

Research Design



The importance of appropriate research methods and designs cannot be overstated. Research methods and designs help us achieve an accurate understanding of various phenomena and ensure conclusions are justified.

Research Questions

Research questions are fundamental to all research projects. Research questions help focus the study, determine the appropriate methodology, and guide each stage of inquiry, analysis, and reporting. Some examples of research questions germane to people analytics include:

- Q_1 : Why has there been an increase in attrition over the past quarter?
- Q_2 : How equitable are promotion nominations across the organization?
- Q_3 : Are there meaningful differences in the favorability of experiences for remote vs. non-remote employees?
- Q_4 : Do new joiners have the training and resources they need to be successful?
- Q_5 : What portion of team performance is attributable to leadership effectiveness?

Research Hypotheses

Research hypotheses are testable statements about the expected outcome of a research project or experiment.

- H_1 : Manager satisfaction is a significant predictor of voluntary attrition.
- H_2 : Promotion nomination rates are not significantly different by gender and ethnicity.

- H_3 : Employee experience favorability is not significantly different between remote and non-remote workers.
- H_4 : New hire training perceptions are positively associated with onboarding experience favorability.
- H_5 : Leadership effectiveness perceptions explain significant variation in team performance.

Internal and External Validity

Internal validity refers to the extent to which confounding variables are controlled. In other words, internal validity reflects the *robustness* of the study.

For example, if a study finds a significant relationship between work location and attrition but considers no other factors or explanations, this would *not* be a robust study. Work location may emerge significant because certain roles for which attrition is higher are more concentrated in one or more geographies. It could also be the case that the company has made acquisitions in new geographies, and the acquired employees have significantly different experiences (and attrition rates) relative to non-acquired employees.

Confounding variables are critically important in the context of internal validity. A confounding variable is an extraneous variable whose presence impacts the variables being studied such that results do not reflect the actual relationships. Studies with weak internal validity often result in spurious associations that *confound* the true relationship between two variables, leading to invalid conclusions and recommendations.

External validity refers to the extent to which study conclusions will hold in other contexts (for other people, in other places, at other times). *Randomization* is fundamental to our ability to generalize and apply findings to other groups or contexts.

If we survey employees to understand sentiments about recent changes in business strategy but exclude groups for which there may be different impacts or perceptions, conclusions about the collective sentiment would be suspect at best.

Research Methods

There are three major categories of research methods: (1) *quantitative*, (2) *qualitative*, and (3) *mixed methods*.

1. Quantitative

- Addresses *what* questions
- Utilizes numerical data (e.g., surveys, systems)
- Primarily deductive

- Used to test hypotheses
- Involves statistical analyses
- More objective
- More generalizable

2. Qualitative

- Addresses *how* and *why* questions
- Utilizes text data (e.g., focus groups, interviews, open-ended feedback)
- Primarily inductive
- Used to formulate theory or hypotheses
- Involves organizing data into categories or themes
- More subjective
- Less generalizable

3. Mixed Methods

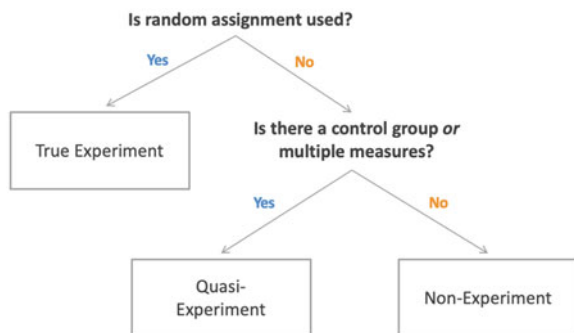
- Integrates the strengths of both quantitative and qualitative methods within a single study, often leading with qualitative approaches to build theory and hypotheses followed by quantitative methods to test hypotheses

Research Designs

In addition to determining whether a quantitative, qualitative, or mixed methods study is most appropriate, researchers also need to decide on the type of study within each of these three. **Research designs** are the types of inquiry within quantitative, qualitative, and mixed methods approaches that issue specific direction for the research procedures (Creswell & Creswell, 2018). There are multiple taxonomies for research designs, and we will simplify to the most common types.

Within the quantitative category, there are three types of designs: (a) *experimental*, (b) *quasi-experimental*, and (c) *non-experimental*. As shown in Fig. 1, it is important to understand the centrality of randomization in this decision.

Fig. 1 Quantitative research designs



Experimental Research

Experimental research is concerned with casual (internal) validity. Randomized experimental designs provide the most rigor with regard to causal validity. However, in social science research contexts, true experiments often are not possible due to ethical considerations.

For example, if we were interested in understanding the causal effect leadership quality has on employee engagement, based on a hypothesis that poor leadership decreases employee engagement, we would need to randomly assign employees to one of two groups that are identical on the basis of all variables that could theoretically explain why employees vary in their levels of engagement. Then, we would need to manipulate the variable of interest (leadership quality) to evaluate if the group of employees subjected to poor leadership (treatment group) reports significantly different levels of engagement relative to the group of employees for whom leadership quality has not been manipulated (control group). In a practical setting, it would of course be unethical to purposefully subject employees to poor leadership with the expectation of reducing engagement—and consequently, productivity, retention, and impact to the organization.

Clinical trials are a common setting for true experiments, as isolating the effects of an experimental drug can be a matter of life or death. In a randomized clinical trial, patients are randomly assigned to an experimental group (patients who receive the drug) or control group (patients who receive a placebo). To protect against placebo effects biasing the results, patients do not know if they receive the experimental treatment or the placebo. Done correctly, these experiments have the highest level of internal validity.

Another example of an experimental design is **A/B testing**. A/B testing is often performed in the context of website optimization, in which two or more versions of the site are shown to customers to identify which version impacts key success metrics more positively. In a people analytics context, we may create two versions of a dashboard and *randomly* assign the permissioned users to each. We could then assess whether utilization rates, average usage time, repeat usage, among other success measures are significantly different between the two groups of users to inform which design is most effective.

In experimental research, it is important to consider the potential influence of the **Hawthorne Effect**, which refers to the tendency of some individuals to modify their behavior in response to the awareness that they are being observed. This term was coined during experiments at Western Electric’s factory in the Hawthorne suburb of Chicago in the late 1920s and early 1930s. One of many studies conducted to understand how work environments effect productivity was known as the “Illumination Experiment.” During this study, researchers experimented with a number of lighting levels in a warehouse in which workers made electrical relays. The researchers found that any change in the lighting—even when introducing poor lighting—led to favorable changes in output. However, these productivity gains disappeared once the attention faded (Roethlisberger and Dickson, 1939).

In a people analytics context, if we inform employees that we are going to monitor productivity over a period of time, it is likely that at least some employees will attempt to modify their behavior in order to increase productivity levels. After all, higher productivity is generally regarded as an ideal across companies and industries. In this case, manipulating some aspect of the typical work context to study a treatment effect, such as flexible work arrangements, may be impacted by this phenomenon; that is, observed differences in productivity may not be attributable to flexible work arrangements but merely due to employees knowing they are being observed.

Quasi-Experimental Research

Quasi-experimental research is an experiment in which participants cannot be randomly assigned.

In the case of our leadership quality example, a quasi-experiment may examine engagement differences between two groups of employees who rate their leader either favorably (Group A) or unfavorably (Group B). A key limitation of this approach is that the groups may be different in important ways beyond leader perception incongruities. For example, Group A employees may be concentrated within a single department, whereas Group B employees may span all other departments. This would indicate that the difference in leadership—and presumably engagement—is driven by factors unique to the department, making it more challenging to isolate the effects of leadership quality on engagement. Perhaps the department with unfavorable leader perceptions has seen significant attrition, or the department is largely first-time people leaders in need of coaching and support.

Another example of quasi-experiments is a pretest-posttest setting in which there are multiple measures. Random assignment could be used in pretest-posttest contexts, in which case this would be characterized as a true experiment, but often this approach is implemented without random assignment. For example, we could test the hypothesized effect of leadership quality on engagement via a pretest-posttest approach. If leaders are selected for a leadership development workshop, we could survey the leaders' teams and collect data on leader effectiveness perceptions and self-reported engagement prior to (baseline) and after the workshop. It is unlikely that leaders were selected for this workshop by a random process; there were likely criteria driving the selection, such as leaders who were identified as critical talent or who achieved a certain performance level. If this study finds that improvements in leadership effectiveness correlate with improvements in engagement, there would be *some* evidence favoring investments in leadership development; however, this would not be sufficient evidence for a causal effect.

Though quasi-experiments are not as robust as true experiments, they are usually more feasible in a people analytics context. True experiments control for confounding variables by way of the research design (randomization ensures equivalent groups), while these factors must be controlled statistically in quasi-

experimental contexts. In chapter “[Linear Regression](#)”, we will discuss how to model relationships among multiple variables in order to study how one variable influences another while holding constant variables that may influence the outcome but are not the primary focus of the research.

Non-Experimental Research

Unlike experimental and quasi-experimental designs, **non-experimental research** does not involve the manipulation of a variable. The goal of non-experiments is not to provide evidence for causal effects, but to study measured variables as they naturally occur and disentangle patterns in the data.

Given the potential for alternative explanations of any observed differences or relationships, non-experimental research tends to have lower internal validity than experimental and quasi-experimental designs. As we have discussed, it is often not possible or ethical to manipulate aspects of a person’s work context or to randomly assign people to groups. In addition, the nature of research questions does not always warrant experiments. In these cases, one of three non-experimental approaches may be considered: (a) *cross-sectional*, (b) *correlational*, and (c) *observational*.

Cross-sectional research compares two or more natural groups of people. For example, we may examine differences in engagement between employees in the Engineering department relative to employees in the Product department. In this case, we would neither manipulate one’s department to determine how department influences engagement nor randomly assign employees to these departments. Department membership exists apart from the research, so these naturally occurring groups can be leveraged for comparisons. There are of course many examples of naturally occurring groups that we would not manipulate, such as gender, ethnicity, generation, education, job family, job level, location, and tenure band. When participant characteristics are used to create groups, these variables are sometimes referred to as *experimenter-selected*—rather than *experimenter-manipulated*.

Correlational research involves studying the statistical relationship between two variables without manipulating some aspect of a person’s natural context. The relationship between leadership quality and engagement could be evaluated using correlational research. However, we would be unable to leverage a correlational design to test a hypothesis positing a causal effect of leadership quality on engagement. We would be limited to understanding how leadership quality and engagement covary; that is, to what extent a change in one variable is associated with a change in the other. Engagement may tend to increase as leadership quality increases, but a correlational design does not lend to understanding the direction of causal influence—if such an effect exists.

Observational research refers to studies in which the researcher gathers information without research subjects being explicitly involved in the recording of data. Collecting data from the company’s Human Capital Management (HCM) system could be an observational research method. For example, if we access data on

terminations to determine the rate of attrition over a specified period, we would not need to interfere by asking past or present employees for this information. We would also do so without manipulating any aspect of the ordinary environment, tagging people to naturally occurring or artificially created groups, or evaluating the association of attrition with another variable. The reality is that such an approach would not be too actionable, however, as this would offer no understanding of what may be influencing attrition.

Review Questions

1. What type of research method and design would be best suited for a study aiming to understand the effect of stay interviews on employee attrition?
2. Why are quasi-experiments less rigorous than true experiments?
3. When evaluating the effectiveness of a new program, what are some reasons an experimental design would not be implemented?
4. What is the role of research questions?
5. What is the role of research hypotheses?
6. What is the difference between internal and external validity, and why are these concepts important in research?
7. What is an example of a mixed methods study?
8. What is the key difference between experimental and non-experimental research designs?
9. What are the differences between cross-sectional, correlational, and observational non-experimental designs?
10. How can the Hawthorne Effect impact the integrity of an experiment?

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

