databases besides the Web of Science Core Collection™. Either way, the Web of Science Core Collection™ is a large database, and it does a much better job of filtering out low-quality outlets than Google Scholar does.

#### 5.2.2.4 Library websites

Another resource worth keeping in mind are **university library websites**. If you are a student or employee of a university, your tuition dollars typically automatically help pay for dozens of database subscriptions which you can access via your library's website. Depending on where you live, public libraries might also provide access at no charge to you. Most universities subscribe to APA PsycArticles® and APA PsycInfo®, and you would typically access them via a platform like EBSCOhost.

#### 5.2.2.5 Journals that IWO psychologists frequently publish in

In addition to databases which contain written works across hundreds of journals, you can also browse the websites of specific journals. As you become more experienced, you will develop a sense of which journals typically produce high-quality content. There are hundreds of journals that IWO psychologists typically publish in.

Below, I've included a small list of journals that may be particularly good for people who are new to IWO psychology, because I think that will give you a concrete launching point. However, by presenting you a small list of examples here, I'm risking biasing you into thinking these are somehow better than any journals I didn't include in the list. There are dozens of high-quality journals I omitted from this list, simply because I didn't want the list to be overwhelming for you. If you're ever wondering what other journals might be good to look into, a quick internet search should be helpful, and you can also email me such questions.

- Journal of Applied Psychology (JAP)
- Personnel Psychology (P. Psych)
- Annual Review of Organizational Psychology and Organizational Behavior
- Journal of Occupational and Organizational Psychology (JOOP)
- Journal of Business and Psychology (JBP)
- Journal of Occupational Health Psychology (JOHP)
- Academy of Management Annals
- Academy of Management Review (AMR)
- Harvard Business Review (HBR)
- Human Resource Management Review
- Academy of Management Journal (AMJ)
- American Psychologist
- Psychological Bulletin

Personally, I recommend keeping a spreadsheet of your favorite journals (including their website URLS, and other information you find useful), and then making a routine for reading through the abstracts of those journals so that you can stay at least somewhat up-to-date on the latest research in your areas of interest.

## 5.3 Now that we know *where* to search, *what* should we type into the search box?

The text that you write into any search box (e.g., on a website or computer database) is known as your **query**. Typically, the search boxes for scholarly databases are built to handle **Boolean queries**, which are queries that use *Boolean logic*. Boolean queries contain special symbols that enable you to be very specific with your queries. Table 5.1 contains the most important symbols you should know.

*Boolean logic* is named after George Boole (1815–1864) who introduced its foundational concepts.

Table 5.1: Boolean query syntax

| Syntax element | Example Boolean query | What it does in a query |
|---|---|---|
| `AND` | `work AND "psychological safety"` | We are requiring both expressions that are on either side of the `AND` operator. In this example, we're saying we don't want any results that *only* contain `work` nor *only* `psychological safety`; we only want results that contain *both* expressions. |
| `OR` | `occupation OR work` | We are requiring *at least one* of the expressions on either side of the `OR` operator. In this example, we're saying we want any results that contain either `occupation`, or `work`, or both `occupation` and `work` at the same time. |

| Syntax element | Example Boolean query | What it does in a query |
| --- | --- | --- |
| `NOT` (sometimes also the minus sign: –) | `work NOT physics NOT joule` | Anything after the `NOT` operator will be excluded from the results. In this case, we're searching for anything that contains `work`, but we don't want anything that's about physics, which has a completely different definition of *work*. |
| `*` (it's an asterisk, but in queries it's typically known as the **wildcard** symbol) | `work*` | We will get results containing any word beginning with `work`, such as `worker`, `working`, `works`, `workable`, etc. Some search engines allow wildcards inside a double quote, but some don't. You can check the specifics of the search engine you're using. |

| Syntax element | Example Boolean query | What it does in a query |
|---|---|---|
| `" "` (i.e., double quotes) | `occupational NOT "occupational therapy"` | Treat the contents inside the double quotes as a single intact phrase (also known as a **string**). In this example, we could end up with results that contain `therapy`, but not any results that contain `occupational therapy` as a sequence of two words. In other words, we could get results containing text such as `occupation-related therapy`, and `therapy for occupations`, but not anything that specifically contains `occupational therapy`. |
| `( )` (i.e., parentheses) | `(work OR organization*) AND (climate OR culture)` | Parentheses will group expressions in the same way you would for regular algebra (from grade school). In this example, we want results that are about `work climate` or `work culture`, which are also known as `organizational climate`or `organizational culture`. |

The other crucial thing to know is the types of database **fields** you can search within a search engine or database. Perhaps the best way to explain the concept of database fields is via an example. If you think of a spreadsheet, a field could be a column that you've labeled for a specific type of data. For example, if I want to create a database which will be a simple spreadsheet to list everyone in my company, such that each row of the spreadsheet will be a new person, then I might create these fields (or columns) in my spreadsheet: `First name`, `Last name`, `Title`, `Department`, `Email`. To make this example more concrete, see Table 5.2.

Table 5.2: Simple example database

| First name | Last name | Title | Department | Email |
|---|---|---|---|---|
| Jane | Smith | Project Manager | Human Resources | janesmith@example.com |
| Alex | Xela | Engineer | Computer Science | alexxela@example.com |
| Ricky | Bobby | Marketer | Marketing | rickkybobby@example.com |

Unfortunately, scholarly databases don't all use the same names for the same types of fields. For example, for a field that contains the title of a journal, one database might call that field `Journal` whereas another database might call it `Outlet`. However, most databases typically make it very easy for you to see a complete list of all the fields you can search, as well as how to format your queries to search those fields.

Thus, equipped with the knowledge of Boolean query syntax and the database-specific formats for how to search in fields, you can combine those two to make highly-specific queries that specify what you want to search in various fields. For example, you could specify you want to search for `psychological* AND safe*` in the `Title` field, and also `Review` in the `Journal` field. That way, you'd get results from any journal with the word `Review` in the journal's title (such journals typically publish literature reviews authored by experienced scholars), and your results would contain things like `psychological safety` or `psychologically safe` in the `Title` of the article.

## 5.4 Specifying your inclusion and/or exclusion criteria

Now that we've discussed the types of scholarly works you could search and how to format your queries, I want to now focus on the important step of creating inclusion and/or exclusion criteria. A *criterion* is *"a principle or standard by which something may be judged or decided"* (Oxford Dictionary of English, 2025). Your **inclusion criteria** are your criteria that define the types of scholarly works you want to *include* in your literature-review. Your **exclusion criteria** are your criteria that define the types of scholarly works you want to *exclude* from your literature-review.

The purposes of inclusion/exclusion criteria are primarily to help you organize your literature-review process. Often, a peer-reviewer will ask you how why you decided to review only the scholarly works you ended up including in your literature-review write-up, or why you *didn't* include other specific scholarly works that the peer-reviewer believes would be good to include. In that situation, having a clearly specified and logical list of inclusion/exclusion criteria will help validate your argument in response to that peer-reviewer. Indeed, if your inclusion/exclusion criteria are nonexistent or vague or illogical, the peer-reviewer might be correct in stating that your literature-review is severely incomplete due to major omissions of various types of scholarly works that you really should've included in the first place.

Importantly, your inclusion/exclusion criteria should be compatible with your goals for your research (e.g., your hypotheses). For example, if you want to figure out the average job-satisfaction of workers in the United States since the Great Recession of 2007, then your inclusion/exclusion criteria might look like this:

- I only want information that was published in high-quality outlets, such as high-quality journals, or official U.S. government databases (because I don't want any results that came from a sloppy measurement of job-satisfaction, nor incorrect methods for calculating the statistics).
  - I don't want any conference posters; due to the high variability in the quality of peer-reviews for conferences.
- I want the information to be based on data that was collected after 2007 (if the date of

data-collection is unknown, I'll exclude those sources).
- I want information about working adults.

The above list is just an example. You could create a much more specific set of criteria. You might even write a list of specific outlets and justify the list on the basis of their quality and your limited time. The appropriate level of detail will depend on your preference and the preference of whichever outlets you're aiming to publish in.

## 5.5 Specifying your list of keywords

When you're searching in a typical search engine or database, you likely aren't writing complete sentences. Instead, you're probably only writing a fragmented series of **keywords**, like: `BBQ restaurants` or `eyeliner tutorial how to`.

Especially in scholarly literature reviews, the quality of your keywords can make a *huge* difference in the types of information you end up finding. This is especially true when scholars use different words to refer to one concept. For example, there is a concept in statistics known as *Simpson's paradox*, but it is also known as *Simpson's reversal*, the *Yule–Simpson effect*, and the *amalgamation paradox*. Imagine if you only searched for scholarly works that included the keyword `Simpson's paradox`, and you ended up not finding several newer publications that use the more recent keywords such as `Yule-Simpson effect` or `amalgamation paradox`.

Thus, it is extremely important to keep a list of all of the keywords that are relevant to your query.

Fortunately, recent developments in large-language artificial-intelligence tools such as ChatGPT are immensely helpful for coming up with a great list of keywords. For example, you could just ask ChatGPT something like, *"Please give me a comprehensive list of keywords I should include in my queries for scholarly databases, to help me find works about...[insert your main ideas here]"* (in Section 5.7, I describe other ways for using large-language A.I. tools to help with literature-reviews).

In the past—before tools like ChatGPT—scholars relied heavily on thesauruses such as the **APA Thesaurus of Psychological Index Terms®** (which requires a paid subscription) and the **APA Dictionary of Psychology** (available for free online). Those resources are still very useful today, but tools like ChatGPT have made the keyword-creation process much easier. In fact, you could try asking ChatGPT to use the APA Thesaurus and APA Dictionary when it creates a keyword list.

## 5.6 Forward citation search and Backward citation search

Aside from entering queries into a searchbox, another excellent set of techniques to discover scholarly-works relevant to your query is the *forward citation search* and *backward citation search*.

When you are conducting a **backward citation search**, you are essentially asking “What came *before* this publication?”. In other words, you’re moving backward across time. Specifically, you are looking at the references that are listed in the *References* section of any good scholarly publication. For example, if I found a great journal article published in October 2020, I can look in the *References* section of that article and look up each of the sources listed in there. Logically, all of the sources in the *References* section were published before October 2020. Thus, we are searching backward across time.

When you are conducting a **forward citation search**, you are essentially asking “What came *after* this publication?”. In other words, you’re moving forward across time. Specifically, you are looking for any subsequent publications that are citing a specified prior publication. For example, if I found a great journal article published in January 1991, I can search databases to find any subsequent publications that reference the January 1991 article. Logically, any publication that references the 1991 article must have been published after January 1991. Thus, we are searching forward across time.

To do a forward citation search, you’ll need to use a search platform that supports that feature, such as Google Scholar (free), or the Scite® platform (requires a paid subscription), or the APA PsycNet® platform (which requires a paid subscription). It is particularly easy in Google Scholar, because all you have to do is click on the `Cited by` button beneath any search result. Keep in mind, the publications that will be listed in one platform may be slightly different than those of another platform, because of their different underlying databases.

Another tool that can help with both backward and forward citation searches is known as the Local Citation Network (free). To use it, you'll just need to enter the DOI or some other identifier for your focal publication.

## 5.7 Using A.I. tools for literature-reviews

Around the year 2023, A.I. tools like ChatGPT exploded in popularity. Since then, those tools have evolved at such a rapid pace that it would be a bit of a waste of time for me to go into lengthy detail about the various features of such tools—because they are very likely to become much more advanced tomorrow! Thus, it seems email, phone, and in-person conversations are the better communication mediums for discussing current A.I. tools, rather than trying to convey those tools in these lecture notes which only get updated once per year.

Still, I here is a general list of use-cases (i.e., contexts or purposes for which something is used) for using tools like ChatGPT for literature-reviews:

- Summarizing a topic, ideally with cited references included for you.
- Getting help to re-specify your research topic more clearly.
- Creating an appropriate set of inclusion/exclusion criteria for the types of scholarly works you want to find.
- Creating a comprehensive list of keywords (including synonyms and near-synonyms) for your queries.
- Discovering publications that fit your inclusion/exclusion criteria and your keywords.
- Summarizing any publications that you're struggling to understand or don't have the time to read completely for now.

**Elicit** (https://elicit.com) is another popular A.I. tools specifically made for reviewing scholarly literature. There are likely many other tools, so look around!

## 5.8 Use reference-management software to organize your collection of scholarly-works

As you are conducting your literature-review, you will start to build a collection of good publications that are relevant to your topic. Many people will immediately download the PDF of each of those publications and sick them in a folder on their computer. That is certainly fine, but there is a much more organized and user-friendly method for managing your collection. I'm talking about reference-management software.

**Reference-management software** (also known as *citation-management software*, or *citation managers*) help you collect, store, organize, and cite digital objects that you may want to reference (e.g., scholarly works such as articles, books, chapters, conference posters, etc.). Typically, reference-management software also enables you to effortless and automatically create a perfectly formatted (e.g., APA style) *Reference* list from all of the sources you've included in your manuscript.

Reference-management softwares rely on meta-data for the digital objects in your collection. Of course, digital objects convey information encoded in that object—such as the words and other information contained in a journal article—but the digital object is also associated with **meta-data**, which is any data that describes the digital object itself. Examples of meta-data include:

- The names of the authors or other creators.
- The date the object was published or created (e.g., publication date, or conference date).
- The title of the object (e.g., book title, article title).
- The object's DOI.
- The name of the outlet in which the object is published (e.g., journal title, or website name).

**Zotero®** is one of the most popular reference-management softwares, probably because of its robust features and it is free! Zotero®, like many reference-management softwares, can automatically import the meta-data for any digital object, based on the object's DOI or URL. In fact, Zotero® has a web-browser extension that will let you import digital objects into your collection by just clicking a button when you're on the website of that digital object.

At the beginning of this course I included a tutorial on how to set up and use Zotero®. I also enjoy talking about Zotero® in general, so please feel free to email me if you have any questions.

# 6 Building a Theory, Model, and Hypothesis

## 6.1 Introduction

This course focuses primarily on the **hypothetico-deductive** approach for inquiry—which I defined in Section 2.2.2—because that is the predominant approach in the current praxis of IWO Psychology research. Other approaches are very valid, especially the qualitative approaches described in Chapter 4. However, for the sake of your immediate employability and acclimation to the current world of IWO psychology, this course will focus more on the hypothetico-deductive approach for quantitative inquiries.

In this chapter, our main goal is to define several of the most important terms that will help you create theories or talk about theories, and to create good hypotheses to test a theory. Along the way, we must define several other terms that describe the processes involved in creating a theory and/or creating a hypothesis.

## 6.2 What is a theory?

A **theory** is an idea or statement about the relationship(s) among two or more concepts (compare with: Jaccard & Jacoby, 2020). The idea and/or statement may or may not be accurate or reflect "reality", but it is still a theory. Thus, Figure 6.1 is a diagram of the simplest structure for a theory.

In the language of mathematics, the diagram in Figure 6.1 is known as a **graph**, which comes from an area of mathematics known as **graph theory**. Laypersons typically use the word

Some scholars (e.g., Jaccard & Jacoby, 2020) prefer to distinguish between a *theory* versus a *conceptual system*, such that a conceptual system is a *mental representation* about the relationship(s) among two or more concepts, whereas a *theory* is an expressed *statement* (usually written or spoken) that represents a conceptual system. Such distinction isn't necessary in this course, but I felt it was worth mentioning.

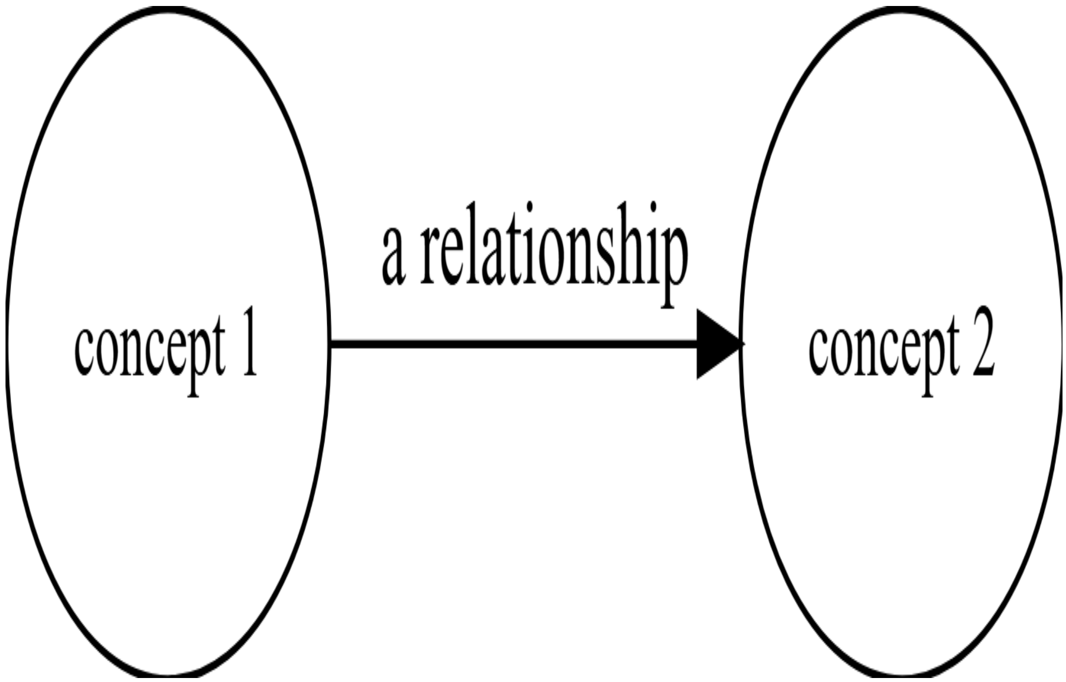



Figure 6.1: The simplest structure for a theory.

*graph* to refer to any data-based diagram (e.g., a bar chart). However, in this book, the word *graph* will always refer to the meaning from graph theory (i.e., the *graph-theoretic sense*).

The concepts and relationships in a theory can be anything you specify, and you can draw multiple valid graphs to represent any given theory.

Importantly, in any graph that represents a theory, each concept and relationship must be labeled, otherwise the correct interpretation of the graph is ambiguous. For example, consider the graph in Figure 6.2, in which the relationship betweens `productivity` and `income` is unlabeled.

Also, a *relationship* in one theory may be a *concept* in another theory, and vice versa. It's all based on the perspective of whoever is writing the theory or drawing a graph for it.

The relationship between `productivity` and `income` could be anything: maybe productivity `increases` income, or productivity `is positively correlated with` income, or maybe productivity `is uncorrelated with` income, or perhaps productivity `is negatively correlated with` income. Obviously, the latter three theories seem to contradict each other, so we

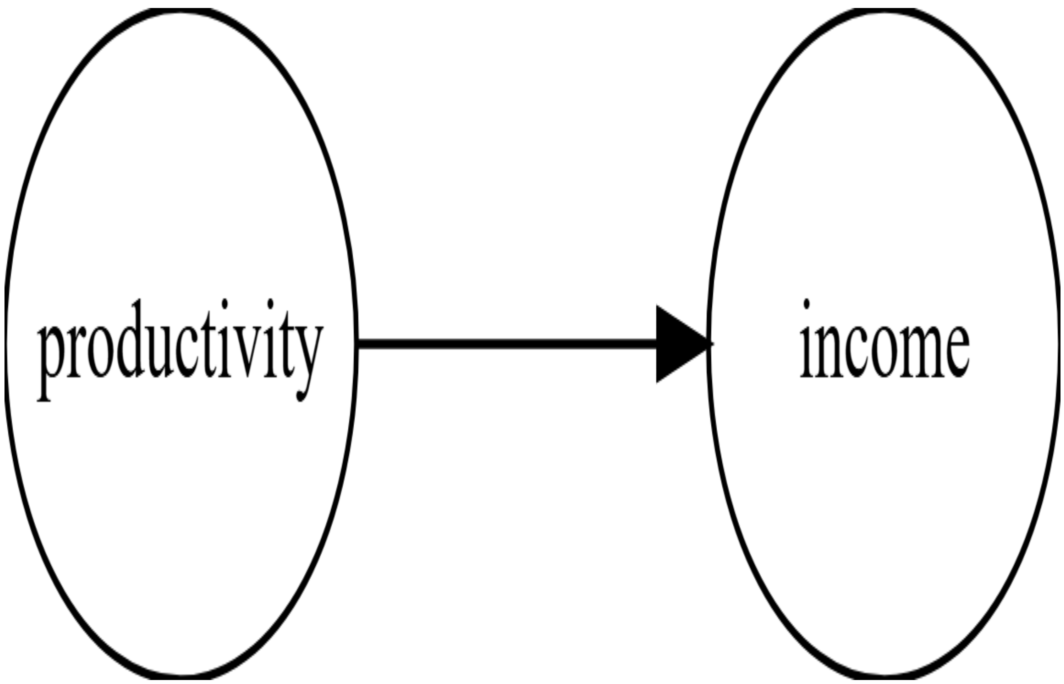



Figure 6.2: This graph has an unspecified relationship between `productivity` and `income`.

would need to do some empirical tests to figure out which one is true. Similarly, there are many theories in science, including IWO psychology, and it's up to us to test them to figure out whether they are accurate reflections of what we perceive as reality.

Another word worth discussing now is **model**. Some scholars (e.g., Jaccard & Jacoby, 2020) consider *models* and *theories* as synonyms for each other. Fundamentally, I agree with that definition, but many (most?) IWO psychologists (and/or behavioral scientists in general) typically use the word *model* to refer to either a part of a theory or a statistically derived mathematical representation of a theory. In quantitative research methods, you will almost always be creating a *statistical* model, such as Equation 6.1, which is a math-symbolic expression of the theory that productivity is correlated positively with income:

$$\text{income} = \beta_0 + \beta_1 \text{productivity} + \varepsilon, \text{ and } \beta_1 > 0 \tag{6.1}$$

Thus, the same theory and/or model can be expressed via words, or via graphs, or via math-symbolic expressions. Equation 6.1 is a math-symbolic expression of this verbal expression: *"Productivity is correlated positively with income"*. Likewise, Figure 6.3 is an equivalent diagrammatic expression from graph theory.

Again, a theory is an idea or statement about the relationship(s) among two or more concepts. Next, let's define what we mean by *concept*.

## 6.3 From concepts to cónstructs and variables

The word **concept** is a generic word to refer to anything that is conceived in the mind. Concepts can refer to things that have tangible manifestations (e.g., the concept of a worker's height) or they can refer to things that are not so clearly tangible, if at all (e.g., the concept of being satisfied with one's career). Also, concepts can be highly specific with rich detail (e.g., the concept of my uncle Steve's productivity in December), or they can be more generic with less detail (e.g., the concept of any worker's productivity in general). When we strip away the rich detail

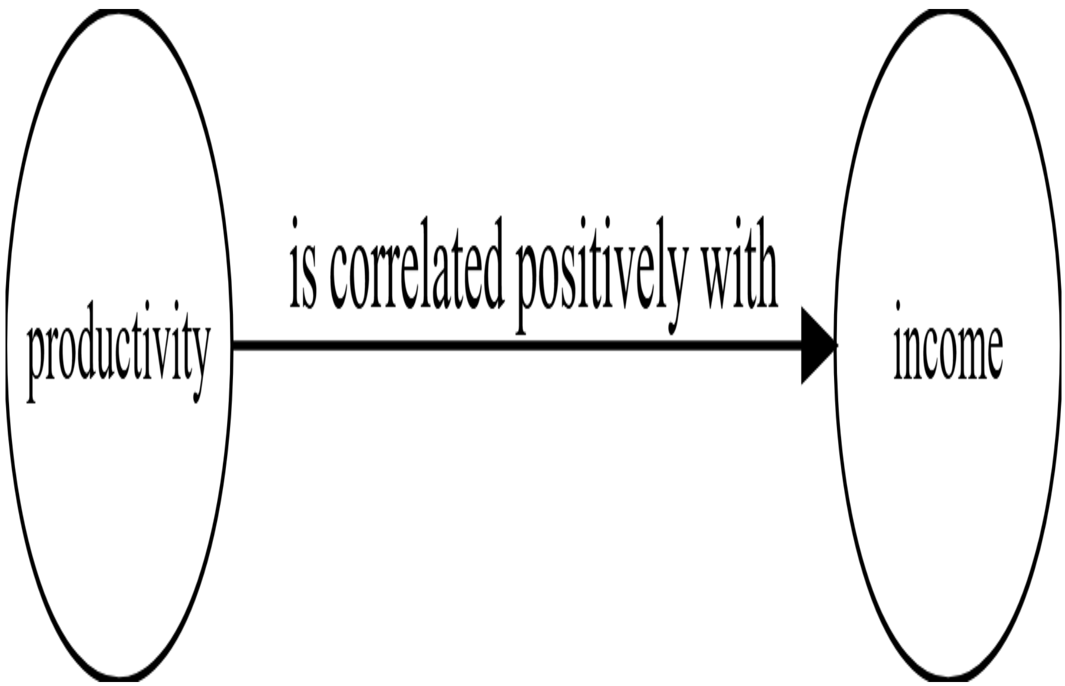



Figure 6.3: A diagrammatic expression from graph theory, equivalent to the verbal expression: *"Productivity is correlated positively with income"*.

of a concept to make it more generic, we are moving toward *greater abstraction*. In other words, a more generic concept is more abstract than a highly-specific concept.

Recall from Section 2.2.2.1 I described that scientists are generally interested in discovering *universal statements* that are true. In other words, IWO psychologists typically aren't interested in knowing whether my uncle Steve's productivity in December is correlated positively with his income. Rather, we want to know whether the productivity of *all* workers at Company ABC is correlated positively with their income, or perhaps whether that's true for *all* workers in the USA, or perhaps the entire world! Thus, IWO psychologists are *typically* interested in concepts that are more abstract, rather than less abstract.

Abstract concepts are thus so special that scholars typically call them *cónstructs* (more typically spelled: *constructs*), rather than calling them *concepts*. Thus, a **cónstruct** is merely a generic term to refer to any concept that is regarded as being more abstract than not. There is no special rule to determine the level of abstraction that is appropriate before you can call a concept a *cónstruct*. For example, the role of *engineer* may be considered a concept, and so is the role of *marketer*. Thus, we could say the concept of *roles* in general is a *cónstruct*. However, we could just as easily say *the engineering role of Steve* is a concept, and so is *the engineering role of Alex*, and that *the role of any engineer in general* is a cónstruct. Thus, the distinction between a concept versus a cónstruct is in the eye of the beholder. In general, IWO psychologists typically refer to most concepts they study as *cónstructs*, perhaps out of habit.

Cónstruct vs. Construct...

The word *construct* can be a verb (i.e., *to build*)—in which case it is pronounced /kən strək(t)/, like *con-STRUCT*—and it can be a noun (i.e., *an idea or theory*)—in which case it is pronounced / kän strək(t)/, like *KAHN-struct*. There are few instances where the meaning isn't disambiguated via context clues. However, I always force the disambiguation via the acute accent diacritic over the letter *o*: *cónstruct*. As far as I know, there are very

> few people who write it like that, but I think it helps make things at least a little easier for new students to understand sentences like: *"we're going to construct a measure of a construct"* (i.e., *"we're going to construct a measure of a cónstruct"*). For me, the disambiguation is worth the extra keyboard strokes. Don't be alarmed if you never see anyone write it like that after this course (and I won't be offended if you prefer to just write *construct*!).

A special type of cónstruct that is commonly studied in science is called a *variable*. In the general vocabulary of science and data-analysis, a **variable** (pronounced: / verēəb(ə)l/, like *VERY-uhbl*) is any cónstruct that represents multiple possibilities. For example, the temperature in any room of a company's building is a variable, because there are multiple theoretical possibilities that the temperature could be (i.e., it could be cold or warm today, and some rooms might be different than others). Once we measure a room and we know its temperature is 70 , then we say 70  is that room's *datum* or *value* or *score* on the *temperature* variable. There are countless (arguably infinite) variables we can think of about people: variables about their appearance (e.g., the presence of tattoos, the person's height), variables about their behaviors (e.g., how fast they talk, how nervous they typically feel), variables about the structure of their social networks (e.g., the types of coworkers they have), and the list goes on and on. Indeed, the cónstructs that are typically studied in IWO Psychology are almost always variables. Keep in mind: variables don't have to represent numbers. A person's hair color is a variable (i.e., some people have black hair, others have brown, blonde, red, etc.), but it isn't necessarily represented by a number. However, because many IWO psychologists focus on quantitative methods, the variables they study are either inherently numeric or they are converted into numbers (e.g., black hair = 1, brown hair = 2, …).

Now that we've defined concepts, cónstructs, and variables, let's now explore a special type of variable known as a *dimension*.

## 6.4 Dimensions and multidimensionality

You've likely heard that we exist in three-dimensional space, or that a drawing on a flat sheet of paper is approximately two-dimensional. We can use our intuition from those examples to arrive at a general definition for the word *dimension*. A **dimension** is any variable whose data can be put into a logically ordered sequence of increasing (or decreasing) *intensity*, *amount*, *magnitude*, or *level* of the variable's essence. Common examples of dimensions are `length`, `weight`, `color saturation`, `sweetness (as a taste)`, and `size`. Each of those is a dimension, because they are variables whose data can be put into a logically ordered sequence of increasing amount. For example, I could describe three foods along the `sweetness` dimension as `not sweet`, `a little sweet`, and `very sweet`. As you can see, as we move from left to right, those data are strictly *increasing* in their level of `sweetness` in that order—and they're strictly *decreasing* in `sweetness` if we move from right to left. In contrast, a variable such as `occupational industry sector` with data such as `agricultural jobs`, `healthcare jobs`, and `education jobs`, is not a dimension, because those data do not form a linear pattern of any increasing (or decreasing) essence.

Thus, it's convenient to think of a dimension as an imaginary *directional* line, such that each subsequent datum has *more* of the variable's essence as you move in one direction of the line, and they have *less* of the variable's essence as you move in the opposite direction of the line.

Whereas dimensions like `length`, `weight`, `color saturation`, and `size` can be measured using tools such as rulers, weight-scales, and optical sensors, `sweetness` is a *subjective perception*. Instead of measuring `sweetness`, we could measure `amount of sugar` with a chemical device, but what if we are actually interested in the *perceived sweetness* of a food, rather than its amount of sugar? For example, diet soda drinks have no sugar, yet we perceive them as sweet and we may legitimately want to compare the perceived sweetness of several non-diet and diet drinks, regardless of whether or not they contain any sugar. To measure the sweetness perceived by a person, we could either ask them to tell us how much sweetness they perceive, or we

*Datum* is the singular word for *data*. For example, if you are doing a survey of everyone's age at the university, and only one person responds to your survey, you have a *datum*. If multiple people respond to your survey, you have *data*.

could measure the parts of their brain that light up when they taste something sweet.

Similarly, in IWO Psychology we're often interested in variables such as `happiness`, `anxiety`, or `extroversion`—none of which can be directly measured with a ruler nor weight-scale. That is why psychologists invented techniques to help measure subjectively perceived things and other abstract psychological cónstructs. We call those techniques *psychometric techniques*, and the topic of measuring such psychological cónstructs is known as *psychometrics*. In Chapter 7, I provide a gentle introduction to the basic ideas of psychometrics, but you will also receive an entire course on psychometrics later in this degree program.

In a nutshell, when we interview or observe people because we want to figure out how happy, anxious, or extroverted they are, we can use psychometrics to convert that information into numbers and that's how we derive numeric variables for those cónstructs. Thus, for example, we can imagine a person's level of `extroversion` can be represented by a number on an imaginary number line, starting from zero and increasing to greater levels of `extroversion`. Thus, we can think of `extroversion`, `happiness`, and `anxiety` as three separate dimensions. Of course, there are many more (arguably infinite) types of dimensions we could describe any person with. For example, each of these is a dimension:

- The number of years the person has been working at their job.
- The number of coworkers the person considers as their close friends.
- The level of stress the person subjectively experiences on a typical day at work (e.g., on a scale from 1 to 10).
- The level of passion the person feels toward their work (e.g., on a scale from 1 to 10).
- The ease with which the person responds aggressively toward hostile comments from others (e.g., on a scale from 1 to 10).

As you can see, any person can be described via a never-ending list of dimensions.

### 6.4.1 Multidimensionality

When a variable is defined as a single dimension, we say it is *unidimensional*.

Sometimes, we may want to define a cónstruct as a combination of multiple concepts of multiple cónstructs (i.e., multiple ingredients). When two or more of those ingredients are distinct dimensions, then we say the overall cónstruct is **multidimensional**.

For example, the owners of Company ABC might define their worker's job performance as a combination of the worker's speed and the worker's accuracy. Speed is a dimension, and accuracy is a separate dimension. A worker could be fast and accurate, or fast and inaccurate, or slow and inaccurate, or slow and accurate, or anywhere else in between! Thus, that particular definition of job performance is a multidimensional cónstruct.

In IWO Psychology there are many popular multidimensional cónstructs, such as:

- personality (typically defined as a combination of five major dimensions: openness to experience, conscientiousness, extroversion, agreeableness, emotional stability);
- job-satisfaction (typically a combination of one's level of satisfaction with multiple distinct aspects of the job, such as one's pay, one's supervisors, one's coworkers, etc.);
- burnout (a combination of one's level of emotional exhaustion, one's amount of cynicism, and one's level of reduced professional efficacy);
- passion for one's occupation (typically a combination of how much one likes the occupation, how much one is motivated to engage in the occupation, and how much one self-identifies with the occupation);
- intelligence (typically defined as a combination of several dimensions of cognitive abilities such as verbal reasoning ability, and quantitative reasoning ability).

## 6.5 Types of relationships in a theory

At the beginning of this chapter, I said a *theory* is an idea or statement about the relationship(s) among two or more concepts. So far in this chapter, I've discussed the definitions of many types of concepts, including cónstructs, variables, and dimensions. Now let's move on to exploring some important types of the other main ingredient in theories: the *relationships*!

Fundamentally, when you write a theory, the relationships you write can be whatever you believe is true. Your theory might or might not reflect reality, but that's a separate issue. Just like we can define any concept in broad or specific detail, we can also define any relationship in broad or specific detail.

Giving examples is perhaps the best way to explain what we mean by a *relationship* between concepts. Let's consider the examples in Table 6.1, and notice it is split into Table 6.1a and Table 6.1b. We'll use those examples to explore some important types of relationships:

- deterministic versus probabilistic relationships;
- causal versus correlational relationships;
- moderated relationships;
- mediated relationships.

### 6.5.1 Deterministic versus probabilistic relationships

As you can see, Table 6.1a shows only relationships with *deterministic* phrasing, whereas Table 6.1b shows only relationships with *probabilistic* phrasing (also known as *non-deterministic*). If something is **deterministic**, it means it is guaranteed. If I throw a regular ball up into the air, it is guaranteed to come down. However, almost all relationships among concepts in psychology (and all behavioral sciences) are **probabilistic**—in other words, we can't typically guarantee that we know what will happen, but we can assign some level of (un)certainty to our statement (i.e., we can assign a *probability* to our statement). Thus, in reading Table 6.1, if you felt Table 6.1b is more reasonable than Table 6.1a, then your instincts served you well!

Table 6.1: Examples of types of relationships in any theory

(a) Deterministic relationships

| Relationship | Hypothetical Example |
|---|---|
| `occurs before` or `occurs after` | *“After working on a repetitive task for 10 minutes,* |
| `increases together with` or `decreases together with` | *“The size of a worker’s house increases as their nu* |
| `increases as the other thing decreases` | *“A worker’s risk of illness decreases as their salar* |
| `causes` or `is caused by` | *“Disrespect within a team causes a decrease in psy* |
| `moderates` or `is moderated by` | *“The relationship between your skill-level and your* |
| `mediates` or `is mediated by` | *“If you have more clients, you’ll have more incom* |

(b) Probabilistic relationships (non-deterministic relationships)

| Relationship | Hypothetical Example |
|---|---|
| `typically occurs before` or `typically occurs after` | *“After working on a repeti* |
| `typically increases together with` or `typically decreases together with` | *“The size of a worker’s ho* |
| `typically increases as the other thing decreases` | *“A worker’s risk of illness* |
| `typically causes` or `is typically caused by` | *“Disrespect within a team* |
| `typically moderates` or `is typically moderated by` | *“The relationship between* |
| `typically mediates` or `is typically mediated by` | *“If you have more clients,* |

### 6.5.2 Causal versus correlational relationships

You likely already have a great intuitive sense of what a causal relationship is. Intuitively, we understand that the sun causes the rooster to crow, and that the rooster’s crow doesn’t cause the sun to rise. In other words, the crow “listens to” or “takes orders from” the sun, but not the other way around.

In contrast, a correlational relationship *doesn’t necessarily* involve causation. In other words, some correlational relationships involve causation, but some don’t. Correlation is almost always defined numerically (which you will learn much more about in the *statistics* course of this degree program). For example, if I write down the time of day the sun rises, and also the time of day the rooster crows, and I do that every day for a week, I will see the sun rises and the rooster crows at around the same time as each other. Those two variables are correlated with each other. In this case, there also happens to be a causal relationship between them, because we know the sun causes the rooster to crow.

However, let's look at another example where two variables are correlated but neither causes the other. Consider a situation in which several companies in the USA keep track of how many `internship applications` they receive each day, and they keep track for an entire year. Looking back on the year's data, they notice the number of `internship applications` on any given day is correlated positively with the `outdoor temperature` of that day. In other words, hotter `outdoor temperatures` are associated with more `internship applications`, as depicted in Figure 6.4. Notice the arrow for the correlational relationship is *bidirectional*, meaning it has an arrow head on both ends, because correlation doesn't imply any causal directionality. All we can say is that the two variables are correlated positively with each other.

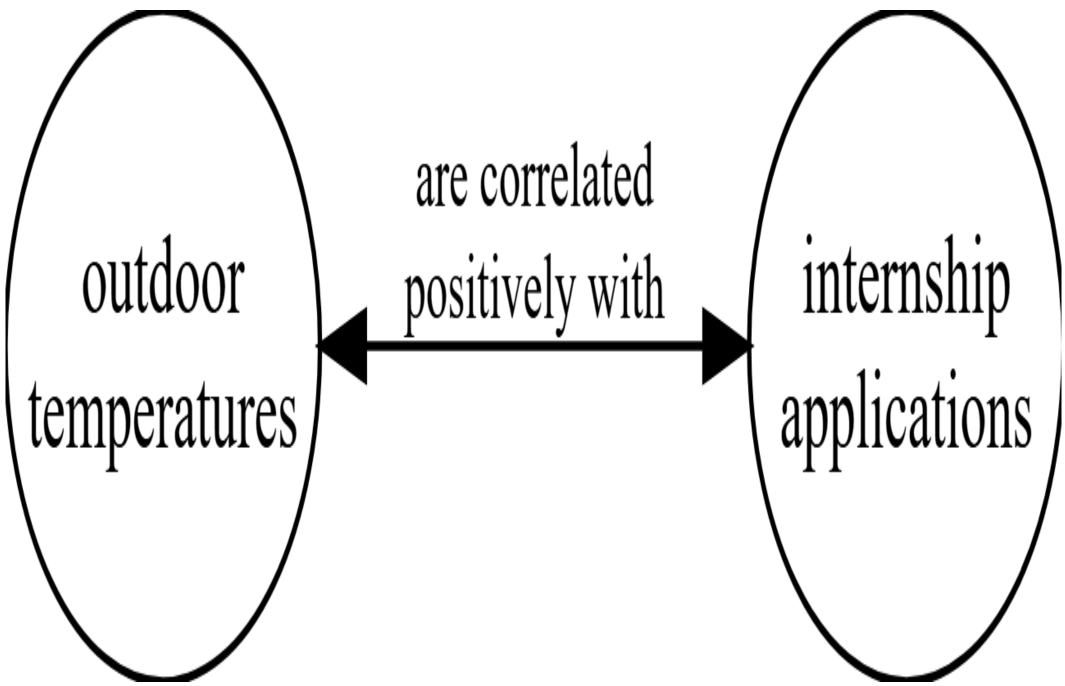



Figure 6.4: `Outdoor temperatures` and `internship applications` are correlated positively with each other, because they increase together. We're not saying either causes the other.

We know it would be ridiculous for the applications to cause cause the outdoor temperatures to go up, and it's unlikely the outdoor temperature is causing a chemical reaction in the applicants' brains that's causing them to want to apply to internships. Instead, the true reason behind the positive correlation is that the applicants are predominantly college students who mostly only have time for internships during their summer break, which happens to be the hottest part of the year. Thus, the onset of summer causes the applications to increase and it also causes the temperature to go up, as depicted in Figure 6.5.

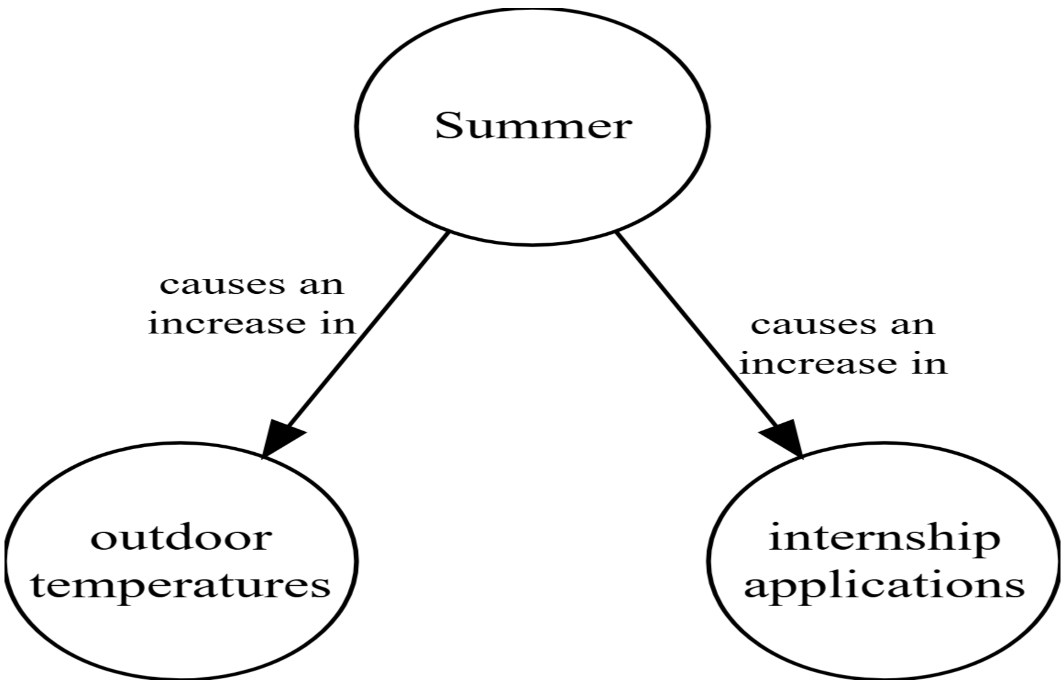



Figure 6.5: `Outdoor temperatures` and `internship applications` are correlated positively with each other, because `Summer` causes them both to increase.

In almost all scientific theories, the relationships have directionality, which is why we use directional arrows to depict relationships in our graphs. When we are expressing a correlational relationship, we use a bidirectional arrow as in Figure 6.4. When we are expressing a causal relationship, we use unidirectional arrows as in Figure 6.5. In causal relationships, the concept that the arrow is coming *out* of is called a **predictor** (e.g., *predictor variable*), also known as an **antecedent** (e.g., *antecedent variable*). The concept that the arrow is pointing *into* is typically called an **outcome** (e.g., *outcome variable*), also known as a **dependent variable**. For example, in Figure 6.5, `Summer` is a predictor, whereas `outdoor temperatures` and `internship applications` are both outcomes.

### 6.5.3 Moderated relationships

A **moderated relationship** (also known as a *conditional relationship* or a *modified relationship*) is any relationship (between two concepts in a theory) whose characteristic(s) can change, *depending on* one (or more) concepts. For example, consider the graph in Figure 6.6 which corresponds to this theory: *"How much time I have available will influence how much time I spend with that person"*.

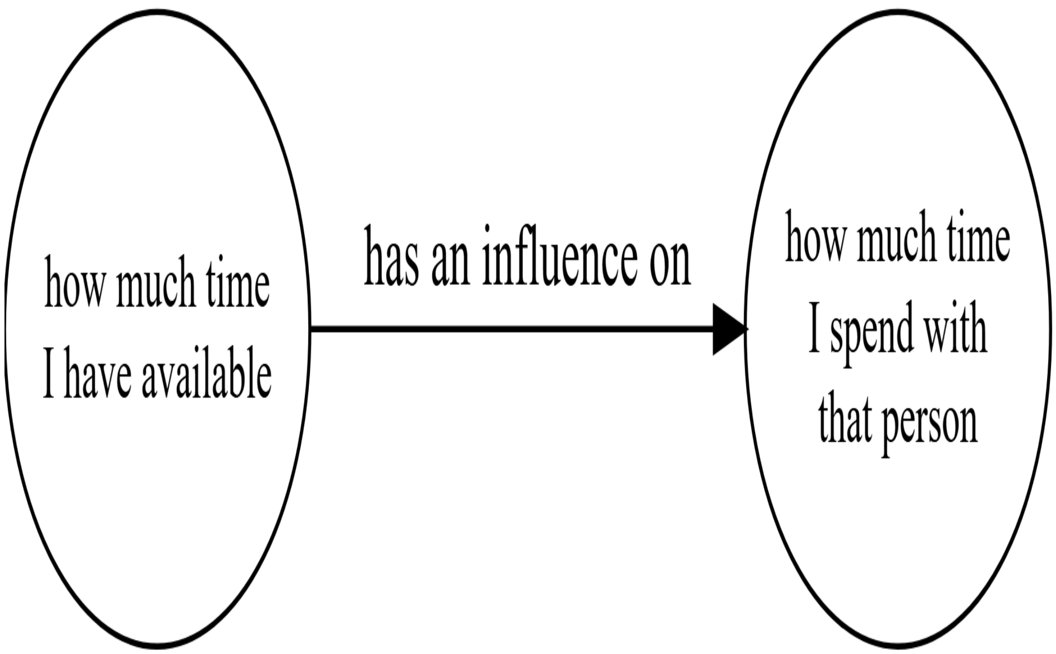



Figure 6.6: A simple theory without any moderated relationships.

That seems like a perfectly reasonable theory, but I think we can all relate to the fact that we'll probably end a conversation more quickly (or even avoid it altogether) if we don't enjoy that person. In other words, we could say: *"How much time I have available will influence how much time I spend with that person, but that depends on how much I enjoy that person".* This fancier theory is depicted in Figure 6.7.

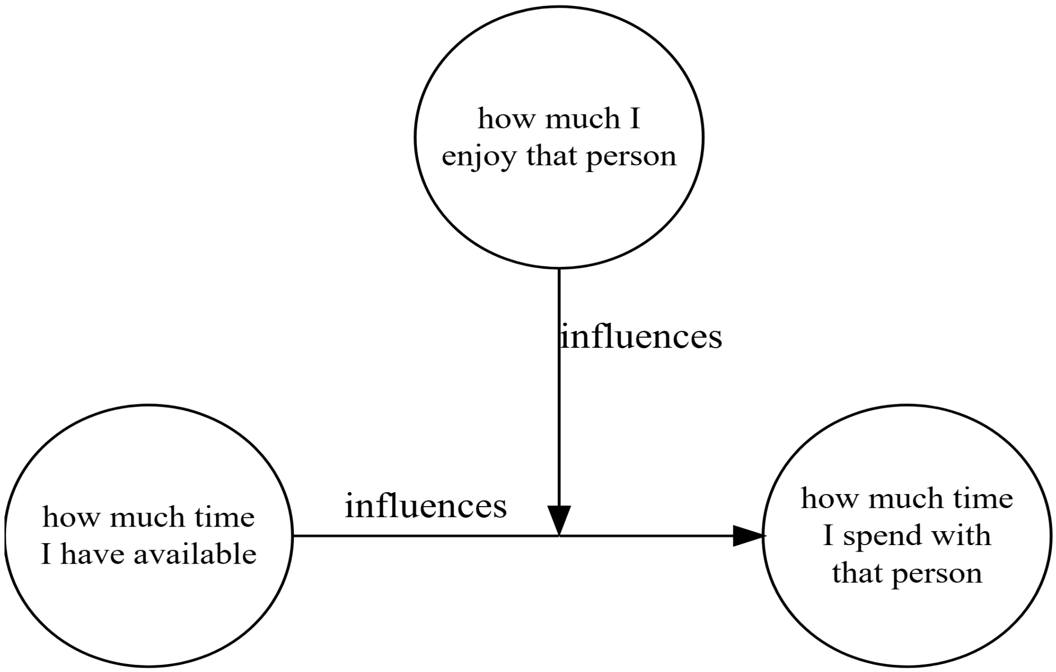



Figure 6.7: A basic theory with a moderated relationship.

In Figure 6.7, `how much I enjoy that person` is a *moderator* (also known as a *moderating predictor*), and it is moderating the relationship between `how much time I have available` and `how much time I spend with that person`. We can also use any of these equivalent verbal forms:

- The relationship between `X` and `Y` *is moderated by* `M`.
- The relationship between `X` and `Y` *is modified by* `M`.
- The relationship between `X` and `Y` *depends on* `M`.
- The relationship between `X` and `Y` *is conditional on* `M`.
- There is an *interaction effect* between `X` and `M` on `Y`.
- There is an *interaction* between `X` and `M` on `Y`.

- There is a *two-way interaction* between `X` and `M` on `Y`.

Another fun fact about moderated relationships is that there are always at least two equivalent ways to draw a graph for a moderation relationship. For example, in Figure 6.8, the position of `how much time I have available` and `how much time I enjoy that person` are swapped, compared to their positions from Figure 6.7. Thus, Figure 6.8 corresponds to this verbal expression: *"How much I enjoy that person will influence how much time I spend with that person, but that depends on how much time I have available"*. As you can see, it is conceptually identical to the theory expressed by Figure 6.7.

You could also specify a theory in which a relationship depends on two, three, four, or as many concepts as you like. For example: *"The relationship between `X` and `Y` is jointly moderated by `M` and `W`"*. In that case, it would be a *three-way interaction effect*, because `X`, `M`, and `W` are all interacting with each other in their relationship with `Y`.

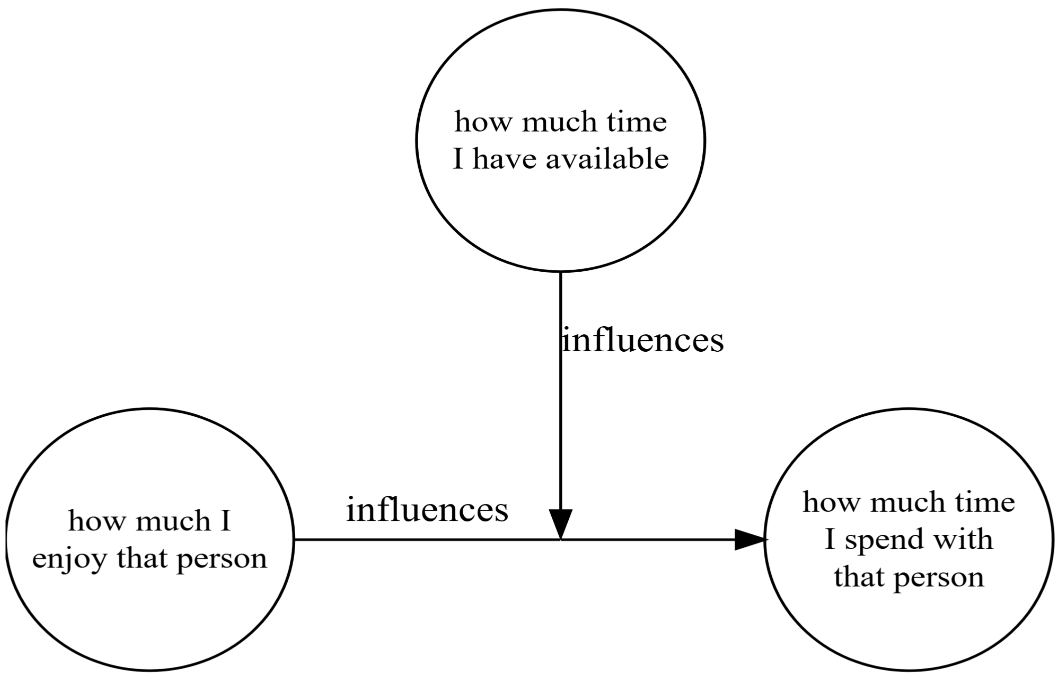
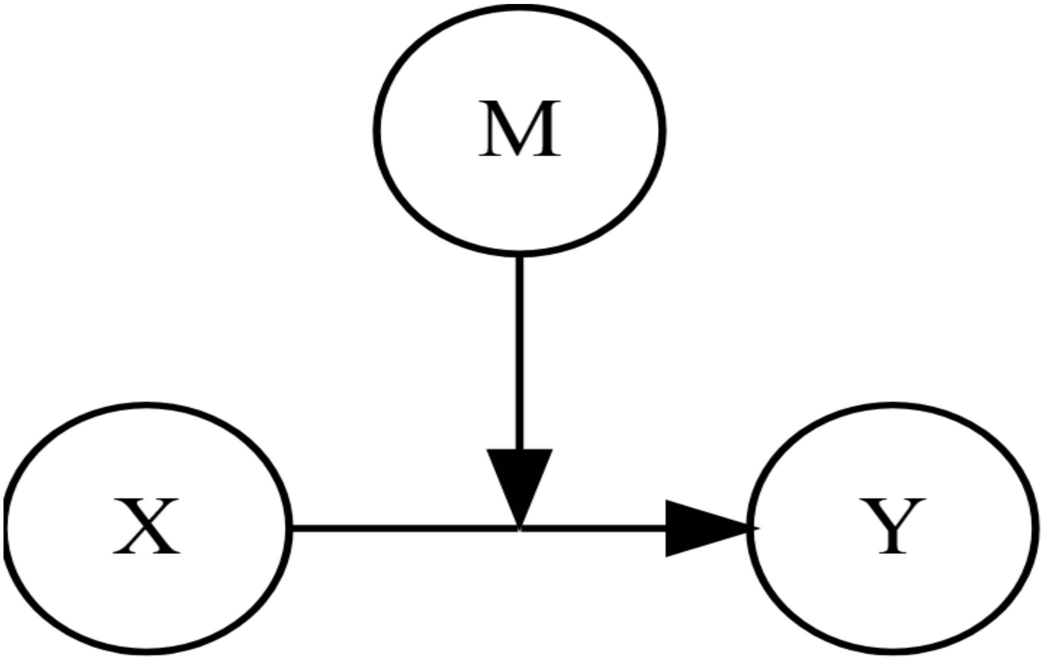
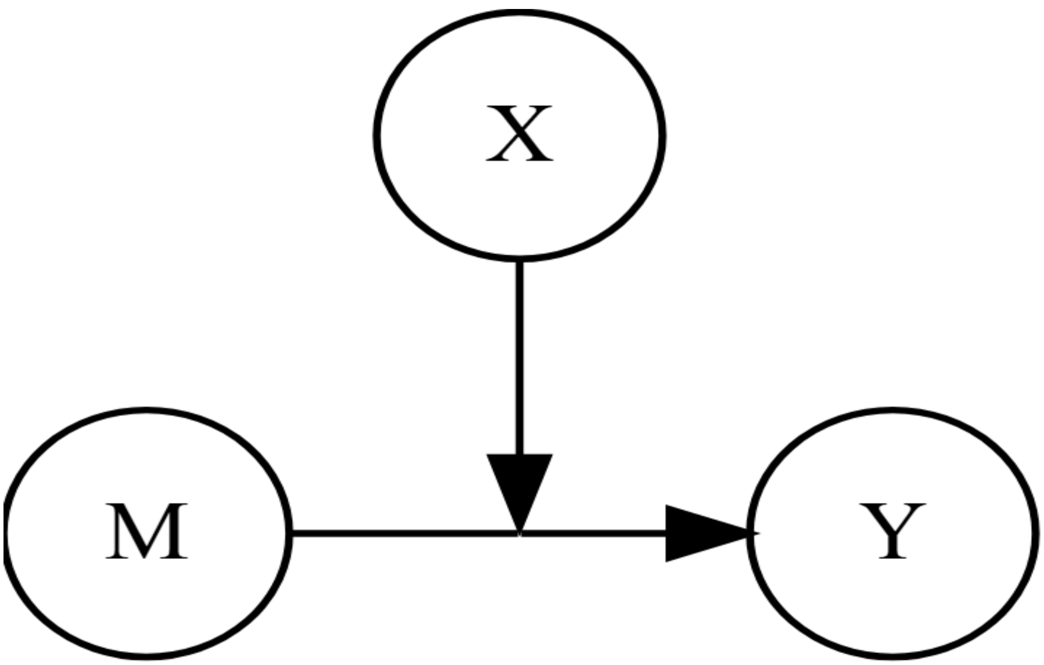



Figure 6.8: The theory expressed in this graph is identical to that of Figure 6.7.

Indeed, in any moderated relationship as the one depicted in Figure 6.9, it doesn't matter which of the interacting predictors are thought of as the "focal predictor" versus the "moderating predictor". The decision is based purely on how you prefer to think about your theory. Thus, Figure 6.9a is conceptually identical to Figure 6.9b.




(a) We think of `X` as the *focal predictor*, and we think of `M` as the moderator.




(b) We think of `M` as the *focal predictor*, and we think of `X` as the moderator.

Figure 6.9: Two equivalent expressions for the same interaction effect between the predictors `X` and `M` on the outcome `Y`.

### 6.5.4 Mediated relationships

Consider this theory from Table 6.1b: *"If you have more clients, you'll typically have more income, which means you'll typically go on more expensive vacations"*—expressed as a graph in Figure 6.11.

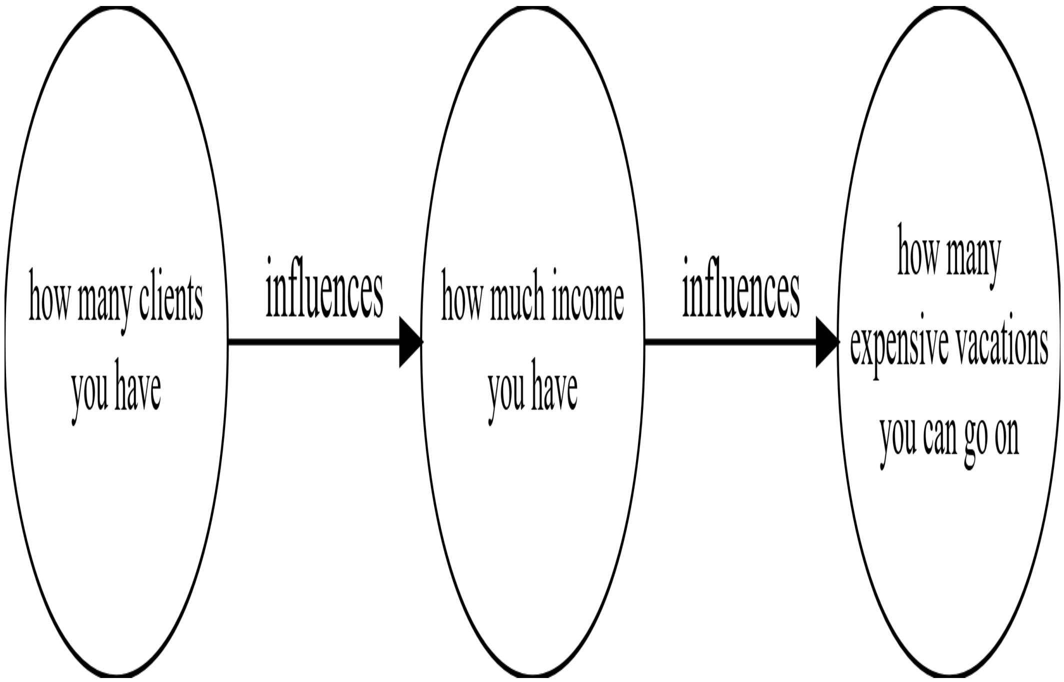



Figure 6.10: A simple mediation model.

In that theory, the relationship between `how many clients you have` and `how many expensive vacations you can go on` is *mediated* by `how much income you have` (i.e., `how much income you have` mediates the relationship). Thus, we would say Figure 6.11 is a *mediation model*, or a theoretical model about mediation.

A **mediated relationship** (also known as an *indirect relationship*) is any relationship (between two concepts in a theory) whose effect is transmitted through at least one intermediating concept (typically called the *mediating variable*, or the *mediator*).

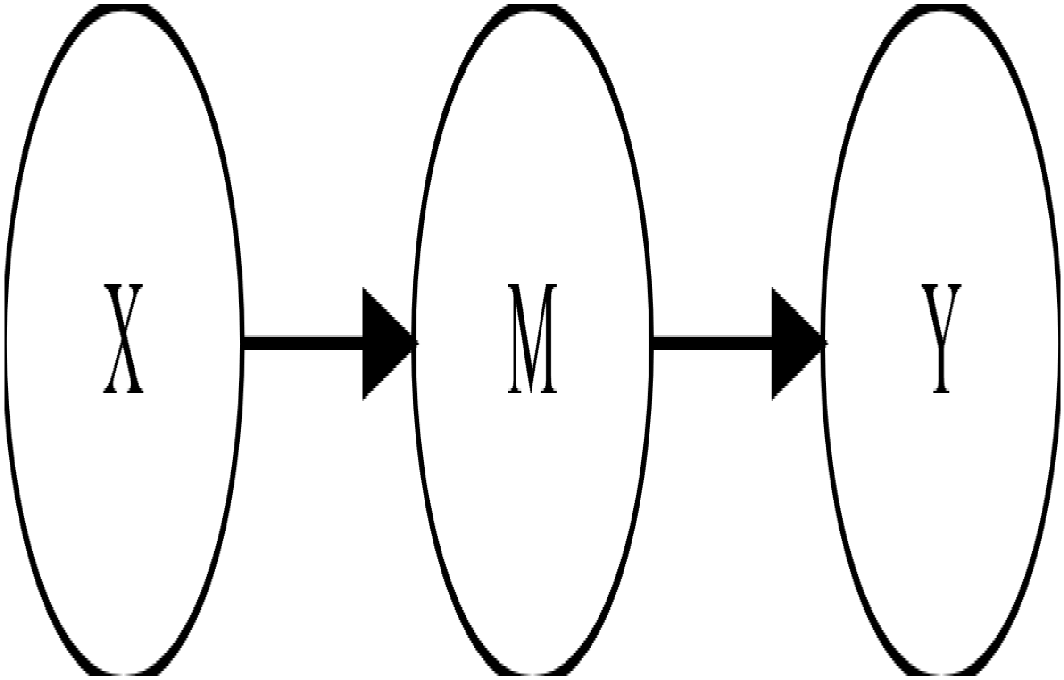



Figure 6.11: A simple mediation model.

These are equivalent verbal expressions about a mediated relationship:

- The relationship between `X` and `Y` is *mediated by* `W`.
- `W` *mediates* the relationship between `X` and `Y`.
- There is an *indirect relationship* between `X` and `Y`, through `W`.

It's also possible to specify a mediation model like the one in Figure 6.12, in which `X` has a direct effect on `Y`, while also having an *indirect* effect on `Y` via `W`.

In this context, we're not talking about the other sense of *mediation* which you might be familiar with from every-day English, in which a person serves as a mediator between two opposing parties in a dispute.

### 6.5.5 Combining moderated and mediated relationships

Consider the theory expressed in Figure 6.13, which includes moderation and mediation—these are often called *conditional-process* theories (Hayes, 2022), as well as other names (see: Edwards & Lambert, 2007).

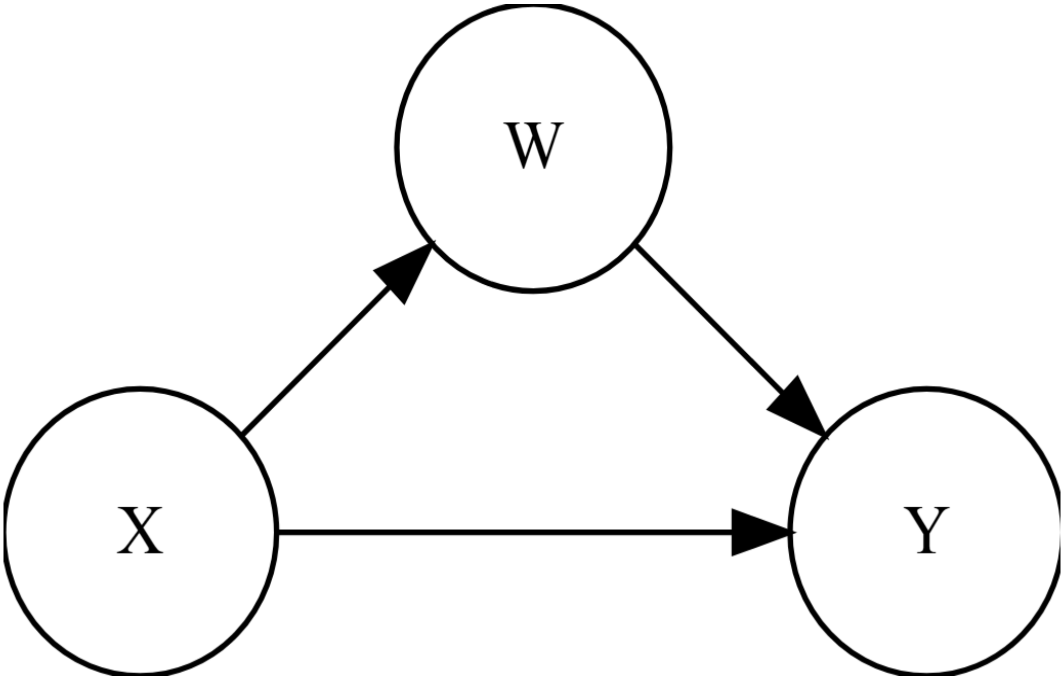



Figure 6.12: A slightly more complicated mediation model.

We could make it even more complicated by adding more predictors, including more mediators and more moderators anywhere in the graph, as well as more outcome variables. We might want to do that because we want the theory to more accurately reflect real life, which is often complicated. However, this brings us to a very important principle when we're building a theoretical model:

> "As the model matches reality better, it necessarily becomes less simple. Or, as it becomes simpler, it necessarily loses some of its match to reality" (Rodgers, 2010, p. 5).

Thus, a good scientific theory strikes a good balance between matching real life close enough to be useful, but not being so complicated that it becomes less useful because it only applies to a few specific situations or it is too complicated to understand.

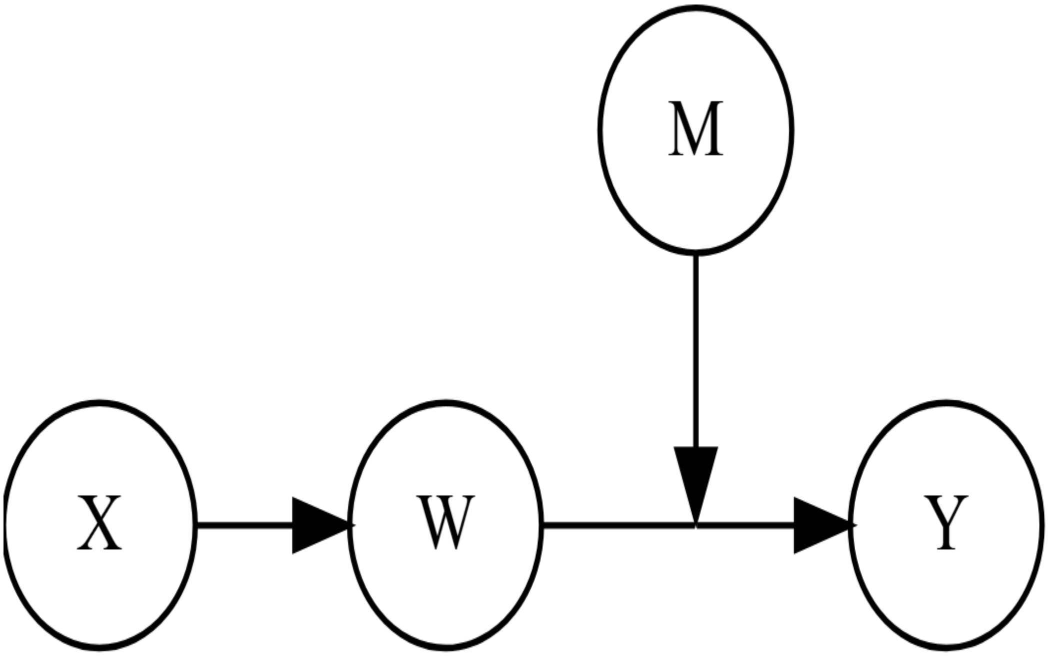



Figure 6.13: A conditional-process theory.

An analogy…

Here is a good analogy for why the best theoretical models strike a balance between simplicity and complexity:
A map is a miniature model of the world. We can add more details to the map to make it match the real world better. However, if we add too many details to the map, the map starts to become too big to handle. Imagine if the map included every pebble on every road, every grain of sand, every blade of grass. If we tried to print that map on paper, it would end up being near the size of the entire world itself! If it were stored digitally on a computer, it would require an unfeasible amount of digital storage space and computer processing power. Thus, a good map is one that is a *close enough* representation of the real world to suit our needs.

This principle segues nicely into the next section: what *else* makes a good scientific theory?

## 6.6 Necessary versus optional features of a good scientific theory

At minimum, a scientific theory must be:

1. empirically testable (e.g., empirically falsifiable);
2. logically internally consistent (i.e., it doesn't contradict itself).

The following features are typically *desirable* in a scientific theory, though not necessarily *required*:

1. increases our understanding of the world;
2. novel (i.e., it's something new);
3. relatively broad in scope;
4. easily understood and communicated to others;
5. appropriately parsimonious (i.e., Occam's razor);
6. consistent with other accepted theories that have achieved consensus;
7. accepted by the broader scientific community;
8. stimulates research;
9. addresses a real-world problem.

## 6.7 Finally: Hypotheses

Some scholars believe a *hypothesis* (plural: *hypotheses*) is fundamentally no different than a theory. In some sense, I agree with that. But I also know most scholars—especially in the behavioral sciences—tend to use the word *hypothesis* to connote the meaning from APA's online *Dictionary of Psychology*:

> **hypothesis**: an empirically testable proposition about some fact, behavior, relationship, or the like, usually based on theory, that states an expected outcome resulting from specific conditions or assumptions.

As you can see, that definition of *hypothesis* is consistent with our definition of a theory that is *empirically testable.* Thus, in the parlance of behavioral scientists, one of the most important features of a hypothesis is that it is *empirically testable*.

> For example…
>
> Here's an example to help reify the distinction between a *hypothesis* versus a *theory*. Consider this theory: *"An increase in an employee's happpiness will tyically cause an increase in their performance."* Now, suppose I don't yet have the resources to design a study that can test the causal mechanism, but I can design a self-reported questionnaire study that can measure whether `employee happiness` is correlated positively with `employee performance`. In that case, my hypothesis might be: *"Self-reported employee happiness will be correlated positively with self-reported employee performance."* As you can see, that hypothesis is more clearly empirically testable than the original theory statement. If you feel like this hypothesis could be written with even greater specificity, you're right! That is a topic in Chapter 7.

Typically, after you describe your theory using words, equations, and/or graphs, you would write one or more hypotheses that can be empirically tested. Then—assuming your hypothese are logically derived from your theory—when you or anyone else conducts an empirical study to test your hypotheses, that process accumulates evidence about whether your *theory* is true or not.

> Research Questions vs. Hypotheses
>
> Sometimes when we're conducting an empirical study, we might seek to collect evidence to help us answer some questions that aren't necessarily tied to any hypotheses. In that case, we might write one or more research questions (regardless of whether or not we also have any hypotheses in our study) that we seek to address via an empirical study. A research question is exactly what it sounds like: it's a question in which you're not making a statement nor prediction about what results you expect.

## 6.8 Suggested Readings

If you enjoyed this chapter, you'll likely enjoy Jaccard and Jacoby's (2020) book: *Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists*. It is part of the The Guilford Press' *Methodology in the Social Sciences* series, which contains *excellent* books about research methods.

McGuire (1997) is a classic article with excellent tips to help inspire your creativity for creating a hypothesis.

I can't help but share this funny bit of trivia: James goes by "Jim" and his middle name is Jay. Jacob (1940–2018) went by "Jack". Thus, the authors are James "Jim" Jay Jaccard and Jacob "Jack" Jacoby. It doesn't get better than that!

# 7 Psychometrics and Sampling

## 7.1 Psychometrics

In Section 6.4, I briefly mentioned that psychologists invented techniques (especially starting in the early 1900s) to help measure subjectively perceived things and other abstract psychological cónstructs. We call those techniques *psychometric techniques*, and the topic of measuring such psychological cónstructs is known as **psychometrics**. You will receive an entire course on psychometrics later in this degree program, but I will give you a gentle introduction to some of its fundamental ideas here (because repeated exposure and repeated recall are excellent ingredients to aid your learning).

Recall from Chapter 6 that IWO Psychologists are often interested in using quantitative methods to study one or more cónstructs and/or other variables. Before we can use quantitative methods to study those concepts, each of them must be measured in a manner that produces a numeric measurement. However, before we can numerically measure those concepts, we must clearly define each of the concepts we're interested in.

An expert in psychometrics is known as a *psychometrician*.

In scientific research, two of the most common types of definitions for any cónstruct are: *conceptual definitions* and *operational definitions*. A **conceptual definition** describes the cónstruct in clear detail, to prevent other people from misunderstanding what the researcher means when they refer to that cónstruct.

For example...

If one of the cónstructs I'm studying is `the performance of workers at ABC Company`, I might provide this conceptual definition: *"For the purposes of this study,* `the`

> *`performance of workers at ABC Company` is defined as the overall quality of all work-related behaviors and subsequent products and/or services produced by the workers".*

An **operational definition** describes the empirical evidence that the researcher will use as a measurement of the cónstruct. The process of creating an operational definition is known as **operationalization**. When we create an operational definition for a cónstruct, we say we've *operationalized* our cónstruct.

> For example...
>
> If one of the cónstructs I'm studying is `the performance of workers at ABC Company`, these are just some of the countless types of operational definitions I might choose if I want the data to be numeric:
>
> - *"For the purposes of this study, `the performance of workers at ABC Company` will be operationalized as `the number of client reports delivered by the workers within the past 12 months`."*
> - *"For the purposes of this study, `the performance of workers at ABC Company` will be operationalized as `the average of the performance ratings given by the workers' supervisors via our questionnaire`."*
> - *"For the purposes of this study, `the performance of workers at ABC Company` will be operationalized as `the average of the performance ratings given by the subject-matter experts we recruited to review the workers' client reports`."*

In general, there are almost always numerous (even countless) operational definitions that we could choose for any given cónstruct. Our chosen operational definition(s) can involve any of several popular types of measurement procedures:

- **Self-report measure**: The person being studied provides information about themself. For example, a worker might answer questions about themself.

- **Other-report measure**: A person provides information about someone else. For example, the researcher asks a supervisor to answer questions about an employee.
- **Observational measure**: This is technically just a type of *other-report measure*, in which the researcher or another trained observer or rater is the person providing the information. For example, a trained research assistant might observe how team members talk during a team meeting.
- **Retrospective measure**: A person provides information about something that occurred in the past. For example, a worker might answer questions about a performance event that occurred in the past month.
- **Concurrent measure**: A person provides information immediately while an event is occurring. For example, a worker might narrate what they perceive while they are observing something occur in the moment.
- **Prospective measure**: A person provides information now about a future event (which might be either purely hypothetical or very likely). For example, a worker may provide information about what they believe would be the impact on the workers' attitudes if the company changed its scheduling policy.
- **Physiological measure**: Also known as *biological measure*: Information such as heart rate, eye movements, hormone levels, brain activity, etc., is collected.

Notice the types of measurement procedures listed immediately above aren't necessarily mutually exclusive. For example, if you asked a supervisor to give you information today about a worker's performance from the past month, that would be a *retrospective other-report measure*.

Furthermore, the operational definition that we choose for any cónstruct can have an influence on the **scale of measurement** (also sometimes called *level of measurement*) for our measure. The *scale of measurement* for any particular measurement process is simply a special description of the type of data that can be produced by that measurement process. In science—including IWO Psychology—the phrase *scales of measurement* typically refers to a specific typology of four types: 1. categorical, 2.

ordinal, 3. interval, and 4. ratio. Perhaps the best way to explain those four scales of measurement is via examples, as described in Table 7.1.

Table 7.1: The four scales of measurement.

| Scale of measurement | Example |
| --- | --- |
| **Categorical**: The variable's data do not form a dimension. That is, they cannot be put into a logical sequence of increasing/decreasing essence of the variable. | The `occupation type` variable would be on a *categorical scale of measurement* if it had these types of data: `engineer`, `marketer`, `accountant`. |
| **Ordinal**: The variable's data form a dimension, but the distance between each possible "level" along the dimension might not be equal. | The `job satisfaction` variable would be on an *ordinal scale of measurement* if it had these types of data: `dissatisfied`, `somewhat dissatisfied`, `somewhat satisfied`, `satisfied`. As you can see, those data form a dimension, but we cannot guarantee that the subjectively perceived "size" of the difference between `dissatisfied` versus `somewhat dissatisfied` is the same as the subjectively perceived "size" of the difference between `somewhat dissatisfied` and `somewhat satisfied`. A respondent might believe there is a small difference between the former, but a large difference between the latter. Another example would be a variable whose data are `never`, `sometimes`, `usually`, `always`. Again, we can't guarantee that the subjectively perceived "distance" between each level is equal. |

| Scale of measurement | Example |
|---|---|
| **Interval**: The variable's data form a dimension, and the distance between each possible "level" is equal, but there is no *natural zero* (also known as a *true zero*). In other words, an interval scale cannot measure the true absence of the thing it is measuring (typically because the true absence is undefined for that thing). | The `extroversion` variable would be on an *interval scale of measurement* if it had this type of data: `0`, `1`, `2`, `3`, `4`. As you can see, it would be measured on a 5-point scale from `0` to `4`, and the distance between each "level" is equal. However, a score of `0` on that scale doesn't mean the person has no extroversion, because there is no such thing as *the complete absence of extroversion* (i.e., it is conceptually undefined, because even the most introverted people interact with other people eventually). In other words, `0` merely denotes the lowest level on that scale (e.g., perhaps "very little extroversion"). Another example would be a `performance rating` variable on a scale of `0`, `1`, `2`, `3`, where `0` is `very poor` and `3` is `excellent`. Again, `0` on this scale doesn't denote the complete absence of any performance at all, thus it isn't a *natural zero*. |

| Scale of measurement | Example |
|---|---|
| **Ratio**: The variable's data form a dimension, and the distance between each possible "level" is equal, and there is a *natural zero* (also known as a *true zero*). In other words, a ratio scale can measure the true absence of the thing it is measuring. | The `how many coworkers` variable would be on a *ratio scale of measurement* if it had this type of data: `0`, `1`, `2`, `3`, … `99`, `100`, …. If the question is *"how many coworkers do you have"*, the datum `0` would indicate the true natural absence of coworkers if the person is self-employed. If the question is *"how many coworkers are your enemies"*, the datum `0` could truly represent zero enemies. Another example would be the `number of correct answers on a hiring test` variable. That type of variable could have a natural zero (e.g., none of the questions were answered correctly). |

In general, if a cónstruct is a dimension, you can choose whether you want to measure it via an ordinal or interval scale of measurement. But whether or not you can use a ratio scale of measurement depends on whether the dimension contains a natural zero (i.e., whether it's possible to measure the true absence of that cónstruct).

The scale of measurement for any cónstruct will largely determine the types of analyses you can do on the data from your measurements. For example, if all of your data are on a categorical scale of measurement, then you won't be able to compute a correlation coefficient nor perform a regression analysis (both of which are popular techniques you will learn in the statistics course later). In general, there are more types of numeric operations you can do on your data as you move up the ladder of scales of measurement: categorical data allow the fewest types of numeric operations; ordinal data allow more than categorical data allow; interval data allow even more; and ratio data allow the most. Thus, whenever you have the ability to choose your measurement process, you should typically choose the one that will allow you to have the highest scale of measurement.

### 7.1.1 Validity and Reliability

#### 7.1.1.1 Validity

Two of the most important topics in psychometrics are *validity* and *reliability* (and, in fact, reliability is part of validity). In science—including IWO Psychology—it may seem like the word *validity* has a dizzying variety of meanings. However, there is a unifying fundamental meaning. To help clarify things, a good place to start is with a definition of *validity* from the *Oxford English Dictionary*:

> The quality of being well-founded on fact, or established on sound principles, and thoroughly applicable to the case or circumstances; soundness and strength (of argument, proof, authority, etc.). (Oxford English Dictionary, 2025a)

Notice the above definition includes the adjective *sound* and the noun *soundness*. The same dictionary provides this definition of *sound*:

> In full accordance with fact, reason, or good sense; founded on true or well-established grounds; free from error, fallacy, or logical defect; good, strong, valid. (Oxford English Dictionary, 2025b)

The same dictionary provides this definition of *soundness*:

> The quality or fact of being in harmony with solid or well-established principles or facts. (Oxford English Dictionary, 2024)

Thus, in the domain of Psychology, a good definition of **validity** is:

> the extent to which a structure (e.g., a tool or psychometric instrument) or process (e.g., a method or experiment) is well-founded on or in accordance

Sometimes a researcher will have a good reason to intentionally choose an ordinal scale of measurement instead of interval or ratio. For example, if they are studying workers at a small company, the researcher might intentionally use an ordinal scale of measurement to measure `age` (e.g., `18-24`, `25-34`, `35-44`, …) because they want to avoid making the participants feel like someone is going to be able to identify them in the dataset via a combination of their specific information (e.g., they might be the only 18-year-old person at the company).

> with facts or reasonable principles, and appropriate for a specified context or purpose.

As you can see from that definition, the validity of any structure or process may depend on the principles you deem *reasonable*, and your specific context or purpose.

> For example...
>
> Using a poorly designed questionnaire on students in a classroom might be a very *invalid* process if my goal is to measure their extroversion, but it can be a very *valid* process if my goal is to teach them about the pitfalls of poor psychometric design!
> Here's another example: a new firefighter readiness exam is developed, and using it on experienced trainees is a highly valid way to accurately evaluate the trainees' readiness for graduation. However, the exam is extremely expensive to administer, such that it is way above the city's budget. In this case, the use of that exam is valid for evaluating the trainees' readiness, but it is *invalid* for doing so within the city's budget.
> As you can see, the validity of these two procedures depends on the intended purpose in each context.

Especially in IWO Psychology, Table 7.2 highlights typical structures or processes for which we may want to evaluate their validity.

Table 7.2: Typical structures or processes for which we may want to evaluate their validity.

| Structure or Process | Example |
|---|---|
| A measurement instrument or process for measuring a cónstruct `X`. | *We want to know the validity of using this psychometric instrument on members of our target population for obtaining an accurate measurement of work passion.* |

| Structure or Process | Example |
|---|---|
| A process for figuring out whether there is any causal influence of `X` on `Y`. | *We want to know the validity of our randomized controlled experiment's design for figuring out whether giving verbal praise towards long-tenured employees will have a meaningful impact on their work motivation.* |
| The structure of a theoretical model `X` as a useful representation of reality. | *We want to know the validity of our JD-R model in terms of how well it fits the empirical data.* |
| The structure of an inference `X`, or a process for inferring `X`. | *We want to know the validity of the logical structure of our inference."* Another example: *"We want to know the validity of our process for inferring the most common types of rewards in organizational contexts.* |
| A process for studying a topic `X`. | *We want to know the validity of using an ethnographic research method to study the emergence of conflict within teams.* |

In the psychometrics course later in this degree program, you will learn about many different labels that psychologists created to describe the various types of evidence about validity, especially the **validity of a measurement process** (also known as **measurement validity**) intended to obtain accurate measurements.

### 7.1.2 Reliability

As I mentioned earlier, *reliability* (also known as *consistency*) is merely one part of validity. In all domains of Psychology, the concept of *reliability* is applied almost exclusively to *measurement instruments* and *measurement processes*. A good definition for **measurement reliability** (also known as the **reliability of measurement**) is this:

> The extent to which a measurement process produces measurement data that are consistent with themselves or consistent with measurement data from other measurement processes measuring the same cónstruct or object.

Indeed, the concept of *reliability* can be applied to other processes that aren't about measurement, but it's uncommon in IWO Psychology.

An important distinction to keep in mind is that a valid measurement process must be reliable, but a reliable measurement process isn't automatically valid. The truth of that sentence may be clearer when translated like this: an appropriate measurement process must be consistent, but a consistent measurement process isn't automatically appropriate.

**For example...**

Imagine there is an emergency-responder readiness exam that consistently gives lower scores to women than men, because the exam's scoring formula places a big emphasis on how many pull-ups the person can do. In this case, the process of using that exam may be *reliable* at producing the same score, over and over, for each examinee regardless of how many times they retake the exam, but that exam process is *inappropriate* (i.e., *invalid*) for evaluating the *true* readiness of the person to be an emergency-responder—because their job has virtually nothing to do with being able to do pull-ups!
A valid measurement process must be reliable, but a reliable measurement process isn't automatically valid.

When scientists—including IWO Psychologists—want to collect evidence about the reliability of a measurement process, it's perfectly normal to collect evidence that doesn't completely address the validity of that measurement process. This is because it is often more straightforward to separately collect evidence that focuses on reliability and then separately collect other evidence that focuses on validity, rather than try to collect evidence that completely addresses reliability and validity together. For that reason, when examining the reliability of a measurement process, IWO Psychologists aren't necessarily focused on whether the measurement data produced by the measurement process are *accurate* nor otherwise *valid*—because the validity of the measurement process and its measurement data will be evaluated later via a separate series of validation tests.

In all fields of Psychology, including IWO Psychology, these are the most popular types of reliability:

A *validation study* is a research project (i.e., a study) to evaluate the validity of a process for some intended outcome. A validation study would typically include several validation tests, to produce multiple lines of evidence.

- Reliability across two or more timepoints with the same measurement instrument and measurement process (also known as **test-retest reliability**, or *consistency across time*). For example: a psychometric questionnaire to measure a person's cognitive abilities is administered today, and then again next month. *We'd like to know how much the measurement data from the two timepoints are similar to each other.* This is especially relevant when we are measuring something that we expect shouldn't change much across the timepoints, such as an adult's cognitive ability in the span of a year, or a person's overall satisfaction with their life in the span of a month.
- Reliability across two or more measurement instruments or measurement processes:
  - Reliability across two or more persons who are separately measuring the same object (e.g., a person, cónstruct, event, etc.) at the same time:
    * Reliability across two or more raters who are each giving a rating at the same time (also known as **interrater reliability**). For example: Two raters are separately rating Alex's performance at the same time. *We'd like to know how similar their ratings are to each other.*
  - Reliability across two or more measurement instruments that are expected to be similar to each other:
    * A special and very important type of this reliability is known as **internal reliability** (also known as **internal consistency reliability**). This specifically refers to the reliability across two or more items in a questionnaire that are all supposed to measure the same concept. In other words, the researcher is treating each item as a "separate instrument", even though the re-

spondent answers all items together in a single questionnaire. For example, if I have a psychometric questionnaire with 10 items that are each intended to measure a worker's `job-satisfaction`, I would want to know how much the data from each of those 10 items are similar to each other. One of the most common methods to calculate internal reliability is known as **coefficient alpha** (also known as **Cronbach's alpha**, or **Cronbach's** $\alpha$), which usually ranges from 0 to 1 (i.e., $0 \leq \alpha \leq 1$) where 1 represents perfect internal consistency. As a rule of thumb, a coefficient $\alpha$ of .70 or higher is typically deemed as acceptable. In the statistics course later, you will learn how to calculate *coefficient* $\alpha$.

* Another common type is called **equivalent-forms reliability**, which typically refers to the reliability between two or more multi-item psychometric questionnaires (or interviews) that we believe are equivalent to each other. For example: imagine we have two versions of a psychometric protocol to measure a person's level of extroversion. *We'd like to know how much the extroversion scores from both versions are similar to each other.*

In all fields of Psychology, including IWO Psychology, the concept of reliability is almost exclusively applied to *numeric* data. Thus, reliability is almost always expressed numerically as a single number called the **correlation coefficient**, typically denoted with the letter $r$. A correlation coefficient is always somewhere between -1 and 1 (i.e., $-1 \leq r \leq 1$). For example, you might see $r = 0.23$, or $r = -0.691$. When $r = 0$, there is no consistency at all among the measurement data. When $r = 1$ or $r = -1$, there is perfect consistency among the measurement data. In Psychology—including IWO Psychology—we almost never see $r = 1$ nor $r = -1$. In Chapter 8, I will describe how to calculate $r$.

## 7.2 Sampling: obtaining a sample from a population

After you've operationalized your cónstructs and you've described your intended measurement process to measure those cónstructs, you'll need to figure out how you're going to get access to the people or other things (e.g., events) you want to measure!

To talk about that, let's get acquainted with some vocabulary.

In the vocabulary of math, a **set** is any collection of things, and the *things* in a set are called the **elements** of that set.

> For example...
>
> A group of people is a *set* whose elements are people. A collection of training programs is a *set* whose elements are training programs. A collection of emotions is a *set* whose elements are emotions.

A **subset** is a set whose elements all come from another set (called the *superset* of that subset).

> For example...
>
> The set of M.S. I–O Psychology students at this university is a *subset* of all the students at this university.

In the vocabulary of statistics (which is part of math), a **population** is any set we want to understand.

> For example...
>
> It could be the set of all biomedical startup companies in Alabama from the year 2000 to 2025, or the set of all nurses in the USA, or the set of all employee meetings that occur at ABC Company during this summer.

A **sample** is any subset of a population. In the vocabulary of statistics, an element of a sample is typically called a *sample element*, or a *sample unit*. The process of collecting a sample is called **sampling**. Ideally, sampling is carefully planned, typically with the goal of obtaining a sample that is similar to the target population you are interested in.

> For example...
>
> If your target population is the set of all biomedical startup companies in Alabama from the year 2000 to 2025, you could obtain a complete list of all of those companies, and then randomly select 100 from that list. Thus, you would have a sample of 100 such companies.

Recall from Section 2.2.2.1 I said scientists (including IWO Psychologists) typically seek to make *universal statements* that are true, and that most of the universal statements that scientists seek to make are about sets that contain such a large number of elements that it would be practically and/or literally impossible to observe every element of those sets to verify the truth of any universal statement about those sets.

Because we typically don't have access to observe all of the elements of the large populations we are interested in, we typically study a sample from that target population instead of directly studying the population. Thus, we seek to make a good *inference* about the target population, based on the sample. One of the most popular techniques for using samples to make inferences about a target population is called *statistics*.

Recall from Section 2.2.2.1: A **universal statement** is any expression that says a specified property is true for *all elements* in a specified set. For example, each of the statements in Table 2.2 is a universal statement because they each make a claim about all elements in a specified set.

I will dive deeper into statistics in Chapter 8.

> When the company *is* the target population...
>
> Sometimes, practitioners of IWO Psychology are hired to make inferences about the employees at a specified company. In that case, the target population is the employees of that company. If the company is small enough, the IWO Psychologist might have sufficient time and resources to conduct the study on all of the employees of the company. In that case, there is no need to select a sample. The IWO Psychologist already has data on their entire target population.

### 7.2.1 Population Parameters vs. Sample Statistics

In the vocabulary of the field of statistics (as a broad area of theory and praxis), a **parameter** is any numeric characteristic that describes a *population*, whereas a **statistic** is any numeric characteristic that describes a *sample*. Typically, we use sample statistics to try to *estimate* their corresponding population parameters, because we typically don't have the ability to measure the entire target population.

> For example...
>
> If my target population is every employee at ABC Company, I might obtain a sample of 100 of those employees. Any numeric characteristic of the sample has a corresponding numeric characteristic for the entire company, as illustrated in Table 7.3.

As you can see, the word *statistic* has at least two meanings: it can be a number that describes a sample, or it can refer to the entire field of statistical theory and statistical praxis (i.e., the practice of inferring something about a target population, based on numeric summaries of samples).

Table 7.3: Examples of sample statistics vs. their corresponding population parameters.

| Sample statistics ($n = 100$ employees at ABC Company) | Population parameters (all employees at ABC Company) |
|---|---|
| Average level of extroversion across those 100 employees | Average level of extroversion across every employee at ABC Company |
| Highest salary across those 100 employees | Highest salary across every employee at ABC Company |
| Lowest salary across those 100 employees | Lowest salary across every employee at ABC Company |
| Combined total number of years those 100 employees have worked at ABC Company | Combined total number of years every employee has worked at ABC Company |

Ideally, we would like all of our sample statistics to match their corresponding population parameters (or at least be close enough). Any difference between a sample statistic and its corresponding population parameter is known as **sampling error**, which we will discuss further below.

### 7.2.2 Some typical methods for sampling

Almost always in science, the goal of sampling is to obtain a **representative sample** (also known as an *unbiased sample*)—that is, a sample that is similar to the target population on all characteristics that are relevant to our study.

When our sample isn't representative of our target population, we say we have an **unrepresentative sample** (also known as a *biased sample*). The reason why we don't want an unrepresentative sample is because we will have a greater amount of sampling error—recall earlier in this chapter I said *sampling error* is the difference between any sample statistic and its corresponding population parameter. Any time we are using a sample instead of our entire target population, we are practically guaranteed to have some amount of sampling error. However, there are things we can do to try to minimize the amount of sampling error in our sample.

One of the best ways to minimize the amount of sampling error in our sample is by using any sampling method that is likely to produce a *large representative* sample.

**In case you're curious...** We almost *never* know any of the parameters of our target population, because typically the only way to know that is by directly observing/measuring our entire target population (and, if we were able to do that, we probably wouldn't be using a *sample* in the first place). Without knowing the parameters of our target population, we can't directly calculate the sampling error in our sample. However, we can try to *estimate* the amount of sampling error in our sample by using statistical techniques that are beyond the scope of this textbook. If you are curious to learn about how to estimate the sampling error, you can either wait until you take the statistics course next semester, or you can learn about it in any decent statistics resource on the internet.

To help us select a large representative sample, there are two different categories of sampling methods: one is called *known-probability sampling*, and the other is called *unknown-probability sampling*. I discuss both of these in the following two subsections.

#### 7.2.2.1 Known-probability sampling

In **known-probability sampling** (also known as **random sampling**, **probability sampling**, or **probabilistic sampling**) we typically must have a **sampling frame**, which is a list that represents everyone in our target population.

For IWO Psychologists, the most common situation in which you might use known-probability sampling is when your target population is the workers at some specified company or other type of organization. Otherwise, if your target population transcends any specific company (e.g., "*all nurses in the USA*" or "*all working adults in Western cultures*"), it becomes practically unfeasible to obtain a sampling frame. In that case, you would need to use *unknown-probability sampling*, which I describe in Section 7.2.2.2.

For example...

If your target population is every employee at ABC Company, your sampling frame could be a list of the names of each employee. If your target population is all of the meetings that occur at ABC Company this year (because your study is focused on making an inference about meetings, rather than just about employees), then your sampling frame could be a list of every meeting that occurred this year at ABC Company.

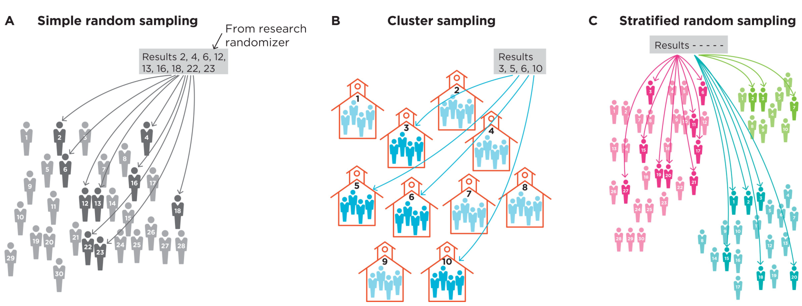



Figure 7.1: Three of the most common yet effective types of known-probability sampling methods (Image source: Morling, 2021)

When you have a sampling frame, one of the simplest and effective sampling methods to obtain a representative sample is called **simple random sampling** (often written as: **SRS**) without replacement (depicted in panel *A* of Figure 7.1). It is effective because it ensures everyone in the target population has an *equal probability of being included into the sample*. Simple random sampling can be achieved with any computer program that generates a random integer. You would assign a unique integer as an ID number to each element of your sampling frame (e.g., 1, 2, … 700). Then, each time you generate a random integer from the computer, that will be the element ID in your sampling frame that you must select to be included into your sample. You would continue generating a new integer and adding that element into your sample until you reach your desired sample size.

> For example…
>
> If your sampling frame contains 700 people, and the computer-generated random integer is 23, then you must include into your sample the 23rd element from your sampling frame. Then you generate a new random integer and continue adding the corresponding element from the sampling frame until you reach your desired sample size.

Before the ubiquity of computers, people would use special books that contained large lists of random numbers, and that's how the person would select a random number.

Another excellent known-probability sampling method (when you have a sampling frame) is called **stratified sampling** (also known as *stratified random sampling*, or *stratified probabilistic sampling*)—depicted in panel $C$ of Figure 7.1. In that method, we use one or more *stratification variables* to divide our target population's sampling frame into mutually exclusive groups (called *strata*, which is plural for *stratum*), such that any given element belongs to one and only one stratum. Then we randomly select elements from each of those strata. We could use any characteristic(s) of our sampling frame as the stratification variable(s) (e.g., gender, ethnicity, age, income bracket). We must also decide whether we want the number of elements in each stratum of our final sample to be proportional to their numbers in the target population (called **proportional stratified sampling**), or not (called **disproportional stratified sampling**).

One of the main advantages of stratified sampling (especially proportional stratified sampling) is that it helps us avoid ending up with a sample that is missing sample elements from any given stratum. In other words, by taking a random sample from each stratum, we ensure that no stratum is missing from our sample—thus helping to ensure our sample is more representative of our target population than not.

If you want to use more than one stratification variable (which is known as *cross-stratification*), you would first organize your sampling frame according to the first stratification variable's strata (e.g., age brackets), and then within each of those strata you would organize the sample elements according to the second stratification variable's strata (e.g., income brackets).

Either way, when using stratified sampling, it is very important that the stratification variable(s) you select are meaningful for the variables you ultimately want to make an inference about.

For example...

If we wanted to use `income bracket` as our stratification variable, first we would need to know the income bracket of each element in our sampling frame. Then, we would organize our sampling frame according to those income brackets. Then, from the first income bracket, we would randomly select multiple elements to be included into our sample. Then we would move on to the next income bracket and randomly select multiple elements from that one to be included into our sample. We'd continue likewise until we've gone through each income bracket.
If we want the proportions of elements in our final sample per each stratum to match those of the target population's strata, then we're aiming to do *proportional stratified sampling*. However, if we want our sample's proportions to be something different than the population's proportions, then we're aiming to do *disproportional stratified sampling*.

Another good known-probability sampling method is called **cluster random sampling** (also known as *cluster sampling*, or *probabilistic cluster sampling*)—depicted in panel *B* of Figure 7.1. In that method, the elements in our sampling frame can be conceptually organized into mutually exclusive clusters, such that any given element exists in only one cluster. Depending on what the sample elements are (e.g., employees, meetings, or other events) the clusters can be based on anything (e.g., departments in a company, or days of the week if your sample elements are events). In **one-stage cluster sampling**, we randomly select a cluster and then we include *all* of its elements into our sample. We continue randomly selecting a new cluster and adding all of its elements into our sample, until we reach our desired sample size. Of course, some clusters won't end up into our sample. In **two-stage cluster sampling**, after we randomly select a number of clusters, we then randomly select a subset of elements within each cluster, rather than including the entire cluster into our sample.

A common scenario in which you might use cluster random sampling is when your sampling frame doesn't have detailed-enough information for you to randomly select individual sample elements, but you know your sampling frame can be organized into mutually exclusive clusters. However, the ability of cluster random sampling to produce a representative sample is decreased if the clusters have very different sizes.

For example...

Imagine we wish to sample 100 employees from ABC Company's 700 employees, but the company is struggling to create a list of every employee. Without such a list, we can't do simple-random sampling nor stratified random sampling, but we can do cluster random sampling.

If we know that each employee is officially part of only one of 12 departments at ABC Company, we can treat each department as a cluster. We would assign an integer ID (e.g., 1, 2, ..., 12) to each department, and then use a random integer generator to randomly select departments. For each department that is selected, we would include into our sample everyone from that department. Then we would generate a new random integer and add everyone from *that* department. We would continue until we've achieved our desired sample size. Of course, some departments won't end up into our sample.

As another example, if we know each employee has only one of 45 official job titles at ABC Company, we can treat each job title as a cluster. We would randomly select job titles (using our random integer generator as explained in the preceding example) and then we include into our sample everyone from the randomly selected job titles. We would repeat until we achieve our desired sample size. Of course, some job titles won't end up into our sample.

Lastly, keep in mind you can combine several of the known-probability sampling techniques. For example, you could use two-stage cluster sampling to randomly select several clusters, and then within each cluster you can use stratified sampling to randomly select elements within each of the strata in each cluster.

There are many other known-probability sampling techniques, but I've described here the most common ones you may encounter in IWO Psychology. If you are curious to dive deeper, consider reading Lohr (2021) and Chaudhuri and Stenger (2005).

#### 7.2.2.2 Unknown-probability sampling

When we don't have a sampling frame, we won't know the probability of selecting any given element of the target population, and we certainly can't ensure that each element will have an equal probability of being selected. In that case, our only option is to use any of the **unknown-probability sampling methods** (also known as **nonrandom sampling**, **nonprobability sampling**, or **nonprobabilistic sampling**).

As I described earlier, in IWO Psychology the most common reason why we wouldn't have a sampling frame is because our target population transcends any specific company (e.g., "*all nurses in the USA*" or "*all working adults in Western cultures*"), which makes it practically unfeasible to obtain a sampling frame. Whenever we don't have a sampling frame, we would need to use *unknown-probability sampling*.

In IWO Psychology, the most common type of unknown-probability sampling method is called **convenience sampling** (also known as *opportunistic sampling*). In that method, we simply include into our sample whomever/whatever population elements we have access to. Typically, we would specify a set of **eligibility criteria**, which is a set of attributes that we want our sample elements to have, and we would screen our potential sample elements to filter out any elements that fail to meet our eligibility criteria.

> For example...
>
> Our eligibility criteria could be as broad as "*any working adult*" or as narrow as "*any nurse who works full-time in any public hospital in the Southeastern USA*".

The usage of the word *screen* or *screening* to connote *filtering out* comes from the usage of mesh metal sieves (also known as *mesh screens*) that allow small objects to fall through the holes in the mesh, thus allowing the user to separate objects from a heterogenous mixture. For example, this is often used to separate seeds from juice (the juice flows through the mesh, whereas the seeds are trapped by the mesh because they are too big to flow out of the holes).

If we post an announcement online (e.g., on social media) to try to recruit participants, that is still convenience sampling because we are seeking eligible participants who happen to be available to us because they happen to see our announcement.

One type of convenience sampling is known as **quota sampling**. In that method, we would specify two or more categories of elements we wish to have into our sample, and then we specify how many elements (i.e., a *quota*) of each category we want to end up with into our sample. Then, we use convenience sampling until we obtain our desired number of sample elements for each category. Each category can have a different quota.

> For example...
>
> When using quota sampling, it's up to the research designer to figure out what the categories should be, and how many elements are desired for each category (they don't all have to be the same amount). For example, you might decide it's important for your study to include 30 workers from the healthcare industry, 30 workers from the tech industry, and 30 workers from education. Or you might decide you want to select a different number of workers in each of several age categories to match the relative proportions published in the U.S. Census.

Keep in mind an important difference between stratified sampling versus quota sampling: in stratified sampling, elements of the population are *randomly* selected from each stratum. In contrast, in quota sampling, the elements from each category are selected by convenience—which is technically not random.

A third type of convenience sampling that is common in IWO Psychology is known as **snowball sampling**. In that method, we would seek sample elements just like we would normally do in convenience sampling and/or quota sampling, but then we would use those first-round sample elements to help us find more sample elements (i.e., 2nd-round sample elements), and then we might use those 2nd-round sample elements to help us find even more sample elements. The pattern would continue until we reach our desired sample size.

> For example...
>
> The most traditional example of snowball sampling would be asking participants to convince their coworkers (or acquaintances, or friends, etc.) to also participate in the

The phrase *snowball sampling* comes from the fact that snowballs grow in size as they roll down a snowy hill. Likewise, the sample size grows as each sample element is used to discover even more sample elements.

research. However, snowball sampling can also be used when the sample elements aren't people, as long as the elements are connected to other elements in some way—but those situations are rare in IWO Psychology.

### 7.2.3 How big must your sample be?

The field of statistics is currently somewhat bifurcated into two separate approaches: frequentist statistics and Bayesian statistics. Frequentist statistics is classic, and most students are taught frequentist statistics by default. Today among statisticians, Bayesian statistics has arguably acquired equal status (possibly greater status), but most students aren't taught Bayesian statistics—because it's still relatively new and most professors were never taught it.

In frequentist statistics, there is a technique known as **power analysis** in which you can calculate the minimum sample size that would be required if you want to achieve a sufficiently high probability of detecting the effect you are testing (conditional on several assumptions). In Bayesian statistics, there are analogous concepts that are similar to frequentist power analysis. Both of those techniques are relatively straightforward, but they are beyond the scope of this general research methods course. You can learn about frequentist and/or Bayesian techniques for estimating required sample sizes via any good internet resource about statistics. For power analysis, you might consider reading Muthén and Muthén (2002) as a practical introduction.

## 7.3 A reminder about pre-registration…

! Consider pre-registering your plans…

In quantitative research, IWO Psychologists carefully plan out the details of any study before they start collecting empirical data or committing to a "point of no return". During that planning phase, the IWO Psychologist would

specify details such as who their target population is, what the theory and hypotheses are, what their conceptual definitions and operational definitions are, how they're going to measure their cónstructs, how they're going to sample, how much sample they need, etc., etc.
It is often a great idea to pre-register those plans as soon as possible, so that you have publicly verifiable time-stamped evidence that you didn't secretly HARK (see Section 2.4).

## 7.4 Suggested Readings

Zickar's (2020) article is an easy-to-read yet deeply insightful overview of how to create a valid psychometric questionnaire.

You will take a Statistics course in this degree program, but if you are eager to learn more about statistics before taking that course, here are some books I can personally recommend:

- Wackerly et al.'s (2008) book is gentle but in-depth.
- Wasserman's (2003) book is popular and concise.
- Shao's (2003) book goes into rigorous detail, but it assumes the reader has quite a bit of background knowledge about statistics and other mathematics.

# 8 Statistics

## 8.1 Introduction

The contents of this chapter are being updated for the Fall 2025 semester. Check back soon.

## 8.2 Suggested Readings

# 9 Causation: Randomized Controlled Experiments & Quasi-Experiments

## 9.1 Introduction

The contents of this chapter are being updated for the Fall 2025 semester. Check back soon.

## 9.2 Suggested Readings

# 10 Causation: The basics of DAGs about causation

## 10.1 Introduction

The contents of this chapter are being updated for the Fall 2025 semester. Check back soon.

## 10.2 Suggested Readings

# 11 Causation: Backdoor & Frontdoor criteria in DAGs

## 11.1 Introduction

The contents of this chapter are being updated for the Fall 2025 semester. Check back soon.

## 11.2 Suggested Readings

# 12 Disseminating the Results of Research

## 12.1 Introduction

The contents of this chapter are being updated for the Fall 2025 semester. Check back soon.

## 12.2 Suggested Readings

# References


Adams, C., & van Manen, M. (2008). Phenomenology. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (1–2, pp. 614–619). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909.n317

American Psychological Association. (2017). *Ethical Principles of Psychologists and Code of Conduct (2002, Amended 2010, 2016)*. American Psychological Association. https://www.apa.org/ethics/code/ethics-code-2017.pdf

Blatter, J. K. (2008). Case Study. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (1–2, pp. 68–71). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909

Cascio, W. F., & Aguinis, H. (2008). Research in industrial and organizational psychology from 1963 to 2007: Changes, choices, and trends. *Journal of Applied Psychology*, *93*(5), 1062–1081. https://doi.org/10.1037/0021-9010.93.5.1062

Charmaz, K., & Bryant, A. (2008). Grounded Theory. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (1–2, pp. 374–377). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909

Chaudhuri, A., & Stenger, H. (2005). *Survey Sampling: Theory and Methods* (2nd ed.). Chapman & Hall/CRC. https://doi.org/10.1201/9781420028638

Clandinin, D. J., & Caine, V. (2008). Narrative Inquiry. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (1–2, pp. 541–544). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909

Craig, R. T. (2009). Metatheory. In S. W. Littlejohn & K. A. Foss (Eds.), *Encyclopedia of Communication Theory* (pp. 657–661). Sage.

Edmondson, A. C., & McManus, S. E. (2007). Methodological Fit in Management Field Research. *The Academy of Management Review*, *32*(4), 1155–1179. https://www.jstor.org/

stable/20159361
Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, *12*(1), 1–22. https://doi.org/10.1037/1082-989X.12.1.1
Fetterman, D. M. (2008). Ethnography. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (1–2, pp. 288–292). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909.n150
Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (Third edition). The Guilford Press. https://www.guilford.com/books/Introduction-to-Mediation-Moderation-and-Conditional-Process-Analysis/Andrew-Hayes/9781462549030
Hollenbeck, J. R., & Wright, P. M. (2017). Harking, Sharking, and Tharking: Making the Case for Post Hoc Analysis of Scientific Data. *Journal of Management*, *43*(1), 5–18. https://doi.org/10.1177/0149206316679487
Jaccard, J. J., & Jacoby, J. (2020). *Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists* (Second edition). The Guilford Press. https://www.guilford.com/books/Theory-Construction-and-Model-Building-Skills/Jaccard-Jacoby/9781462542437
Katzell, R. A., & Austin, J. T. (1992). From then to now: The development of industrial-organizational psychology in the United States. *Journal of Applied Psychology*, *77*(6), 803–835. https://doi.org/10.1037/0021-9010.77.6.803
Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, *2*(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4
Legood, A., van der Werff, L., Lee, A., den Hartog, D., & van Knippenberg, D. (2023). A Critical Review of the Conceptualization, Operationalization, and Empirical Literature on Cognition-Based and Affect-Based Trust. *Journal of Management Studies*, *60*(2), 495–537. https://doi.org/10.1111/joms.12811
Lohr, S. L. (2021). *Sampling: Design and Analysis* (3rd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429298899
McGuire, W. J. (1997). Creative hypothesis generating in psy-

chology: Some useful heuristics. *Annual Review of Psychology*, *48*(1), 1–30. https://doi.org/10.1146/annurev.psych.48.1.1

Morling, B. (2021). *Research Methods in Psychology: Evaluating a World of Information* (Fourth edition). W. W. Norton & Company. https://wwnorton.com/books/9780393536263/

Murphy, K. R., & Aguinis, H. (2019). HARKing: How Badly Can Cherry-Picking and Question Trolling Produce Bias in Published Results? *Journal of Business and Psychology*, *34*(1), 1–17. https://doi.org/10.1007/s10869-017-9524-7

Muthén, L. K., & Muthén, B. O. (2002). How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. U.S. Department of Health, Education, and Welfare.

Newman, A., Donohue, R., & Eva, N. (2017). Psychological safety: A systematic review of the literature. *Human Resource Management Review*, *27*(3), 521–535. https://doi.org/10.1016/j.hrmr.2017.01.001

Oxford English Dictionary. (2024). *Soundness, n., sense 3.a.* Oxford University Press. https://doi.org/10.1093/OED/4396907815

Oxford English Dictionary. (2025a). *'Of…validity' in validity, n., sense 2.a.* Oxford University Press. https://doi.org/10.1093/OED/9947174010

Oxford English Dictionary. (2025b). *Sound, adj., sense II.8.a.* Oxford University Press. https://doi.org/10.1093/OED/1626491124

Porter, L. W., & Schneider, B. (2014). What Was, What Is, and What May Be in OP/OB. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 1–21. https://doi.org/10.1146/annurev-orgpsych-031413-091302

Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, *65*(1), 1–12. https://doi.org/10.1037/a0018326

Salas, E., Kozlowski, S. W. J., & Chen, G. (2017). A century of progress in industrial and organizational psychology: Discoveries and the next century. *Journal of Applied Psychology*, *102*(3), 589–598. https://doi.org/10.1037/apl0000206

Shao, J. (2003). *Mathematical Statistics* (2nd edition). Springer.

Society for Industrial and Organizational Psychology, Inc. (2016). *Guidelines for Education and Training in Industrial-Organizational Psychology*.

Somekh, B. (2008). Action Research. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (1–2, pp. 4–7). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909.n4

Wackerly, D., Mendenhall, W., & Scheaffer, R. L. (2008). *Mathematical Statistics with Applications* (7th edition). Thomson Brooks/Cole.

Wasserman, L. (2003). *All of Statistics: A Concise Course in Statistical Inference* (First Edition). Springer.

Zickar, M. J. (2020). Measurement Development and Evaluation. *Annual Review of Organizational Psychology and Organizational Behavior*, *7*(1), 213–232. https://doi.org/10.1146/annurev-orgpsych-012119-044957