

Survey Design and Data Analysis

Survey Design and Data Analysis

A Practical Guide To Crafting Survey
Items and Analysing Quantitative Data

SEUNG YOUN (YONNIE) CHYUNG, ED.D.



Survey Design and Data Analysis Copyright © 2025 by Seung Youn (Yonnie) Chyung, Ed.D. is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/), except where otherwise noted.

Contents

Preface	viii
Acknowledgements	xii
<u>1. Plan to Conduct Self-Administered Surveys</u>	
1.0 Introduction	2
1.1 Various data collection methods	3
1.2 Purposes of conducting surveys in organizations	9
1.3 Surveys conducted with a population or a sample	14
1.4 Sampling methods	18
1.5 Chapter activity: How does your organization use surveys?	25
<u>2. Understand the Basic Structure of Survey Items</u>	
2.0 Introduction	28
2.1 Survey anatomy	30
2.2 Survey items capturing facts or measuring perceived quality/value	32
2.3 Survey items measuring constructs	36
2.4 Survey items with different levels of measurement scales	40
2.5 Chapter activity: What do they measure, facts or perceived quality/value?	45
<u>3. Explore Various Rating Scales</u>	
3.0 Introduction	49
3.1 Discrete vs. continuous rating scales	51
3.2 Numerical rating scales	53
3.3 Verbal descriptor scales including the Likert scale	56

3.4 Graphic rating scales and Web-based sliders	58
3.5 Chapter activity: When to use and not to use specific rating scales?	60

4. Design Surveys Using Evidence-Based Principles

4.0 Introduction	62
4.1 Survey questions or statements written in positive or negative wording	64
4.2 Rating scales with or without a midpoint	71
4.3 Ascending vs. descending order of response options	79
4.4 Chapter activity: How would you revise the current survey questionnaire?	83

5. Develop Surveys for Different Purposes

5.0 Introduction	87
5.1 Purpose of the survey and knowledge to be gained from the survey	88
5.2 Sample surveys used during evaluation	95
5.3 Sample surveys used during performance analysis	111
5.4 Sample surveys used during instructional design	123
5.5 Chapter activity: How would you design your own survey questionnaire?	127

6. Administer Surveys on Paper or Online

6.0 Introduction	132
6.1 Paper-based survey questionnaires	134
6.2 Online surveys with Google Forms	136
6.3 Online surveys with Qualtrics	141
6.4 Online surveys with SurveyMonkey	147
6.5 Chapter activity: How do different online survey systems work?	151

7. Prepare Survey Data for Analysis and Visualize Data

7.0 Introduction	154
7.1 Data Screening	156

7.2 Chapter activity: How would you do your data screening?	165
7.3 Visualizing data and results	167
7.4 Chapter activity: How would you visually present the data?	173
Appendix A	174
About the Author	175
Feedback and Suggestions	176

Preface

Audience for this book

This book is intended for professionals (or aspiring professionals) in Learning and Development (L&D), Human Resources Development (HRD), Organizational Development (OD), and Performance Improvement (PI) who want to learn how to design self-administered survey questionnaires using closed-ended questions and response scales, administer those questionnaires, and analyze the resulting quantitative data using Excel.

If you are one of these professionals, begin by assessing your current knowledge and skills in survey design and data analysis. Consider where you fall within the following five levels—novice, advanced beginner, competent, proficient, or expert. If you identify yourself as a novice or advanced beginner, this book will serve as a practical guide to help you build foundational knowledge and skills. If you are at the competent, proficient, or expert level, you may find this book valuable as a reference or job aid for specific aspects of survey design and data analysis.

Approaches used in this book

Knowledge should be drawn from credible and reliable sources, such as empirical research and the insights of subject matter experts. This book uses this *evidence-based approach*, drawing on both research findings and expert knowledge to teach you essential concepts and principles for designing survey questionnaires.

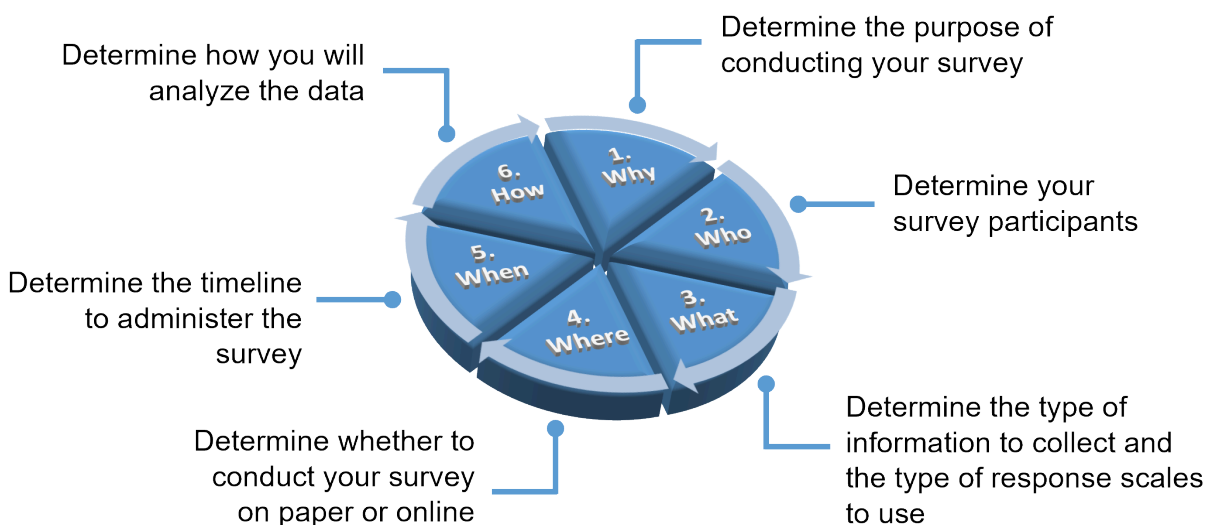
Equally important to teaching strategies are learning strategies. One of the most effective ways to gain new skills is through hands-on practice. This book also uses this *learning-by-doing approach*, providing numerous examples and practical exercises. These activities are designed to help you get up to speed with the knowledge and skills needed to confidently design self-administered survey questionnaires in just a few weeks.

Content of this book

Throughout this book, the terms *surveys*, *survey items*, and *data* specifically refer to *self-administered surveys*, *closed-ended survey items*, and *quantitative data*, respectively, unless otherwise noted.

This book is designed to help you understand the overall process of survey design and data analysis using the 5W1H method. You will follow the five Ws and one H steps in the following sequence—*Why*, *Who*, *What*, *Where*, *When*, and *How*—as illustrated in Figure P-1:

Figure P-1 The 5Ws and 1H Method Used in Survey Design and Data Analysis Process



The seven chapters of this book will walk you through a step-by-step procedure for designing surveys and analyzing survey data:

1. Plan to conduct surveys (*Why* and *Who*)
2. Understand the basic structure of survey items (*What*)
3. Explore various rating scales (*What*)
4. Design surveys using evidence-based principles (*What*)
5. Develop surveys for different purposes (*What*)
6. Administer surveys on paper or online (*Where* and *When*)

7. Prepare survey data for analysis and visualize data (*How*)

In [Chapter 1](#), you will begin with key principles and initial steps for planning a survey. You will learn to recognize surveys as one of several data collection methods—each having advantages and disadvantages. You will also explore the types of projects where surveys are appropriate and determine whether to collect data from an entire population or a sample of that population.

In [Chapter 2](#) and [Chapter 3](#), the focus shifts to designing closed-ended survey items with suitable response scales. You will learn how to generate the type of quantitative data needed for analysis. Additionally, you will explore different survey designs to capture factual data and measure constructs such as perceived quality and value.

In [Chapter 4](#) and [Chapter 5](#), you will dive deeper into evidence-based principles for crafting effective survey items and response scales. You will also explore different types of survey questionnaires tailored for specific projects such as evaluation, performance analysis, and instructional design.

In [Chapter 6](#), you will explore various online survey tools for administering the surveys you have designed. You will compare the features, strengths, and limitations of several popular online survey tools such as Google Forms, Qualtrics, and SurveyMonkey, to help you select the tools that best fits your needs.

In [Chapter 7](#), you will learn the importance of preparing raw data through different data screening techniques. You will also learn how to summarize and visually present your data using charts in Excel. Using hypothetical datasets provided in the chapter, you will practice data screening and chart creation to reinforce these skills.

[Appendix A](#) offers additional guidance on analyzing survey data using Excel's Data Analysis Toolpak. With hypothetical survey scenarios and datasets, you will learn how to perform basic statistical tests, including correlation, paired samples t-test, and independent samples t-test with Excel.

Throughout the book, you will find numerous examples that illustrate survey design principles, along with sample survey items you can adapt for your own use. Each chapter concludes with an activity designed to help you apply what you've learned and build confidence in survey design and data analysis.

Use of this book

After studying and practicing with this book for 50+ hours, you should feel reasonably confident in using surveys in a variety of organizational projects, including evaluation, performance analysis, and instructional design.

If you are an instructor at a post-secondary institution, you may use this book as a primary textbook for a short-term course on survey design and data analysis, or as a supplemental resource for a semester-long course involving survey-based data collection.

If you are a practitioner, this book serves as a self-guided resource to assist your survey design and data analysis efforts in real-world projects.

For those learning to conduct evaluations, this book serves as a valuable companion to my other book, [*10-Step Evaluation for Training and Performance Improvement \(2019\)*](#), published by Sage.

Ultimately, I hope this book, *Survey Design and Data Analysis*, equips both practitioners and emerging researchers with fundamental knowledge and skills needed for effective survey design and basic data analysis, laying the groundwork for more advanced topics such as scale development and statistical data analysis.

Seung Youn (Yonnie) Chyung, Ed.D., Professor
Organizational Performance and Workplace Learning
Boise State University
Boise, Idaho
USA

Acknowledgements

First, I am deeply grateful to my colleague, **Jennifer Shamsy, M.S.**, for many reasons.

As part of my ongoing research, I lead the [Workplace-Oriented Research Central \(WORC\) Lab](#) at Boise State University, where teams of students and I engage in collaborative research projects. Between 2017 and 2019, Jennifer, even after she graduated from the OPWL program, participated in my research lab and was a key contributor to my *evidence-based survey design* research efforts, which resulted in a series of journal articles and conference presentations. That experience also helped lay the groundwork for the development of a new course, **OPWL 506 Survey Design and Data Analysis**.

While developing the course, I wrote this book to serve as the primary textbook for that course. During its development, Jennifer continued to assist me as a research assistant and edited the entire manuscript, greatly enhancing its clarity and readability.

After I taught the course a few times, Jennifer took over and began teaching it herself. Her hands-on approach and thoughtful feedback to students greatly enriched the course experience.

Over the years, Jennifer continued to contribute her insights based on her teaching experience, offering valuable suggestions that led to further enhancements to the book.

Her contributions have been truly invaluable in bringing this book to life.

Thank you, Jennifer!

My sincere appreciation also goes to the following colleagues at Boise State's eCampus Center, who developed the Open Educational Development Series (OEDS) and guided me through the process of converting my text into this Open Educational Resource (OER). Their support ensured copyright and accessibility compliance, as well as a professionally designed front page for the book.

- **Sarah Saia**, OEDS Coordinator, eCampus Center
- **Steph Hamilton**, Copyright and Accessibility Specialist, eCampus Center
- **Arlie Sommer and Trevor Warr**, Book Cover Design, eCampus Center
- **Amy Vecchione**, Assistant Director of eCampus Research and Innovation, eCampus Center

This OER format allows me to update content more easily, share knowledge more widely, and invite collaboration from a broader community of practitioners and educators. **Thank you, Sarah, Steph, Arlie, and Amy!**

1. PLAN TO CONDUCT SELF-ADMINISTERED SURVEYS

1.0 Introduction

As a professional in Learning and Development (L&D), Human Resources Development (HRD), Organizational Development (OD), or Performance Improvement (PI), you often need to make data-driven decisions in areas such as performance analysis (also known as needs assessment), evaluation, change management, and instructional design. To support these decisions, it is essential to collect data from various sources. One effective method for doing so is through self-administered survey questionnaires.

Learning Objectives

After completing this chapter, you will be able to:

- Recognize self-administered surveys as one of several data collection methods
- Identify various types of projects where you can use self-administered surveys to collect data
- Explain the advantages and disadvantages of using self-administered surveys
- Determine whether to administer your survey to an entire population or a sample
- Select appropriate sampling methods for survey-based projects

1.1 Various data collection methods

When collecting data, you take on the role of a researcher. As a researcher, you can use a variety of data collection methods such as the following:

- **Self-administered survey:** Individual informants provide information to you by completing your survey questionnaire on their own. The survey questionnaire may consist of closed-ended questions, open-ended questions, or a combination of both.
- **One-on-one interview:** You ask each informant a series of questions to gather in-depth information. These interviews typically include open-ended questions, but may also incorporate some closed-ended questions.
- **Focus group:** You facilitate a discussion with a group of 10-12 informants from diverse backgrounds to learn about their experiences and opinions.
- **Observation:** You collect data by using your senses (watching, listening, talking, feeling, and sometimes interacting) in an environment, either as a participant or non-participant.
- **Extant data review:** You review and analyze existing data and documents to extract relevant information.
- **Test:** You ask informants to complete written or performance tests to demonstrate their knowledge or skills.

No single data collection method is ideal for all types of situations. Each has its own strengths and limitations (see Table 1). Therefore, it is important to select the method, or a combination of methods, that best fits your project's goals and context.

Unlike interviews, where you ask questions directly, self-administered surveys (hereafter referred to as *surveys*) are completed by the respondents themselves. For this reason, your survey questionnaire must be designed clearly so respondents can easily understand the purpose of the survey, the meaning of each item, and how to respond appropriately.

Surveys can be administered on paper or online. Paper surveys may be collected on-

site or returned by mail. Additionally, survey questionnaires can be integrated into interviews or focus groups as supplementary tools.

This book focuses on conducting surveys within organizational settings. It is specifically written for professionals in Learning and Development (L&D), Human Resource Development (HRD), Organizational Development (OD), and Performance Improvement (PI)—those who act as survey researchers, collecting data from both internal and external stakeholders to inform evidence-based decisions that support performance improvement.

Table 1 *Strengths and Limitations of Various Data Collection Methods* (adapted from Table 36, pp. 116-117 in Chyung, 2019¹)

Self-administered survey

Method	Strengths	Limitations
On site	<ul style="list-style-type: none"> • Can administer it to a group of people at once • Can handle participants' questions in person immediately 	<ul style="list-style-type: none"> • Costly to print out copies of the questionnaire • Necessary to use a survey administrator for each site • Time-consuming to enter data into an electronic file
Via mail	<ul style="list-style-type: none"> • Can administer it to a large number of people in different locations • Participants can complete it during their preferred time 	<ul style="list-style-type: none"> • Costly to print out and mail copies of the questionnaire • Notoriously low survey return rates • Time-consuming to enter data into an electronic file
On line	<ul style="list-style-type: none"> • Can administer it to a large number of people in different locations • Participants can complete it during their preferred time • Data are already stored in an electronic file, ready for statistical analysis 	<ul style="list-style-type: none"> • Notoriously low survey return rates • Survey administrators and participants must have access to technology

One-on-one interview

Method	Strengths	Limitations
In person	<ul style="list-style-type: none"> • Can observe interviewees' body language • Can retrieve detailed information* <p>(* common to all types of one-on-one interviews)</p>	<ul style="list-style-type: none"> • Interviewers and interviewees need to be in the same location • Costly and time-consuming to complete multiple interviews* <p>(* common to all types of one-on-one interviews)</p>
By telephone	<ul style="list-style-type: none"> • Interviewers and interviewees do not need to be in the same location • Most people have access to telephone 	<ul style="list-style-type: none"> • Cannot observe interviewees' body language • Interviewers and interviewees must have access to telephone
Via video-conferencing	<ul style="list-style-type: none"> • Interviewers and interviewees do not need to be in the same location • Can observe interviewees' body language 	<ul style="list-style-type: none"> • Interviewers and interviewees must have access to video-conferencing technology

Focus group

Method	Strengths	Limitations
In person	<ul style="list-style-type: none"> • Participants' body language provides additional data • Can retrieve information from a group of people at once* • Participants can learn other people's views and generate group opinions on a chosen topic through discussions* <p>(* common to all types of focus groups)</p>	<ul style="list-style-type: none"> • Discussions can be influenced by the physical environment (e.g., room setup, sitting arrangement) • All participants must be in the same location at the same time • Time-consuming to monitor and record participants' responses* <p>(* common to all types of focus groups)</p>
By telephone-conferencing	<ul style="list-style-type: none"> • Participants do not need to be in the same location 	<ul style="list-style-type: none"> • Cannot observe participants' body language
Via video-conferencing	<ul style="list-style-type: none"> • Participants do not need to be in the same location 	<ul style="list-style-type: none"> • Difficult to observe participants' body language or reaction to other participants' responses • Participants must have access to technology

Observation

Method	Strengths	Limitations
On site (participant observation)	<ul style="list-style-type: none"> • Can understand phenomena in natural settings • Can retrieve additional information for better contextual understanding when observers interact with participants • Can be used to verify self-reported data 	<ul style="list-style-type: none"> • Participants may behave differently than usual due to the presence of observers (a.k.a. the Hawthorne effect) • Time-consuming for the observers to become accepted as part of the natural settings • Costly to use a large number of spot observations through random sampling
Video-recorded (non-participant observation)	<ul style="list-style-type: none"> • Can obtain information unobtrusively as observers are not present at the setting • No/little interruption to the participants and setting • Can be used to verify self-reported data • Can be replayed for careful analysis 	<ul style="list-style-type: none"> • Limiting to rely on recorded information • Lacking opportunity to interact with participants to obtain contextual information • Costly to use a large number of spot observations through random sampling

Extant data review

Method	Strengths	Limitations
In print	<ul style="list-style-type: none"> • Can reveal actual historical information* • Can be used to verify self-reported data* <p>(* common to all types of extant data review)</p>	<ul style="list-style-type: none"> • Not easy to share collected data with other members of the evaluation team (compared to electronic files) • Time-consuming to enter data into an electronic file when needed
In electronic files	<ul style="list-style-type: none"> • Can share and access easily • Can readily conduct statistical analysis on some quantitative data 	<ul style="list-style-type: none"> • Time-consuming to analyze large amounts of information • Difficult to detect errors in information without comparing it to other sources* <p>(* common to all types of extant data review)</p>

Test

Method	Strengths	Limitations
Written (on site or online)	<ul style="list-style-type: none"> • Can reveal actual knowledge • Can be used to verify self-assessed knowledge or confidence levels 	<ul style="list-style-type: none"> • Difficult to measure knowledge • Unreliable results potentially caused by participants' test anxiety
Performance	<ul style="list-style-type: none"> • See Observation 	<ul style="list-style-type: none"> • See Observation

Notes

1. Chyung, S. Y. (2019). *10-step evaluation for training and performance improvement*. Sage.

1.2 Purposes of conducting surveys in organizations



To survey in the physical world means to examine and record the condition or characteristics of something. For example, land surveyors use specialized tools to accurately measure and document the location of property lines.



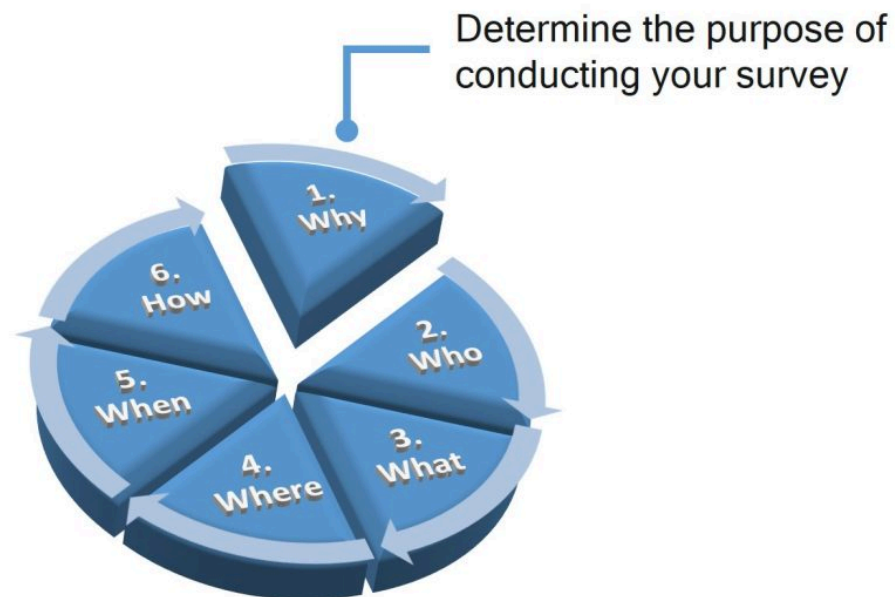
In contrast, **to survey in the social science fields** involve collecting information about people’s experiences, opinions, attitudes, or motivation behind their behaviors, most of which are not directly observable. For instance, to understand people’s satisfaction with a product or their level of support for a political party, you would develop a survey questionnaire and ask individuals to respond to a series of survey items. Just as land surveyors’ physical tools must be precise and reliable, survey questionnaires used in the social sciences must be carefully designed to accurately measure and record human thoughts and behaviors.

Before you begin designing your own survey questionnaire, it is important to understand the overall process of survey design and data analysis. This process can be effectively organized using the 5W1H method—a framework familiar to journalists¹), but equally useful in research. For survey design and data analysis, we will follow the 5W1H steps in the following sequence (also see Figure 1):

- **Why:** Determine the purpose of conducting your survey
- **Who:** Determine your survey participants
- **What:** Determine the design of the survey questionnaire including the type of information you want to collect and the type of response scales you want to use
- **Where:** Determine whether to conduct your survey on paper or online
- **When:** Determine the timeframe in which you administer the survey and complete data collection
- **How:** Determine how you will analyze the data

The first step in the survey design and data analysis process is to address **the Why** (see Figure 1)—Why are you conducting this survey? What is the specific purpose or goal behind your data collection effort?

Figure 1 *The Why Step in the Survey Design and Data Collection Process*



Organizations often conduct surveys as part of social science research to gather valuable insights from employees. As a professional in Learning & Development (L&D), Human Resource Development (HRD), Organizational Development (OD), or Performance Improvement (PI), you may use surveys to better understand key employee characteristics and perspectives. These can include:

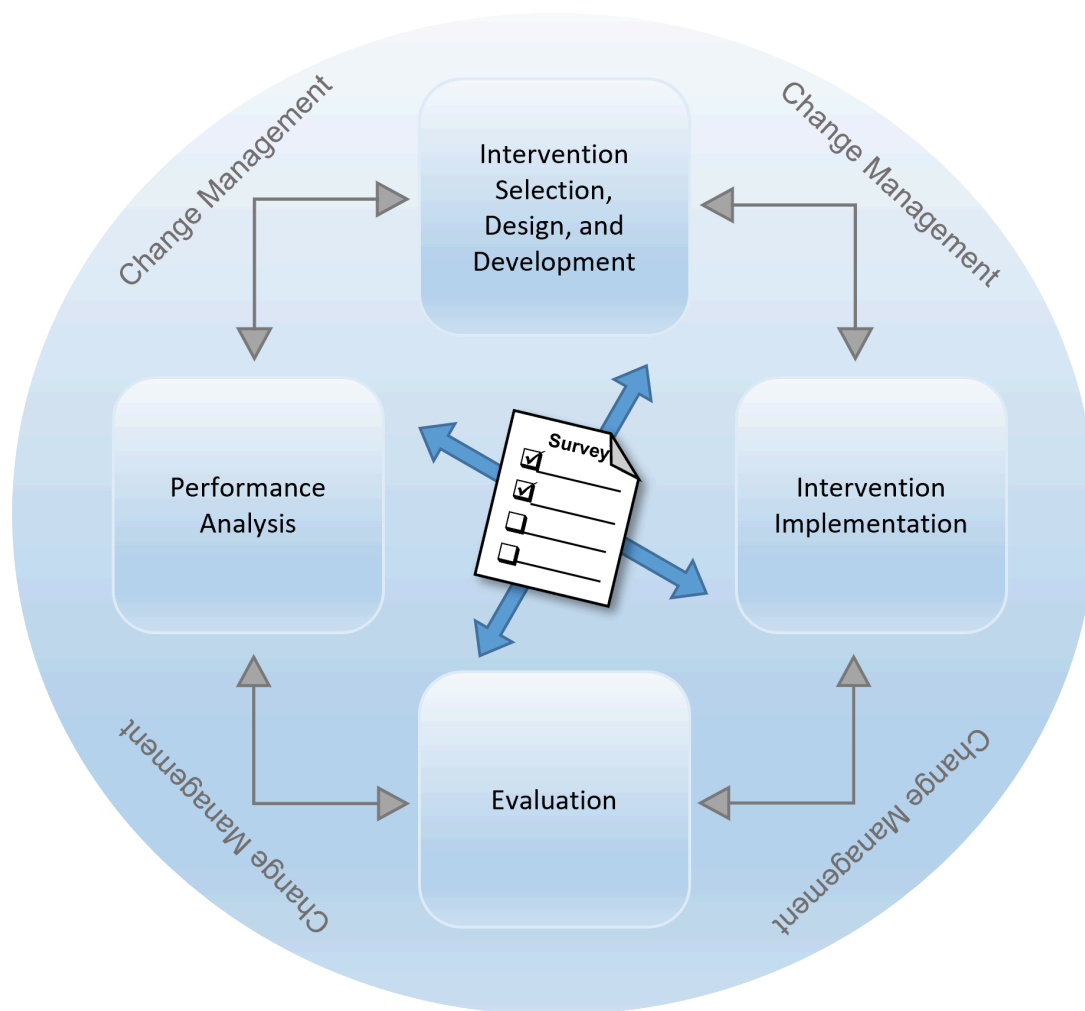
- Engagement levels
- Opinions on leadership styles
- Job satisfaction and overall morale
- Perceptions of organizational culture
- Satisfaction with completed training programs

In many cases, your survey may involve not only internal stakeholders but also external audiences, such as clients, customers, or community members, depending on the scope of your project.

From a performance improvement perspective, you will find the need to conduct surveys across all five phases of your performance improvement process², as illustrated in Figure 2. Specifically, surveys can play a critical role in identifying information during your:

- Evaluation projects,
- Performance analysis (needs assessment) projects,
- Instructional design projects, and/or
- Change management projects.

Figure 2 Surveys Conducted during All Phases of Performance Improvement Process



During evaluation projects, you need to gather input from various stakeholders to

assess the quality of the evaluated programs or products, or to identify areas for improvement. You may collect such data through surveys as well as interviews, focus groups, observations, tests, and reviews of extant data reviews.

Surveys during evaluation projects may address questions such as:

- How user-friendly is the learning management system and how well has it been supporting the development of employee competency?
- Has safety awareness improved among employees since they completed the safety awareness training program?
- Did the onboarding process help new employees get up to speed with their job within the expected timeframe?
- How frequently do supervisors provide coaching to their employees?

During performance analysis projects, you need to collect information to identify performance gaps and analyze environmental or personal factors contributing to those gaps. You may discover such information from performers and other stakeholders via surveys, as well as observations, interviews, focus groups, tests, and/or existing data.

Surveys in performance analysis projects may explore:

- To what extent are employees aware of the organizational goals and performance expectations?
- How valued do employees feel and how does their perception impact their performance?
- What are the key facilitators and barriers to successful performance?
- Are the gaps primarily due to environmental or personal factors, and why?

During instructional design projects, it is essential to understand learner characteristics, supervisor expectations, and subject matter experts' input to inform the instructional design process and desirable learning outcomes. Such information can

be gathered through surveys, interviews, focus groups, and tests with those stakeholders as well as observations and extant data reviews.

Surveys during instructional design projects may investigate:

- What job tasks do subject matter experts or exemplary performers identify as essential for employees?
- How confident and motivated are employees in learning and improving their job performance?
- What tasks do employees find most difficult to understand and perform?

Surveys can also support the **implementation and maintenance of interventions** and the **change management process**. In these phases, the goal is to assess how well solutions are being applied and sustained. Survey questions during intervention and change management phases may include:

The purpose of conducting surveys during these phases is often to find out information about:

- How feasible is the recommended solution?
- How effectively is the intervention being implemented?
- What barriers, if any, have employees encountered during implementation?

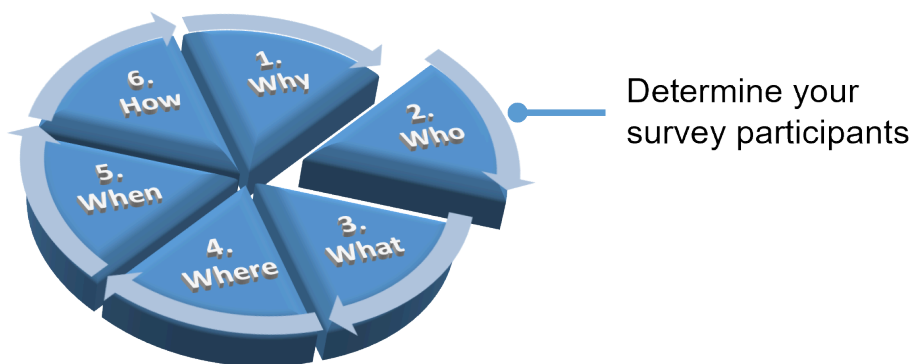
Notes

1. Singer, J. B. (2008). [Five Ws and an H: Digital challenges in newspaper newsrooms and boardrooms](https://doi.org/10.1080/14241270802262468). *The International Journal of Media Management*, 10(3), 122-129. <https://doi.org/10.1080/14241270802262468>
2. Dessinger, J. C., Moseley, J. L., & Van Tiem, D. M. (2012). [Performance improvement/HPT model: Guiding the process](https://doi.org/10.1002/pfi.20251). *Performance Improvement*, 51(3), 10-17. <https://doi.org/10.1002/pfi.20251>

1.3 Surveys conducted with a population or a sample

Determining a clear reason and purpose for conducting surveys is an essential first step (**the Why**). Next, you need to identify whom you will survey and decide whether you will survey the entire population or just a sample. This second step—**the Who**—plays a critical role in effective survey design and data analysis (see Figure 3).

Figure 3 *The Who Step in the Survey Design and Data Collection Process*



You conduct a survey with people who you believe have the information you need. These individuals might be employees, customers, or other relevant stakeholders. But, how many of them actually have the information you are looking for? That group is your population.

If the population is small, you can survey every member of the population. However, if the population is large, you must decide whether to survey the entire group or just a sample. This decision also affects how you will analyze the data—using descriptive statistics, inferential statistics, or both (King & Minimum, 2008¹).

A **population** is the entire group of people whose characteristics you want to understand.

A **sample** is a subset of the population that you decide to study when it is impractical or too costly to study the entire group.

So, start by identifying the population that you wish to study. Ask yourself the following key questions:

- Who are the people I want to describe or explain with the data I collect?
- Who are the people I do not intend to describe with this data?

Below are some examples:

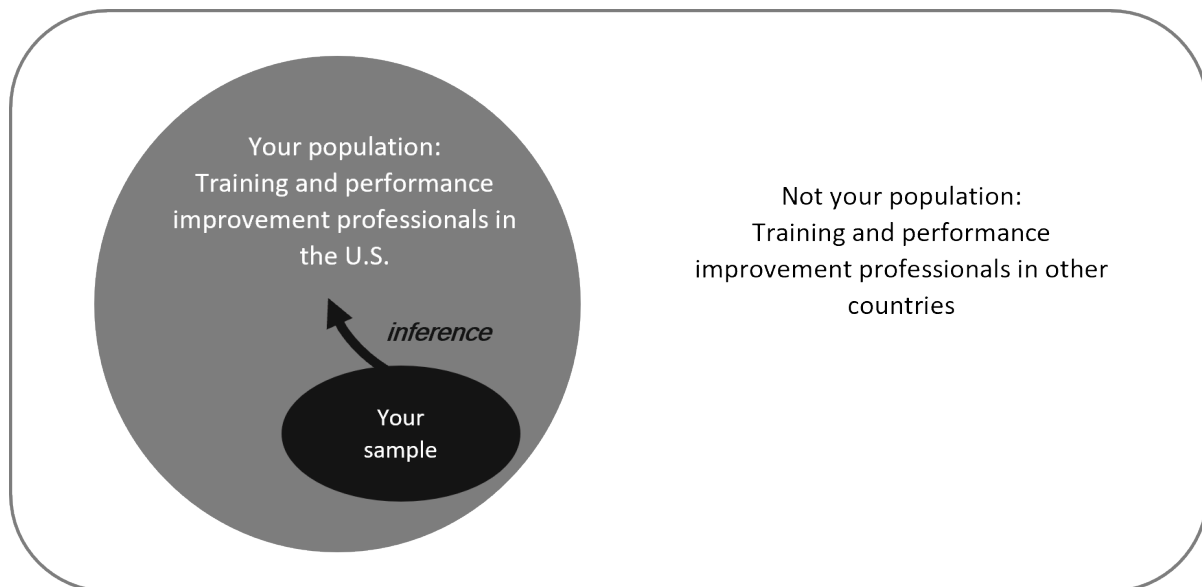
Examples

1. If you want to learn about the general characteristics of the U.S. population, you might refer to data from the U.S. Census Bureau. However, such data usually won't answer specific questions such as whether a new incentive plan will affect employee performance. As an L&D, HRD, OD, or PI professional, you likely don't need to study a national population anyway. Instead, your target populations are usually much smaller groups such as employees in a specific department in your organization, the entire staff of your organization, or members of the local community.
2. If you are studying job satisfaction among a few dozen members of your sales department, then those employees represent your population and that's a small one. Employees in other departments fall outside your population, as you aren't seeking to describe them.
3. Similarly, if you're interested in understanding how U.S. Coast Guard cadets perceive leadership effectiveness, then they are your population. Cadets in the U.S. Navy or Army are not included.

Next, determine whether it's **feasible** and **necessary** to survey with the entire population or if you should use a sample instead.

- If your population is small, you can survey everyone—no sampling needed. For example, when conducting a climate or employee engagement survey in a small organization, it may be practical and even beneficial to survey all employees of the organization. If your survey collects quantitative data, you may be able to analyze results quickly even with a large number of responses, unlike qualitative data which takes more time to process.
- However, if your population is large—for instance, training and performance improvement professionals across the U.S.—surveying everyone may not be practical. In this case, you would select a sample from the population and use the results to make inferences about the population (as illustrated in Figure 4).

Figure 4 *Sample Results Are Used to Infer the Population Results*



When you randomly select a sample from the population, it has a higher likelihood of accurately representing the population. You can then apply **inferential statistics** to analyze your data and draw conclusions about the population based on sample results. Inferential statistics include methods such as *t-tests* (see Appendix A for more details).

By contrast, **descriptive statistics** are used to summarize or describe the characteris-

tics of either a population or a sample. These include measurements such as mean, mode, median, standard deviation, frequency, and percentage.

In survey-based studies, you may use both descriptive and inferential statistics.

Examples

Suppose your organization has 3,000 employees (that's your population).

- If you survey a random sample of 250 employees, you may use descriptive statistics to report the sample's mean scores and standard deviations.
- You can then use inferential statistics, such as a *t-test*, to test your research hypothesis and infer results about the entire 3,000 employee population.

Notes

1. King, B. M., & Miniun, E. W. (2008). *Statistical reasoning for the behavioral sciences* (5th ed.). John Wiley & Sons.

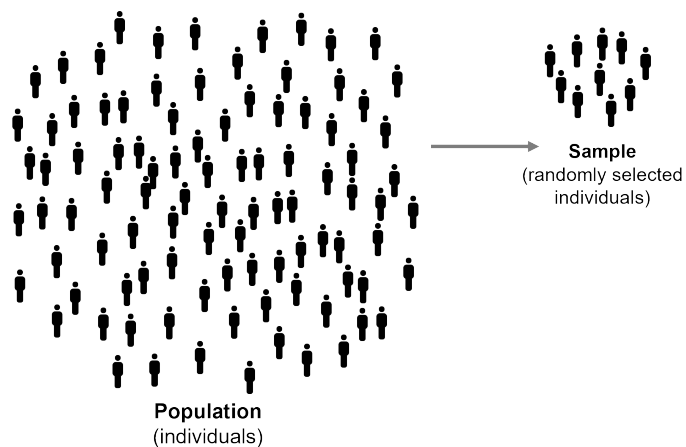
1.4 Sampling methods

If you decide to survey a sample and use inferential statistics to generalize findings to the entire population, you need to select an appropriate sampling method. Researchers have developed several sampling strategies to ensure that the sample accurately represents the population (Fraenkel & Wallen, 1990¹). For example, let's return to the earlier scenario: You want to select a sample of 250 employees from a population of 3,000 employees. The method you choose for selecting those 250 employees will influence the quality and generalizability of your survey results.

1.4.1 Simple random sampling

With simple random sampling, you randomly select individuals from the entire population. This method increases the chance that the sample characteristics are similar to those of its population. For example, a group of 250 employees selected from a population of 3,000 employees would likely reflect the overall demographics, experiences, or opinions of the entire employee (Figure 5).

Figure 5 *Illustrating Simple Random Sampling*

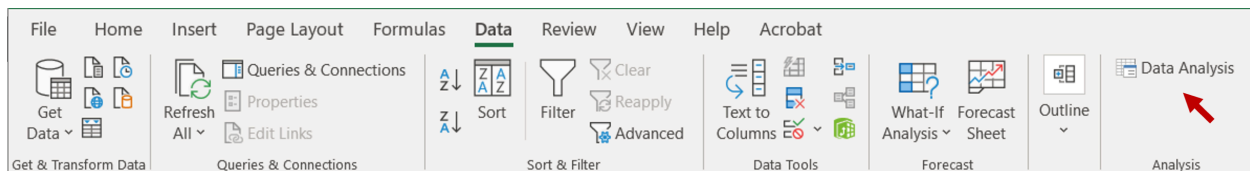


To carry out simple random sampling using Microsoft Excel, assign a unique ID number from 1 to 3,000 to all 3,000 employees. Next, use the `=RANDBETWEEN(1,3000)` function to generate 250 random numbers. Then, select 250 employees whose ID numbers match the randomly generated numbers.

Alternatively, you may use Excel's **Sampling** function from the **Data Analysis Toolpak**. First, you may need to activate the **Data Analysis** menu in Excel. In Excel,

1. Click the **File** tab, and click **Options**.
2. Click **Add-Ins**, and then in the **Manage** box, select **Excel Add-ins**.
3. Click **Go**.
4. In the **Add-Ins** available box, select the **Analysis ToolPak** checkbox, and click **OK**.
5. After successfully loading the Analysis ToolPak, the **Data Analysis** tool appears in the **Analysis** group on the **Data** tab (Figure 6).

Figure 6 *Data Analysis Under the Data Tab*



To randomly select 250 employee IDs from 3,000 IDs,

1. Open an Excel file that contains 3,000 Employee IDs in a column. You may use [Ch1.4-SamplingData.xlsx](#) for practice (Figure 7).

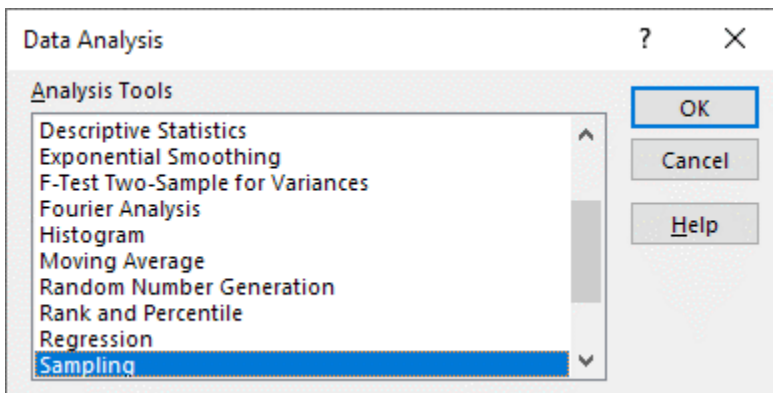
Figure 7 *An Excel File Containing Population Members' IDs*

	A	B	C	D
1	1			
2	2			
3	3			
4	4			
5	5			
6	6			
7	7			
8	8			
9	9			
10	10			
11	11			
12	12			
13	13			
14	14			

Population N = 3,000

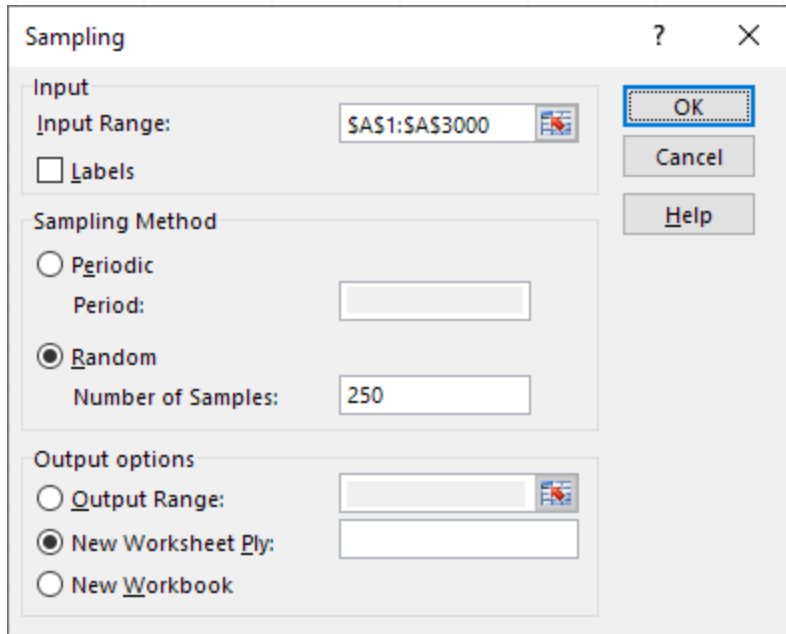
2. Select **Data > Data Analysis** from the menu.
3. Select **Sampling** and click **OK** (Figure 8).

Figure 8 *Selecting Sampling in the Data Analysis Window*



4. In the **Input Range** field, enter **\$A\$1:\$A\$3000** (or select from A1 to A3000). Click **Random**, enter **250** in the **Number of Samples** field, and click **OK** (Figure 9).

Figure 9 *Entering Input in the Sampling Window*

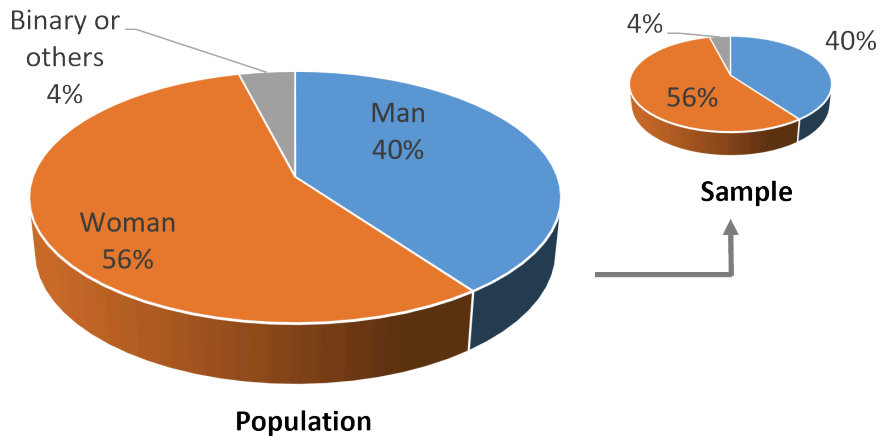


5. In a new worksheet, a randomly selected 250 IDs will be presented. You may sort the numbers in the column from smallest to largest.

1.4.2 Stratified random sampling

Use stratified random sampling when you want your sample to reflect the same proportions of specific subgroups (e.g., gender, department, branch) as in the overall population. For example, suppose your population of 3,000 employees includes 40% of employees who identify as men, 56% as women, and 4% as nonbinary or another gender identity. To maintain that same proportion in a sample of 250, you would randomly select 100 employees (40% of 250) who identify as men, 140 employees (56% of 250) who identify as women, and 10 employees (4% of 250) who identify as nonbinary or another gender identity (Figure 10).

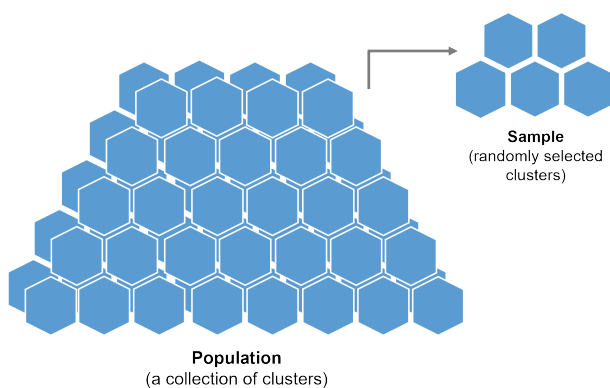
Figure 10 *Illustrating Stratified Random Sampling*



1.4.3 Cluster random sampling

Let's say, your organization has 3,000 employees distributed across 60 branches, with each branch considered a cluster. On average, each branch (cluster) has approximately 50 employees. Instead of randomly selecting individuals from the entire 3,000 employee pool, you can randomly select five branches (clusters) as illustrated in Figure 11 (i.e., 5 random branches x approx. 50 employees at each branch = approx. 250 employees). However, this method may not work well for selecting a sample of 250 employees if the number of employees per branch varies significantly. In that case, you can first perform cluster random sampling to select branches, and then apply simple random sampling or stratified random sampling within each selected cluster to ensure a more accurate and representative sample.

Figure 11 *Illustrating Cluster Random Sampling*



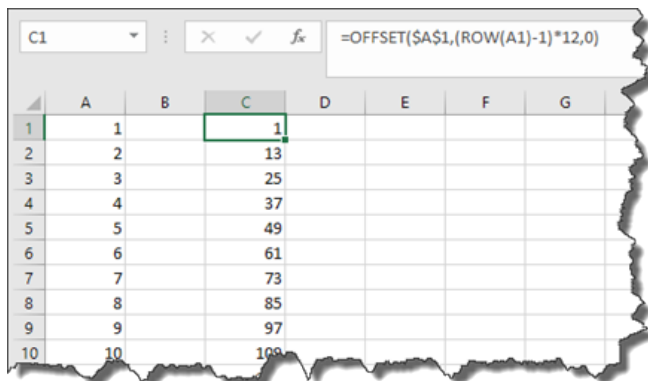
1.4.4 Systematic sampling

With systematic sampling, you select every *n*th individual from a list—in this case, every 12th employee from a coded list of 3,000 employees. To do this in Excel:

1. Assign a unique ID to each employee in Column A, from 1 to 3,000.
2. In cell C1, enter the following formula: =OFFSET(\$A\$1,(ROW(A1)-1)*12,0)
3. Copy this formula down for 250 rows in Column C.

This will return the ID of every 12th employee from the list, giving you a systematic sample of 250 employees (Figure 12).

Figure 12 *Selecting Every 12th ID in Excel*



The screenshot shows an Excel spreadsheet with a formula bar at the top displaying the formula =OFFSET(\$A\$1,(ROW(A1)-1)*12,0). The spreadsheet has columns A through G and rows 1 through 10. Column A contains the numbers 1 through 10. Column C contains the numbers 1, 13, 25, 37, 49, 61, 73, 85, 97, and 109, which are every 12th number from the list in column A. The formula bar shows the formula =OFFSET(\$A\$1,(ROW(A1)-1)*12,0).

	A	B	C	D	E	F	G
1	1		1				
2	2		13				
3	3		25				
4	4		37				
5	5		49				
6	6		61				
7	7		73				
8	8		85				
9	9		97				
10	10		109				

1.4.5 Convenience sampling

With convenience sampling, you select participants who are easy to access or willing to participate. For example, you might survey 250 employees from a couple of departments where you have strong working relationships and support for your survey study. This is the least rigorous sampling method because the sample may be biased and not representative of the large population. As a result, findings based on convenience samples should be interpreted with caution, especially when generalizing to the full population.

To determine a sample size based on the population size, confidence level, and margin of error, you may use a sample size calculator freely available on the web such as: [SurveyMonkey's sample size calculator](#).

Notes

1. Fraenkel, J. R., & Wallen, N. E. (1990). *How to design and evaluate research in education*. McGraw-Hill.

1.5 Chapter activity: How does your organization use surveys?

Chapter Activity: How does your organization use surveys?

This activity includes Part A and Part B.

A. Reflect on how your current or previous organization has used surveys. Summarize your insights by answering the following questions:

1. What was the purpose of the survey? Briefly describe the main goal. If a copy of the survey questionnaire is available, please share it in class.
2. Was the survey conducted with a population or a sample? If a sample was used, identify the sampling method used (if you are aware of it).
3. How was the survey administered?
 - a. face-to-face
 - b. via telephone
 - c. by mail
 - d. online
4. Was the chosen type of survey administration (your response to Q3 above) appropriate based on the survey's purpose (your response to Q1 above)?
5. Were the expected advantages of that survey administration type (as shown in Table 1 in Section 1.1) achieved?
6. What could have been done differently to make the survey more effective? Discuss this from the organization's point of view, the respondents' perspective, or both.

B. Now, plan a new survey for your organization or a hypothetical one. Respond to the following questions:

1. What is the purpose of the survey? Briefly describe the main goal.
2. Will you survey a population or a sample, and why? Explain your decision. If using a sample, specify which sampling method will be used, and why.
3. How will you administer the survey? Select one method and explain your choice:

- a. face-to-face
 - b. via telephone
 - c. by mail
 - d. online
4. Why are you using a survey instead of another data collection method? Discuss the advantages of using a survey in this case. Also, identify any potential challenges or limitations you should watch out for during the process.

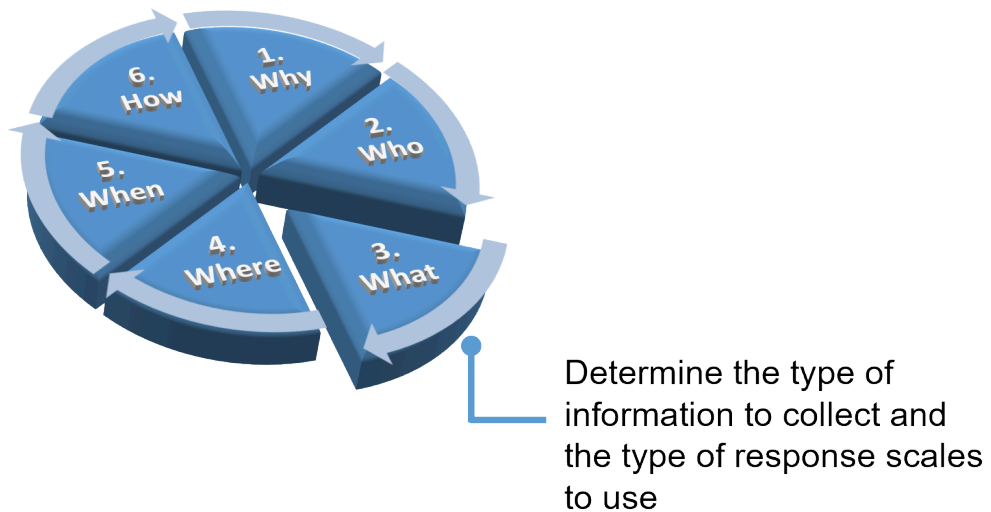
2. UNDERSTAND THE BASIC STRUCTURE OF SURVEY ITEMS

2.0 Introduction

Now that you have a clear purpose for conducting your survey and understand whether you will involve a population or a sample, the next step in the survey design and data analysis process is to actually design your survey questionnaire.

This **What** step (Figure 13) involves identifying the specific information you need to collect and selecting appropriate response scales. These topics will be explored in this chapter (Chapter 2) and expanded upon in the next three chapters (Chapters 3, 4, and 5).

Figure 13 The What Step in the Survey Design and Data Analysis Process



Learning Objectives

After completing this chapter, you will be able to:

- Describe the survey anatomy

- Distinguish survey items that intend to capture facts and that intend to measure perceived quality/value
- Recognize a set of survey items that intend to measure constructs
- Recognize different levels of measurement scales (nominal, ordinal, interval, ratio)

2.1 Survey anatomy

You design survey questionnaires to gather specific information you want to learn from your respondents. When creating a survey questionnaire with closed-ended survey items, you need to consider two essential components:

1. How you will request information from respondents
2. How you will capture their responses

You can request information using a question or a statement (or a phrase). To capture that information, you can use either a closed-ended response scale or an open-ended format. We will refer to each pairing of a question or statement and its corresponding response format as a **survey item**. This book focuses on designing closed-ended survey items—that is, items that include a response scale—regardless of whether the information is requested through a question or a statement, as illustrated below. More examples of survey items are presented in Table 2.

When requesting information using a question

How friendly were the staff?

- Not friendly at all
- A bit friendly
- Somewhat friendly
- Quite friendly
- Very friendly

When requesting information using a statement

The staff were friendly.

- Strongly disagree
- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

Table 2 *Different Ways to Request and Capture Information*

Capture information	Requesting information using a question	Request information using a statement (or a phrase)
With a response scale (closed-ended)	<p>Did you have a good experience?</p> <ul style="list-style-type: none"> ○ Yes ○ No <p>Which one represents your experience?</p> <ul style="list-style-type: none"> ○ Terrible ○ Mediocre ○ Good enough ○ Best ever 	<p>I had a good experience.</p> <ul style="list-style-type: none"> ○ Strongly disagree ○ Somewhat disagree ○ Neutral ○ Somewhat agree ○ Strongly agree <p>My experience was:</p> <ul style="list-style-type: none"> ○ Terrible ○ Mediocre ○ Good enough ○ Best ever
Without a response scale (open-ended)	How would you describe your experience?	Please describe your overall experience:

2.2 Survey items capturing facts or measuring perceived quality/value

Survey items can be designed to capture different types of information from respondents. Broadly, these fall into two categories:

1. Factual information (what something *is* or *was*)
2. Perceived quality or value (how someone *feels about* or *evaluates* something)

2.2.1 Capturing facts

Fact-capturing survey items focus on objective or descriptive information. These may include:

- demographic information such as name, date of birth, gender identity, racial or ethnic background, professional and academic experience, and years of employment
- experiential information such as whether the respondent has completed a specific training program, used a particular tool, or participated in a project.

These items are not designed to assess opinions or judgments but to collect non-evaluative, concrete data. Although many of these survey items use closed-ended response formats, they typically do not use rating scales. Instead, they use options like checkboxes, dropdowns, or 'select all that apply' formats. Below are few examples.

What is your date of birth?

Month [][] Date [][] Year [][][][]

Which racial group(s) do you belong to? Select all that apply.

- White
- Black or African American
- American Indian and Alaska Native
- Asian
- Native Hawaiian and Other Pacific Islander
- Please describe: _____

How many times have you used this device?

- Never
- 1-2 times
- 3-5 times
- More than 5 times

How did you find out about this position? Select all that apply.

- Job fair
- Newspaper
- Recruiter
- Friend/colleague
- Other source. Please describe: _____

When did you start this program?

- Spring, 2024
- Summer, 2024
- Fall, 2024
- Winter, 2024

Did you have to use the help desk to get assistance?

- Yes
- No

2.2.2 Measuring perceived quality or value

Some survey items are designed to measure how respondents perceive the quality, effectiveness, or value of a program, product, service, or individual. These items typi-

cally use a rating scale and require respondents to reflect on their opinions, attitudes, or personal experiences. Unlike fact-capturing items, these survey items are subjective and help assess how well something is received or evaluated by the target audience. Below are examples of survey items that measure perceived quality or value using rating scales:

How well did the trainer facilitate the discussion?

- Very ineffectively
- Somewhat ineffectively
- Somewhat effectively
- Very effectively

The program helped me understand the areas that I need to improve to become a more effective manager.

- Strongly disagree
- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

The content of the program was easy to learn.

- Strongly disagree
- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

Management seeks input from employees when making major decisions.

Very untrue 1 2 3 4 5 6 7 Very true

I am engaged in meaningful work.

- Yes
- No

How knowledgeable about the products was the salesperson?

- Not at all
- A little bit
- Somewhat
- Quite a bit
- Very much

How was the service provided by the agent?

- Totally ignored me
- Was very slow and kept me waiting unnecessarily
- Completed my request, but took longer than expected
- Completed my request promptly, but made some errors
- Completed my request promptly and accurately

Would you recommend this product to others?

- Yes
- No

Whether you are capturing factual information or measuring perceived value or quality, each survey item collects information about an abstract concept, also known as a construct. The following section provides further details on survey items designed to measure constructs.

2.3 Survey items measuring constructs

As discussed earlier, a survey is just one of the many data collection methods available to you. You choose to conduct a survey when it is difficult or impossible to directly observe the information you need or obtain it from other sources.

For example, if you need to know how well nurses are performing patient rounding, the most direct method for collecting this information would be to observe the nurses' performance. Alternatively, you could review rounding checklists if they are formally implemented. You could also ask nurses to respond to a survey item asking how many rounds they completed during their shift, or how easy or difficult to perform patient rounding. However, self-reported data may not be as accurate as data obtained from direct observations or checklist reviews.

Though surveys have some weaknesses, they are not necessarily less useful than other data collection methods. In fact, there are situations where surveys are the most appropriate method, and sometimes the only way, to obtain the information you need. When the information you require is abstract in nature and resides in people's mind—such as their perceptions of the quality or value of something, which cannot be *directly* observed or recorded—you would use a survey to retrieve that information.

These abstract pieces of information, which cannot be directly observed or easily assessed, are called **constructs**. Constructs are hidden until studied and revealed, and are often referred to as *latent variables* in research. Interestingly, even facts such as age, gender, or time can be considered socially-agreed constructs. Therefore, survey questionnaires that ask for facts and perceptions about the quality or value of something are also measuring constructs.

Constructs can be simple or complex. Simple constructs such as employment status or job title can be measured using a single survey item for each. For example:

Are you an employee of Sawtooth Health System?

- Yes
- No

What is your job title?

- Instructional designer
- E-learning developer
- Graphic designer
- Programmer
- Describe: _____

In contrast, more complex constructs, particularly those related to people’s inner thoughts, feelings, or perceptions, usually require multiple survey items to measure them accurately.

For example, if you are evaluating trainees’ satisfaction with a workshop, asking a single question such as “How satisfied are you with the workshop?” may not provide enough insight. Satisfaction is a multifaceted construct, and a single survey item does not capture the various elements that contribute to it. To measure satisfaction more effectively, you would use multiple survey items that assess specific aspects of the workshop, such as the quality of the content, the relevance of the material to their job tasks, the amount of information provided, etc. as shown in Exhibit 1.

***Exhibit 1** A Set of Sample Survey Items That Intends to Measure Trainee Satisfaction as a Construct*

How satisfied are you with the following aspects of the workshop?

Question	Not satisfied	A little bit satisfied	Somewhat satisfied	Quite a bit satisfied	Very satisfied
Quality of the content	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Amount of the content	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Relevance to my job	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Workshop length	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Trainer's preparedness	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

For more complex constructs—such as intelligence and organizational culture—a single survey item is not sufficient. Instead, you need to use a battery of survey items to measure the multiple dimensions that make up the construct. For example, the Dimensions of the Learning Organization Questionnaire (DLOQ), developed by Marsick and Watkins (2003)¹, includes 43 survey items designed to measure seven dimensions of a learning organization and additional 12 survey items that measure learning organizational results. Exhibit 2 presents a few sample survey items from the DLOQ:

Exhibit 2 Sample Survey Items of the DLOQ

Question	Almost Never 1	2	3	4	5	Almost Always 6
1. In my organization, people openly discuss mistakes in order to learn from them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. In my organization, people identify skills they need for future work tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. In my organization, people help each other learn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Therefore, when designing a questionnaire, it's important to consider constructs you intend to measure. Some constructs are simple, such as capturing factual information, while others—particularly those measuring people's perceived quality or value of something—are more complex and require multiple survey items. To effectively measure these complex constructs, you need to ensure that the survey items you're using are all aligned with the same construct. Additionally, it's beneficial to use the same rating scale across these survey items. This allows for a consolidated summary of the results (e.g., reporting an average score based on the multiple survey items that measure the same construct). We will explore response scales and how they support this process in the following sections.

Notes

1. Marsick, V. J., & Watkins, E. W. (2003). [Demonstrating the value of an organization's learning culture: The dimensions of the learning organization questionnaire](#). *Advances in Developing Human Resources*, 5(2), 132-151.

2.4 Survey items with different levels of measurement scales

To select an appropriate response scale for each survey item, it's important to understand the different levels of measurement scales. Stanley Smith Stevens, a professor of psychology at Harvard University, conceptualized four different levels of measurement scales: nominal, ordinal, interval, and ratio scales (Stevens, 1946¹).

2.4.1 Nominal scales

Nominal scales categorize items by assigning names or labels without implying any assumed rank-order among them. Since nominal scales do not involve rank-ordered ratings, they are not suitable for survey items intended to *rate* respondents' perceived quality or value (e.g., questions like, "Was this terrible, bad, good, or excellent?").

Instead, fact-capturing survey items often use nominal response scales. While numerical values (e.g., 1, 2, 3, 4, etc.) can be assigned to these options, these numerical values are used solely for identification purposes and do not imply any order or ranking. For example:

What sex were you assigned at birth (as listed on your original birth certificate)?

- Male
- Female
- Intersex

Here, the options can be coded numerically (Male = 1, Female = 2, Intersex = 3) for convenience in data processing. However, it doesn't imply that Female and Intersex are higher or better than Male, or vice versa. Similarly, consider the following examples:

What is your current job title?

- Instructional designer
- Trainer

- E-learning developer
- Other

How did you find out about this program? Select all that apply.

- Word-of-mouth
- From a brochure
- From a radio advertisement
- From the Web search
- At a conference
- Other source. Please describe: _____

2.4.2 Ordinal scales

Ordinal scales consist of a group of rank-ordered level where the distance between a set of two consecutive levels is not necessarily the same as the distance between another set of two consecutive levels. For example, consider the following scale, which ranks frequency of appreciation. While the levels are ranked in order of how often someone feels appreciated, the distance between *Never* and *Seldom* may not be the same as the distance between *Seldom* and *Often* or between *Often* and *Always*. Nonetheless, this type of ordinal scales can measure people's perceived quality or value.

How often did you feel appreciated for the service you provided?

- Never
- Seldom
- Often
- Always

2.4.3 Interval scales

Interval scales contain a group of rank-ordered levels where the distance between each pair of consecutive levels is consistent and equal. Interval scales can also be used to measure people's perceptions. However, one important characteristic of internal scales is that they do not have an absolute zero point. Zero on an interval scale is an arbitrary zero value. For example, in the following response scale, the *Very dissatis-*

fi ed option is assigned the value 0. However, zero does not mean there is a complete absence of satisfaction—it simply represents the lowest point on the scale.

How satisfied are you with the product?

Very dissatisfied 0 1 2 3 4 5 6 7 8 9 10 Very satisfied

2.4.4 Ratio scales

Ratio scales have an equal distance between intervals, and they have an absolute zero value, which indicates the absence of the attribute measured (e.g., a value of zero in income = no income). Examples are measurements of age, length, weight, height, income, and time, as shown below:

What is your current annual income? \$ _____

What is your current age? _____

2.4.5 Project the type of data you need before designing survey items

When designing survey items, it's important to project the type of data you intend to collect before selecting appropriate response scales.

When capturing factual information, you may use nominal scales for categorical data (e.g., gender or ethnic groups, job titles), ordinal scales for rank-ordered data (e.g., military ranks), interval scales for rank-ordered and meaningful intervals (e.g., years).or ratio scales for precise measurements with a true zero value(e.g., income, age).

On the other hand, when measuring perceived quality or value, the design process is more complex as a variety of rating scales for measuring perceptions are available. Ordinal or interval scales are commonly used to measure perceived quality and value of something (e.g., job satisfaction, employee performance, or customer service).

In terms of flexibility in usage, ratio scales produce the most flexible data, followed by interval scales, ordinal scales, and nominal scales in that order. For example, after you collected actual income or age values (ratio), you can group the values into interval or ordinal categories and generate frequencies as shown below:

\$1 – \$20,000: N = 12
\$20,001 – \$40,000: N = 45
\$40,001 – \$60,000: N = 52
\$60,001 – \$80,000: N = 79

Similarly, after you have collected interval data such as satisfaction levels measured on a 11-point scale (interval data), you may group the data into several ordered categories such as low (0-3), medium (4-7), and high (8-10) satisfaction levels (ordinal data), if needed.

However, once you have collected nominal or ordinal data, you cannot report the data as interval or ratio data. For example, asking for age in categories like the survey item below (nominal scale) will yield categorical data, not continuous data. If you need exact ages for analysis, this survey item would not be suitable:

Your current age:

- in the 20s
- in the 30s
- in the 40s
- in the 50s
- in the 60s

In conclusion, before designing your survey items, **identify the type of data you need and select response scales that align with the appropriate level of measurement level.**

Notes

1. Stevens, S. S. (1946). [On the theory of scales of measurement](#). *Science*, 103(2684), 677–680.

2.5 Chapter activity: What do they measure, facts or perceived quality/value?

This activity has two parts: Part A and Part B.

A. Using the sample survey questionnaire provided in Exhibit 3:

1. Identify any survey items that measure a construct. If multiple items measure the same construct, identify them as a group.
2. Determine whether each item measures a fact or perceived quality/value.
3. Identify the response scale type for each item: nominal, ordinal, interval, or ratio.

Exhibit 3 Sample Survey Questionnaire to Identify What Individual Survey Items Measure

1. Workshop topic: _____

2. Workshop date: _____

3. How did you hear about the workshop? Please select all that apply.

- Co-worker
- Family/Friend
- Flyer/Newspaper ad
- Email
- Website
- Other. Please describe: _____

4. How would you rate your experience with the following elements of the workshop?

Question	Way below average	Slightly below average	Average	Slightly above average	Way above average	Don't know
Initial registration process	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overall information pace	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Facilitator's knowledge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Classroom environment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Handouts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. How likely are you to recommend this workshop to others?

- Very unlikely
- Somewhat unlikely
- Somewhat likely
- Very likely

6. Overall, how satisfied are you with the workshop?

- Very unlikely
- Somewhat unlikely
- Somewhat likely
- Very likely

7. What would you suggest we change to make the workshop better for the future participants?

8. Write your job title:

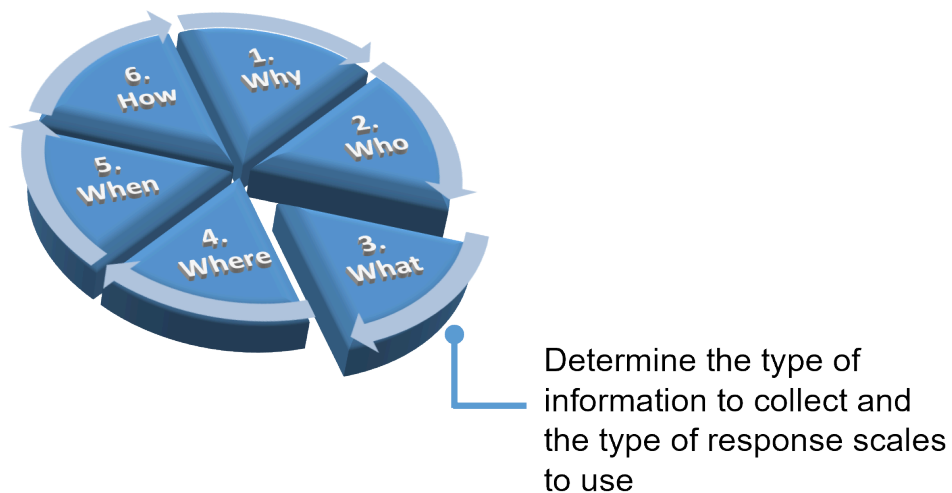
B. Now, look around and find a real-world survey questionnaire used in an organization—This could be from your current workplace, a doctor's office, a retail store, or another setting where surveys are used to gather feedback or information.

1. Post a copy of the survey questionnaire (if it is a paper version, you may recreate the survey items or post a photo of the survey taken with your cell phone).
2. Identify one or more survey items that measure a construct.
3. Determine whether each item measures a fact or perceived quality/value.
4. Identify the response scale type for each item: nominal, ordinal, interval, or ratio.
5. Discuss any changes you would recommend to improve the effectiveness of the survey questionnaire.

3. EXPLORE VARIOUS RATING SCALES

3.0 Introduction

We are still in the **What** stage, now focusing more on how to determine the appropriate type of response scale.



[A Copy of **Figure 13** *The What Step in the Survey Design and Data Analysis Process*]

Various rating scales can be used to measure the perceived quality or value of something, typically falling under ordinal or interval scale categories. These rating scales may be presented using numbers, descriptive terms, and/or graphics, resulting in either discrete or continuous data. In this chapter, we will explore the different types of rating scales.

Learning Objectives

After completing this chapter, you will be able to:

- Distinguish between discrete and continuous rating scales

- Identify numerical rating scales
- Recognize verbal descriptor scales, such as the Likert scale
- Recognize graphic rating scales
- Identify web-based sliders

3.1 Discrete vs. continuous rating scales

Rating scales can be categorized based on whether they produce discrete or continuous data.

Discrete rating scales offer a limited set of fixed response options. Examples include the Likert scale and other Likert-type scales, as shown below. When using these verbal descriptor rating scales, respondents select one of the predefined options, yielding discrete data. The recorded data can be summarized using frequencies or percentages. Or, numerical values can be assigned to the options to calculate average scores.

Example 1	Example 2	Example 3
<ul style="list-style-type: none">○ Very untrue (%)○ Somewhat untrue (%)○ Somewhat true (%)○ Very true (%)	<ul style="list-style-type: none">○ Strongly disagree (%) (1)○ Disagree (%) (2)○ Neutral (%) (3)○ Agree (%) (4)○ Strongly agree (%) (5)	<ul style="list-style-type: none">○ Terrible (1)○ Poor (2)○ Mediocre (3)○ Fair (4)○ Good (5)○ Very good (6)○ Excellent (7)

A partially-labeled Likert scale still generates discrete data:

Strongly disagree	1	2	3	4	5	Strongly agree
-------------------	---	---	---	---	---	----------------

Similarly, the following partially-labeled rating scales (a.k.a. semantic differential) also produce discrete data:

The workshop was:

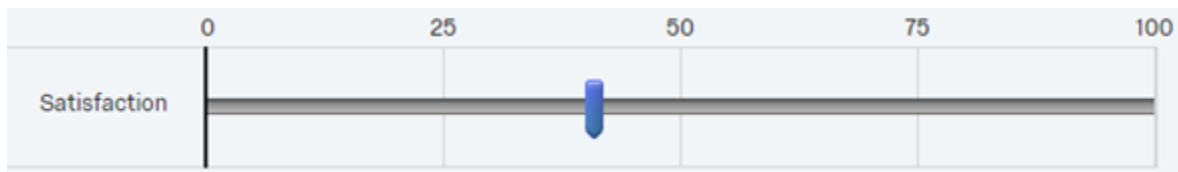
Boring 1 2 3 4 5 Interesting

Irrelevant 1 2 3 4 5 Relevant

Unhelpful 1 2 3 4 5 Helpful

Continuous rating scales allow measurement on a continuum, capturing both whole numbers (e.g., 1, 2, 3) and fractional values (e.g., 1.2, 3.324, 114.59). To collect this type of data, the scale must let respondents select any point along the continuum.

A common example is a web-based slider, which can be configured to record values with multiple decimal places. For instance, if you use a slider to measure satisfaction from 0 to 100, a respondent might place the bar at **40.95**, reflecting a precise level of response.



3.2 Numerical rating scales

Numerical (or numeric) rating scales are designed with numbers (integers) along with a few verbal labels or visual cues. They can be designed to measure various aspects—such as levels of change, satisfaction, or quality—using different ranges of numerical values (e.g., an 11-point scale from 0 to 10, a 7-point scale from -3 to +3, or a 5-point scale from 1 to 5).

These scales are typically ordinal or interval in nature and are considered discrete rating scales, as they produce specific integer values as data. For example, with a 5-point numerical rating scale, respondents can select only one of the five numbers, generating values like 1, 2, 3, 4, or 5. Numerical rating scales are not intended to record data in fractional numbers such as 1.24, 3.26, or 4.38.

Numerical rating scales are usually partially labeled, with only the two endpoints labeled, as shown below.

Poor	1	2	3	4	5	Excellent
------	---	---	---	---	---	-----------

Unlikely	1	2	3	4	5	6	7	Likely
----------	---	---	---	---	---	---	---	--------

Very dissatisfied	-3	-2	-1	0	+1	+2	+3	Very satisfied
-------------------	----	----	----	---	----	----	----	----------------

A type of numerical rating scale, known as the *semantic differential*, consists of a set of partially-labeled N-point scales with two opposing end labels. Typically, 5-point or 7-point scales are used. The bipolar adjectives at each end describe contrasting traits of the object being measured. For example, a semantic differential scale designed to measure a trainers' performance might include the following set of opposing descriptors:

Please rate the trainer's performance:

[negative]	1	2	3	4	5	6	7	[positive]
Unready	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Prepared
Distant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Friendly
Inexperienced	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Knowledgeable
Amateur	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Professional
Rude	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Respectful
Terrible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Outstanding

Numerical rating scales may include more than two ending labels. As shown below, some provide one or more intermediate labels between the endpoints, to offer more context for each response option. Numerical rating scales can be unipolar or bipolar. The first two examples below illustrate unipolar scales, while the third example shows a bipolar scale.

0	1	2	3	4	5	6	7	8	9	10
No pain					Moderate pain					Worst pain ever

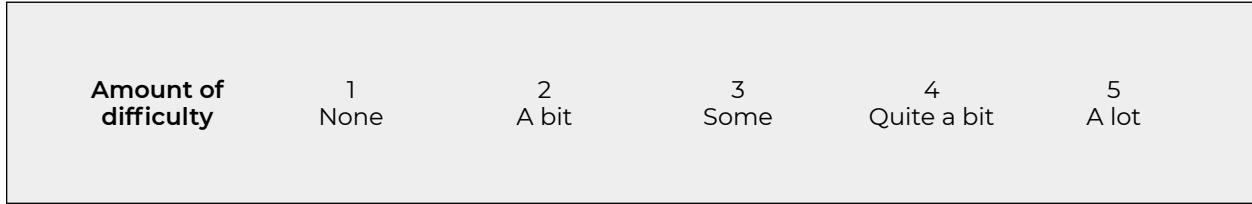
0	1	2	3	4	5	6	7	8	9	10
No change		A bit			Some			Much		Most possible change



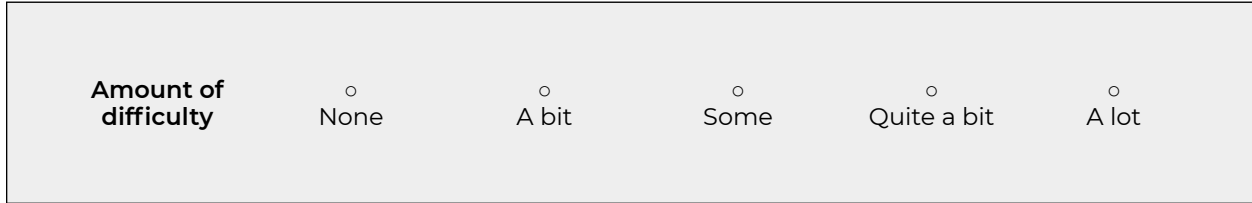
Numerical rating scales are relatively easy to develop and use; however, they should be applied with caution. Descriptive labels added to numerical rating scales assist respondents in selecting a number on the scale that best represents their experience or opinion. However, partially-labeled numerical rating scales leave room for individual interpretation, which can vary widely among respondents.

For example, let's say, Respondent A and Respondent B both have the same level of mild pain from the same symptom, but Respondent A may choose 1 and Respondent B may choose 2 or 3 on the 11-point pain-measuring scale shown above.

This variation may lead you to consider fully labeling each point on a numerical rating scale, as shown below:



However, fully-labeled rating scales are better categorized as **verbal descriptor scales** rather than numerical rating scales, because now respondents are making their selections based primarily on the descriptive labels rather than the numbers. In these cases, the numbers (e.g., 1 through 5) serve mainly as coding values for data analysis and can even be removed from the scale entirely, as shown below. More information about verbal descriptor scales is presented in the following section.



3.3 Verbal descriptor scales including the Likert scale

As explained in the previous section, numerical rating scales present numerical values for respondents to choose from, while verbal descriptor scales use short verbal phrases as the response options.

Verbal descriptor scales are typically ordinal in nature and can be displayed either horizontally or vertically. Numeric values may be added next to the descriptive terms for codification purposes; however, this does not automatically make the scale an interval scale. That said, some verbal descriptor scales may still be treated as interval scales during data analysis, depending on the context and assumptions.

<input type="radio"/> Significant improvement needed (1)	<input type="radio"/> Some improvement needed (2)	<input type="radio"/> Met expectations (3)	<input type="radio"/> Exceeded expectations (4)	<input type="radio"/> Not applicable
<input type="radio"/> Poor (1)	<input type="radio"/> Fair (2)	<input type="radio"/> Good (3)	<input type="radio"/> Very good (4)	<input type="radio"/> Excellent (5)

Verbal descriptor scales, as shown in the two examples above, can be bipolar (e.g., job knowledge) or unipolar (e.g., quality). A common type of bipolar verbal descriptor scale is the Likert scale. Rensis Likert, an American social psychologist, first introduced a five-point psychometric scale for measuring a series of attitude-related propositions in the 1930s (Likert, 1932¹). The wording used in the initial Likert scale was: *Strongly Approve*, *Approve*, *Undecided*, *Disapprove*, and *Strongly Disapprove*. Over time, the wording evolved from *Approve* to *Agree*, resulting in the Likert scale we know today: *Strongly Agree* and *Agree* on one side, *Disagree* and *Strongly Disagree* on the other side, and a midpoint in the middle.

Although originally a 5-point scale, the Likert scale has been adapted into various formats, including odd-numbered (e.g., 7-point) and even-numbered (e.g., 4- or 6-point)

versions. Even-numbered Likert scales omit a midpoint to encourage a more decisive response. Both odd- and even-numbered versions are discrete rating scales, offering a limited set of fixed response options.

4-point Likert scale	5-point Likert scale	6-point Likert scale	7-point Likert scale
<ul style="list-style-type: none"> • Strongly disagree • Disagree • Agree • Strongly agree 	<ul style="list-style-type: none"> • Strongly disagree • Disagree • Neutral • Agree • Strongly agree 	<ul style="list-style-type: none"> • Strongly disagree • Moderately disagree • Slightly disagree • Slightly agree • Moderately agree • Strongly agree 	<ul style="list-style-type: none"> • Strongly disagree • Moderately disagree • Slightly disagree • Neutral • Slightly agree • Moderately agree • Strongly agree

There are also Likert-type rating scales that measure aspects other than agreement—such as satisfaction, likelihood, importance, or quality. For example, the perceived quality of service can be measured with the following Likert-type scale that asks how dissatisfied or satisfied customers feel.

○ Very dissatisfied	○ Somewhat dissatisfied	○ Neutral	○ Somewhat satisfied	○ Very satisfied
-----------------------------------	---------------------------------------	---------------------	------------------------------------	--------------------------------

Notes

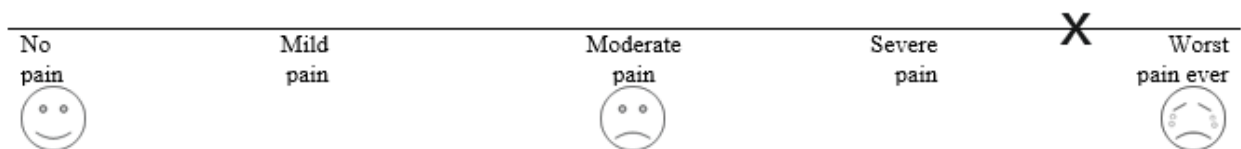
1. Likert, R. (1932). A technique for the measurement of attitudes. In R. S. Woodworth (Ed.), *Archives of Psychology* (Vol. 22, No. 140, pp. 5-55). The Science Press.

3.4 Graphic rating scales and Web-based sliders

Graphic rating scales were developed in the early 20th century as a method for performance appraisal. These scales typically consist of several simple descriptive terms arranged on a horizontal line, representing different levels of a particular trait (Freyd, 1923¹). Graphic rating scales were originally designed as continuous rating tools, allowing respondents to place a checkmark at any point on the line to indicate their rating level. Then, the distance from one edge to the marked point was measured with a stencil. It can be time-consuming to use graphic rating scales in paper questionnaires and manually record the data.

Today, graphic rating scales are not commonly used. However, you can consider a pain-measuring numerical rating scale presented as a graphic rating scale, where patients are asked to mark their pain level with an 'X' (Figure 14).

Figure 14 A Graphic Rating Scale Designed to Measure Pain Levels



When using Web-based survey systems, some of the scales previously described can be easily converted to sliders. **Sliders** are typically designed as continuous rating scales, allowing respondents to select any value within a defined range. For example, on a continuous rating scale from 0 to 10, respondents might select 3.0, 4.2, or 8.1. However, sliders can also be configured to “snap to grid,” limiting responses to specific intervals. That is, the slider shown in Figure 15 with a range from 0 to 10 would function as an 11-point discrete scale, offering only 11 different values to choose from. Thus, you should be aware that sliders with the appearance of a continuous rating scale may produce discrete data (Chyung, Swanson, Roberts, & Hankinson, 2018²). Sliders can also be designed with engaging visuals such as stars, gauges, or thermostats (Figure 16).

Figure 15 A Numerical Rating Scale on a Web-Based Slider Provided in Qualtrics

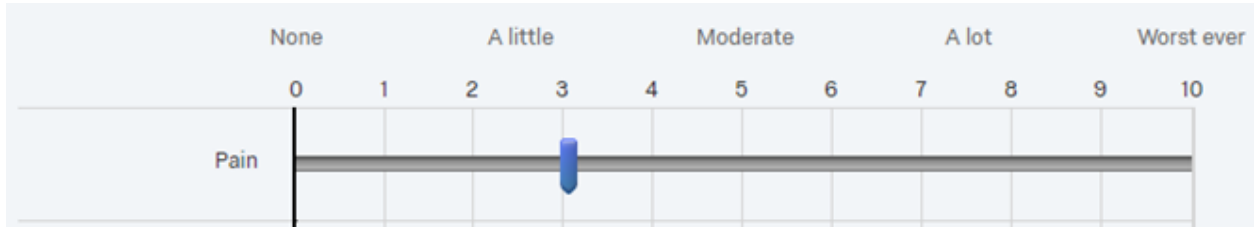
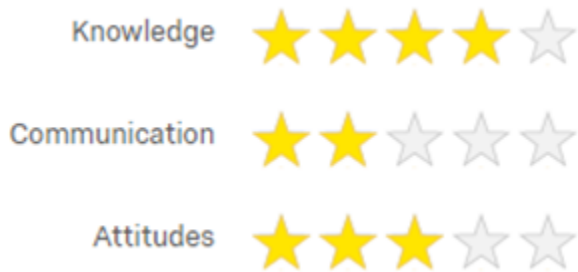


Figure 16 Graphic Sliders Available in Qualtrics



Notes

1. Freyd, M. (1923). The graphic rating scale. *Journal of Educational Psychology*, 14, 83-102.
2. Chyung, S. Y., Swanson, I., Roberts, K., & Hankinson, A. (2018). [Evidence-based survey design: The use of continuous rating scales in surveys](#). *Performance Improvement Journal*, 57(5), 38-48.

3.5 Chapter activity: When to use and not to use specific rating scales?

What are the pros and cons of using the following types of rating scales?

- numerical rating scales as discrete rating scales
- numerical rating scales as continuous rating scales (e.g., web-based sliders)
- verbal descriptor rating scales

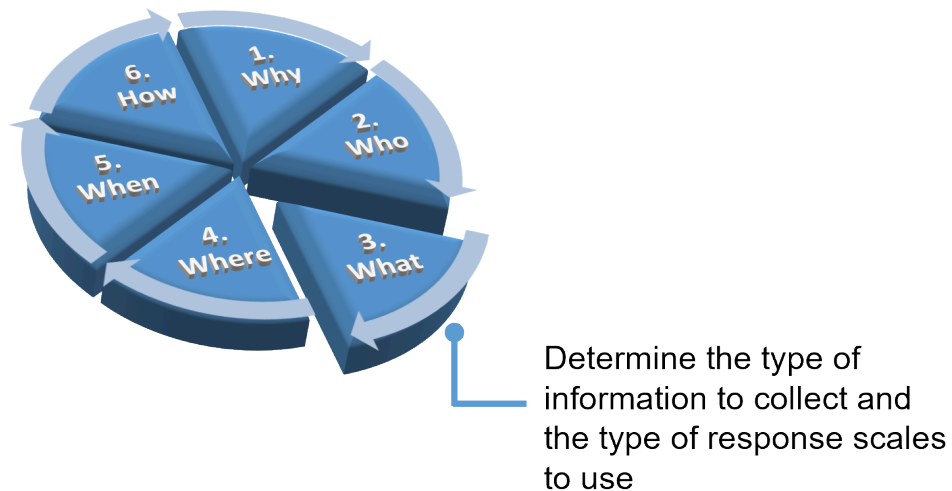
Discuss the use of these three types of rating skills from the following perspectives, providing specific examples:

- From a survey designer's perspective
- From a survey respondent's perspective

4. DESIGN SURVEYS USING EVIDENCE-BASED PRINCIPLES

4.0 Introduction

Within the **What** step, we continue focusing on how to design survey items using appropriate types of response scales.



[A Copy of **Figure 13** *The What Step in the Survey Design and Data Analysis Process*]

As explained earlier, when designing survey items, you need to decide:

- how to request information (as a question or a statement) and
- how to capture the requested information (which response scales to use).

Despite appearing simple, these decisions require a solid understanding of how your choices can affect the quality and accuracy of the data you collect. For example, your survey data may be influenced by decisions such as:

1. Using only positively-worded survey items, or mixing both positively- and negatively-worded items (Chyung, Barkin, & Shamsy, 2018¹)
2. Including or excluding a midpoint in rating scales (Chyung, Roberts, Swanson, & Hankinson, 2017²)
3. Presenting the response options in ascending or descending order (Chyung, Kennedy, & Campbell, 2018³)

There is a substantial body of research on these topics. The following sections provide survey design principles derived from research evidence, which enable you to make informed, evidence-based design decisions.

Learning Objectives

After completing this chapter, you will be able to:

- Explain the effects of using positively and/or negatively-worded survey items
- Explain the effects of including or excluding a midpoint in response scales
- Explains the effects of ordering response options in ascending or descending order

Notes

1. Chyung, S. Y., Barkin, J., & Shamsy, J. (2018). [Evidence-based survey design: The use of negatively-worded items in surveys](#). *Performance Improvement Journal*, 57(3), 16-25.
2. Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). [Evidence-based survey design: The use of a midpoint on the Likert scale](#). *Performance Improvement Journal*, 56(10), 15-23.
3. Chyung, S. Y., Kennedy, M., & Campbell, I. (2018). [Evidence-based survey design: The use of ascending and descending order of Likert-type scale options](#). *Performance Improvement Journal*, 57(9), 9-16.

4.1 Survey questions or statements written in positive or negative wording

When requesting information from respondents, you can use only positively-worded statements (or questions) or a mix of positively- and negatively-worded statements (or questions). (Hereafter, for simplicity we will refer to both statements and questions as *statements*.)

4.1.1 Using only positively- or negatively-worded statements

Survey questionnaires commonly use positively-worded statements. To write one, frame the statement in terms of desirable conditions or outcomes related to the construct being measured.

For example, think about the desirable outcomes of customer service you provide. You want the customers to feel like the service was completed professionally and in a timely manner, and you want them to feel satisfied with the service provided. Using those positive attributes that represent the quality of the service, you would develop the following positively-worded statements:

The service was professionally provided.
The service was completed in a timely manner.
I am satisfied with the service provided.

These statements reflect positive attributes that represent quality customer service.

Likely, if you're measuring the usability of a learning management system, you might use:

The menu items are intuitive.

The navigation structure is user-friendly.

But, should you use only positively-worded statements? Or should you mix positively- and negatively-worded statements?

Some researchers advise against using a mix of positively- and negatively-worded statements in a single survey, for the following reasons:

- Respondents may misread negatively worded items and choose a wrong response, leading to threats to data validity (Weem, Onwuegbuzie, & Lustig, 2003¹)
- Mixing both types of statements can reduce construct validity. For example, in factor analysis, negatively-worded survey items may load onto a separate factor, not because of conceptual differences, but simply due to wording difference (Salazar, 2015²).

4.1.2 When mixing positively and negatively-worded statements

There are situations where it is appropriate to include negatively-worded statements.

1. You may include negatively-worded items when the survey topic itself is negative in nature, such as in surveys measuring depression or anxiety:

I feel hopeless.

I cannot stop worrying about different things.

2. When conducting a cause analysis or analyzing organizational culture, you may find yourself wanting to use a few negatively-worded statements to specifically assess areas of concern or dysfunction.
3. You may also use a balanced scale, which includes an equal number of posi-

tively- and negatively-worded items, which can reduce acquiescence bias (Baumgartner & Steenkamp, 2001³; Cloud & Vaughan, 1970⁴)—the tendency for some respondents to consistently agree (or disagree) with statements regardless of their content (Cronbach, 1942⁵).

4. Including both positively- and negatively-worded items also helps you detect careless responding. For example, when a respondent agrees with a positively-worded statement such as “I like it” and its opposite “I dislike it”, it suggests that the respondent did not pay attention when selecting options. In such cases, you might decide to treat this respondent’s data as invalid and exclude it from analysis. If you decide to include negatively-worded statements in your survey, you will need to choose between wording approaches (see Table 3):
 - Polar opposites – using directly opposite terms (e.g., *Excellent vs. Terrible*)
 - Negated positives – adding “not” to the positive term (e.g., *Excellent vs. Not excellent*)
 - Double negatives – adding “not” to the opposite term (e.g., *Excellent vs. Not terrible*)
 - It is not advisable to use double-negative statements, as they unnecessarily increase cognitive load, cause confusion, and possibly result in incorrect data. Therefore, we will exclude double-negatives from further discussion.

So, which should you use—polar opposites or negated positives? Different researchers seem to have different opinions about this issue (e.g., Schriesheim & Hill, 1981⁶, Weijters & Baumgartner, 2012⁷).

The choice may depend on the survey context, the type of respondents, and the clarity of wording.

Table 3 *Positives, Polar Opposites, Negated Positives, and Double-Negatives*

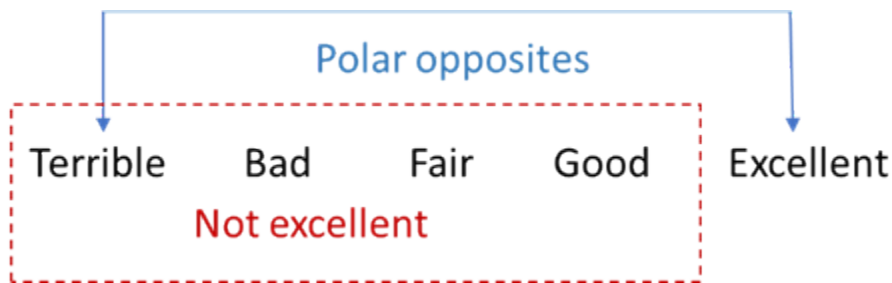
Wording	Ingredient	Type
1. Excellent	A positive word	Positively worded
2. Terrible	The polar opposite of the positive word	Negatively worded
3. Not excellent	“Not” the positive word (a.k.a. negated positive)	Negatively worded
4. Not terrible	“Not” the polar opposite word (a.k.a. double-negative)*	Negatively worded

While choosing between *Terrible* and *Not excellent* as the negative wording, it is important to think about whether the two convey the same meaning.

The service was terrible.
The service was not excellent.

Although both have a negative meaning of *Excellent*, *Not excellent* is not exactly the same as *Terrible*. *Terrible* is the polar opposite of *Excellent* while *Not excellent* simply refers to anything less than *Excellent*, which could include *Good*, *Fair*, *Bad*, and *Terrible* (as illustrated in Figure 17). Therefore, in this case, it would be more precise to use the polar opposite, *Terrible*, as the negative wording.

Figure 17 Polar Opposites vs. Negated Positives



In other cases, especially when using dichotomous descriptors such as *Clear* vs. *Unclear* or *True* vs. *Untrue*, negated positives may convey nearly the same meaning as the polar opposites. In these situations, either statement can effectively serve the purpose:

The objectives are not clear. (negated positive)
The objectives are unclear. (polar opposite)

However, using a polar opposite is not always a better option. In some cases, it may shift the intended meaning of the item. For example, although *Misunderstand* is technically the polar opposite of *Understand*, a more appropriate negative phrasing might be *Feel confused*.

The examples helped me understand the concept.

The examples caused me to misunderstand the concept.

The examples made me feel confused about the concept.

Table 4 presents sample sets of commonly used positive wordings in surveys and their corresponding polar opposites.

Table 4 *Sample Sets of Positive and Polar Opposite Wording*

Positive	Polar opposite	Positive	Polar opposite
Agree	Disagree	Helpful	Unhelpful
Appropriate	Inappropriate	High (quality)	Low (quality)
Comfortable	Uncomfortable	Improve	Worsen
Competent	Incompetent	Interesting	Boring
Correct	Incorrect	Like	Dislike
Easy	Difficult	Likely	Unlikely
Effective	Ineffective	Relevant	Irrelevant
Fast (service)	Slow (service)	Satisfied	Dissatisfied
Friendly	Unfriendly	Supportive	Unsupportive
Good	Bad	Useful	Useless
Happy	Unhappy	Well done	Poorly done

4.1.3 Designing negatively-worded statements with reverse-coding in mind

When using negatively worded statements, it's important to plan for reverse-coding if you intend to compare them with positively worded items. For example, suppose you used the following two survey items to detect careless responses. You would reverse-code the data obtained from the negatively-worded item.

[Positively-worded]

I received excellent service.

- Strongly disagree (1)
- Somewhat disagree (2)
- Neutral (3)
- Somewhat agree (4)
- Strongly agree (5)

[Negatively-worded and reverse-coded]

I received terrible service.

- Strongly disagree (5)
- Somewhat disagree (4)
- Neutral (3)
- Somewhat agree (2)
- Strongly agree (1)

In data from careful respondents, you would expect the reverse-coded values to match or at least closely align with those from the corresponding positively worded items. For example, “I strongly agree that I received excellent service (5)” may correspond with “I strongly disagree that the I received terrible service (5).”

Even when you include negatively-worded statements without the specific intention of screening careless responses, you may still need to reverse-code the responses to those items. Doing so will help avoid confusion during your analysis and reporting.

[Positively-worded]

I am satisfied with the inventive plan.

- Strongly disagree (1)
- Somewhat disagree (2)
- Neutral (3)
- Somewhat agree (4)
- Strongly agree (5)

[Negatively-worded and reverse-coded]

I feel gender bias when it comes to promotion.

- Strongly disagree (5)
- Somewhat disagree (4)
- Neutral (3)
- Somewhat agree (2)
- Strongly agree (1)

In summary, you want to keep in mind the following principles when deciding to use positively- and/or negatively-worded statements:

1. Be aware of possible issues associated with mixing positively- and negatively-worded statements (e.g., threats to construct validity).
2. Develop positively-worded statements that describe the desirable conditions or outcomes of what is being measured.
3. When mixing negatively-worded statements with positively-worded statements,

carefully choose either polar opposites or negated positives, and ensure that the reverse-coded data closely match that of positively-worded statements.

4. Do not use double-negatives.

Notes

1. Weem, G. H., Onwuegbuzie, A. J., Lustig, D. (2003). [Profiles of respondents who respond inconsistently to positively- and negatively-worded items on rating scales](#). *Evaluation & Research in Education*, 17(1), 45-60.
2. Salazar, M. S. (2015). [The dilemma of combining positive and negative items in scales](#). *Psicothema*, 27(2), 192-199.
3. Baumgartner, H., & Steenkamp, J-B. E. M. (2001). [Response styles in marketing research: A cross-national investigation](#). *Journal of Marketing Research*, 38(2), 143-156.
4. Cloud, J., & Vaughan, G. M. (1970). [Using balanced scales to control acquiescence](#). *Sociometry*, 33(2), 193-202.
5. Cronbach, L. J. (1942). [Studies of acquiescence as a factor in the true-false test](#). *Journal of Educational Psychology*, 33(6), 401-415.
6. Schriesheim, C. A., & Hill, K. D. (1981). [Controlling acquiescence response bias by item reversals: The effect on questionnaire validity](#). *Educational and Psychological Measurement*, 41, 1101-1114.
7. Weijters, B., & Baumgartner, H. (2012). [Misresponse to reversed and negated items in surveys: A review](#). *Journal of Marketing Research*, 49(5), 737-747.

4.2 Rating scales with or without a midpoint

When using bipolar verbal descriptor scales such as the Likert scale or Likert-type scales, you have to decide whether to include or exclude a midpoint. Commonly used midpoints are “Neutral” and “Neither [positive] nor [negated].”

When should you include a midpoint, and when should you exclude it? What impact would this decision have on your data? Answering these questions requires a solid understanding of the odd- and even-numbered rating scales, the type of data they produce—ordinal or interval—and how respondents might use or misuse a midpoint.

4.2.1 Odd-numbered scales vs. even-numbered scales

A bipolar rating scale with a midpoint is an odd-numbered scale (e.g., a 5- or 7-point scale).

- Strongly disagree (1)
- Disagree (2)
- Neutral (3)
- Agree (4)
- Strongly agree (5)

- Strongly dissatisfied (1)
- Moderately dissatisfied (2)
- Slightly dissatisfied (3)
- Neither dissatisfied nor satisfied (4)
- Slightly satisfied (5)
- Moderately satisfied (6)
- Strongly satisfied (7)

A bipolar rating scale without a midpoint is an even-numbered scale (e.g., a 4- or 6-point scale).

- Strongly disagree (1)
- Disagree (2)
- Agree (3)
- Strongly agree (4)

- Strongly dissatisfied (1)
- Moderately dissatisfied (2)
- Slightly dissatisfied (3)
- Slightly satisfied (4)
- Moderately satisfied (5)
- Strongly satisfied (6)

Whether odd- or even-numbered, Likert-type scales are generally considered ordinal. This is especially true for even-numbered Likert-type scales without a midpoint because they clearly do not meet the criteria for interval scales. For example, the following distances between two adjacent options are unlikely equal:

- Distance between *Strongly disagree* and *Disagree*
- Distance between *Disagree* and *Agree*
- Distance between *Agree* and *Strongly agree*

However, could odd-numbered Likert-type scales with a midpoint be treated as interval scales assuming the following four distances are reasonably equivalent?

- Distance between *Strongly disagree* and *Disagree*
- Distance between *Disagree* and *Neutral*
- Distance between *Neutral* and *Agree*
- Distance between *Agree* and *Strongly agree*

Let's explore this further in the next section.

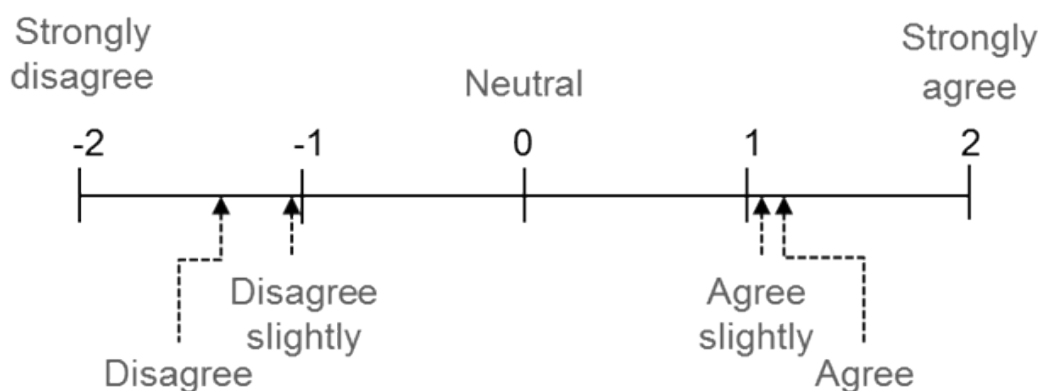
4.2.2 The Likert scale as an ordinal or interval scale

For a 5-point Likert scale to function as an interval scale, survey respondents must perceive the five anchors as being evenly spaced and aligned with numerical values such as -2, -1, 0, +1, and +2, indicating that there is an approximately equal distance between any set of two consecutive points on the scale (Tuckman, 1994¹).

However, research has shown that survey respondents often perceive the 5-point Likert scale as ordinal rather than interval (Worcester & Burns, 1975²). As illustrated in Figure 18, the survey respondents' perceptions of *Disagree* and *Agree* did not align perfectly with numerical values, -1 and 1, respectively. They perceived *Disagree* to be a little bit toward *Strongly disagree*, and *Agree* toward *Strongly agree*.

Interestingly, the researchers found that when modifiers such as *Slightly* were added to *Disagree* and *Agree*, survey respondents perceived *Disagree slightly* and *Agree slightly* as more moderate. This shift in perception brought the anchors closer to the ideal -1 and +1, respectively, helping the scale function more like an interval scale, as shown in Figure 18.

Figure 18 *Changed Perceptions After Adding a Modifier Slightly to Disagree and Agree on the Likert Scale (Based on Worcester & Burns, 1975)*



Based on this research evidence, it is recommended that you use the modified phrases of the second and fourth anchors when designing a 5-point Likert scale: i.e.,

- Strongly disagree
- Slightly disagree (or Somewhat disagree)
- Neutral
- Slightly agree (or Somewhat agree)
- Strongly agree

Why does it matter whether you treat a Likert scale as an ordinal or an interval scale? It matters because the level of measurement affects the types of statistical analyses you can appropriately conduct. Interval data allows more statistical analyses than ordinal data. While ordinal data are typically analyzed using frequencies and percentages, interval data can also be summarized using means and standard deviations.

For example, consider the following survey items using a 4-point Likert scale:

My organization values creativity.

- Strongly disagree
- Disagree
- Agree
- Strongly agree

The results of this ordinal data could be summarized in this way:

- 2% Strongly Disagree
- 5% Disagree
- 25% Agree
- 60% Strongly agree
- 8% Missing data

In contrast, a 5-point Likert scale with modified phrasing may allow you to treat the data as interval (although some researcher may argue that it is still an ordinal scale):

My organization values creativity.

- Strongly disagree

- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

In this case, you might code the data with numeric values (e.g., 1 to 5) and calculate mean and standard deviation of responses. Furthermore, average scores obtained from multiple measures (e.g., pre and post, or Branch A vs. Branch B vs. Branch C) can be compared with each other. For example, if Branch A improved its scores from 3.0 to 4.0, and Branch B improved its scores from 3.0 to 5.0, you can conclude that Branch B improved twice as much as Branch A (because equal distances between response options are assumed). With data obtained from the 4-point Likert scale, this type of a conclusion cannot be made (because equal distances between response options are not assumed).

4.2.3 A midpoint as a dumping ground

A midpoint allows survey respondents to express their neutral feelings towards a presented subject (i.e., feeling equally toward agreement and disagreement). However, research shows that some survey respondents may use the midpoint as a dumping ground, especially when they:

- Do not know enough about the content asked in the survey
- Think their answer depends on other factors
- Are not motivated to respond to the survey
- Want to select a more socially acceptable response, especially when their true opinion is negative

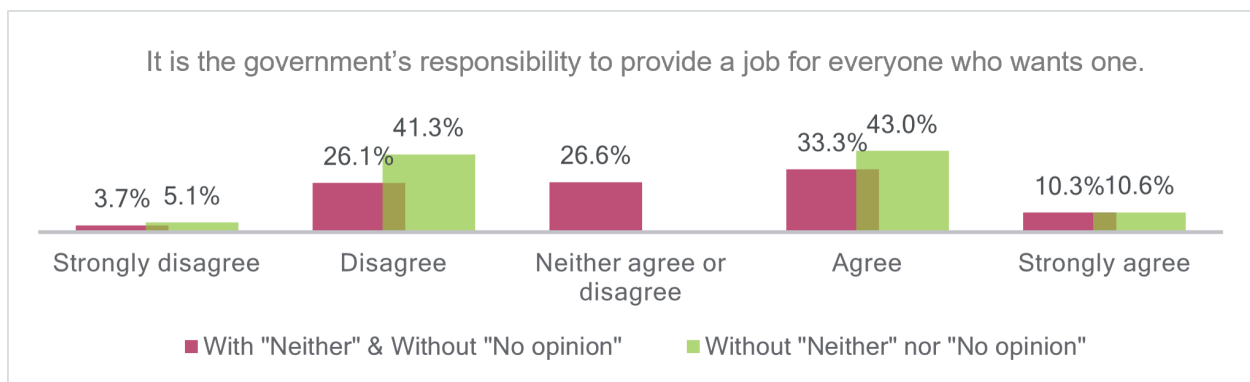
This means including a midpoint can lead to incorrect data. So, should you simply exclude the midpoint?

Not so fast! Omitting a midpoint can create other problems. When a midpoint is absent, respondents who genuinely feel neutral about the survey topic may:

- Be forced to choose a side (negative or positive), resulting in inaccurate data
- Choose not to respond to the survey question, resulting in missing data

In other words, either choice—including or excluding a midpoint—can potentially lead to inaccurate responses. For example, Figure 19 illustrates how response distributions can change when a midpoint was provided vs. not provided (Johns, 2005³). In this public opinion survey, when the *Neither* option was not available, the 26.6% of *Neither* responses may have been re-distributed mostly into *Agree* and *Disagree*. But it's unclear which version reflects true responses—the inflated *Agree* and *Disagree* responses when *Neutral* is not allowed or the decreased *Agree* and *Disagree* responses when *Neutral* is allowed. So, what should you do?

Figure 19 Data Distribution Difference When a Midpoint is Available and When It Is Not (Based on Johns, 2005, Table 3, p. 249)



4.2.4 Add I don't know, not applicable, or no opinion

To address the potential problems described in the previous section—whether or not to include a midpoint—you can offer an additional option such as *I don't know*, *Not applicable*, or *No opinion*. Not all survey items would need these additional options so you should use your best judgment.

I Don't Know – Let's say, you are aware that some trainees missed one or more sessions of a multi-day workshop. If your survey asks about a specific activity they may not have experienced, it would be appropriate to include an *I don't know* option. You could still assign 1 through 5 to the five Likert scale options, but exclude *I don't know* responses when calculating average scores.

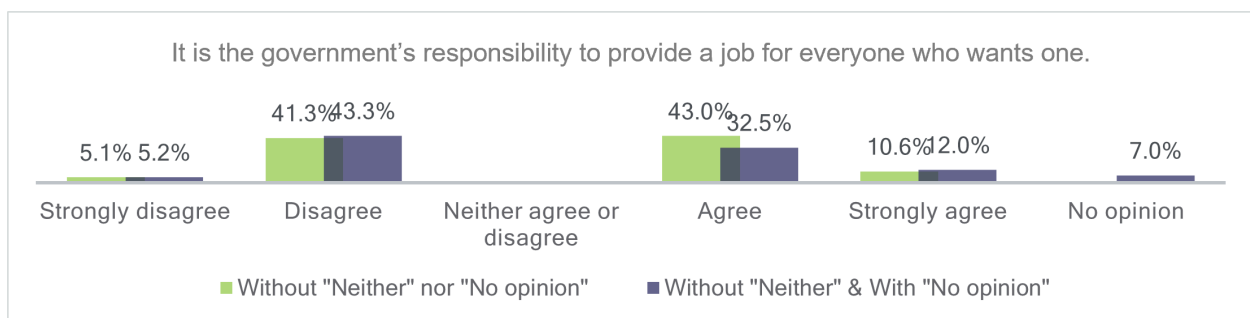
I found the role-play practice very useful.

- Strongly disagree (1)
- Somewhat disagree (2)
- Neutral (3)
- Somewhat agree (4)
- Strongly agree (5)
- I don't know (0)

Not Applicable – You might add a *Not applicable* option when survey items do not apply to all respondents. For example, if you are using a standardized survey questionnaire for both face-to-face and e-learning courses, some questions that apply to face-to-face training do not apply to e-learning courses, and vice versa. Rather than forcing a response that could introduce inaccurate data, *Not applicable* offers a clear opt-out.

No Opinion – In some cases, you might consider offering *No opinion*, particularly in public opinion or customer satisfaction surveys. For example, in the public opinion survey described earlier (John, 2005), adding a *No opinion* option to a 4-point Likert scale redistributed some of the *Neutral* response from the 5-point version. It appears that the 26.6% who originally chose *Neither* on the 5-point scale (Figure 19) were reallocated across *Strongly disagree* (0.1% increase), *Disagree* (2% increase), *Agree* (10.5% decrease), *Strongly agree* (1.4% increase), and *No opinion* (7%) on the revised scale (Figure 20). It is curious to see if some of the respondents who selected *Agree* when *No opinion* was not available (suggesting the social desirability bias) now selected *No opinion*.

Figure 20 Data Distribution Difference When No Opinion Is Available and When It Is Not (Based on Johns, 2005, Table 3, p. 249)



In conclusion, both including and excluding a midpoint on a rating scale come with advantages and drawbacks. Rather than asking *if* you should include a midpoint, it is more useful to consider *when* it is appropriate to do so. You may choose to include a midpoint if you want to allow respondents to express their neutral feelings or if you plan to treat the Likert scale as an interval scale in your analysis. Conversely, you may want to exclude the midpoint if you have a reason to believe that respondents might use the midpoint as a dumping ground. Regardless of your decision, it is important to use the same type of scale consistently across related survey items in your questionnaire, especially if you plan to aggregate responses (e.g., by calculating an average score). Additionally, consider including options such as *I don't know* or *Not applicable* when appropriate, to further improve the accuracy and interpretability of your data.

Notes

1. Tuckman, B. (1994). *Conducting educational research* (4th ed.). Harcourt Brace College Publishers.
2. Worcester, R. M., & Burns, T. R. (1975). A statistical examination of the relative precision of verbal scales. *Journal of the Market Research Society*, 17(3), 181-197.
3. Johns, R. (2005). [One size doesn't fit all: Selecting response scales for attitude items](#). *Journal of Elections, Public Opinion and Parties*, 15(2), 237-264.

4.3 Ascending vs. descending order of response options

Another important design decision is whether to present response options in ascending or descending order because “the order in which response alternatives are presented to respondents may have a significant influence upon their selections” (Krosnick & Alwin, 1987, p. 202¹). This psychological phenomenon found in surveys is known as the *response order effect*.

4.3.1 Survey data impacted by various response order effects

Two common types of response order effects are the recency and primacy effects (Holbrook, 2008a²).

- The *recency effect* typically occurs in oral surveys. When options are presented orally, respondents tend to select one of the options that they heard most recently. This phenomenon can be observed during an order at a restaurant when you listen to several salad dressing options and end up choosing the one that you heard toward the end of the list.
- The *primacy effect* occurs in visual surveys. When options are presented visually, respondents tend to select one of the options that they see at the beginning of the list. You may have observed this phenomenon when you administer surveys on paper or on the web where respondents see a list of options on the rating scale and select one of the options they see first. When text is written from left to right horizontally, they tend to select one on the left side of the list options, which is referred to as a *left-side selection bias*. This brings up an issue when you use an ascending order or descending order of the Likert response scale options.

Survey respondents may also be influenced by other biases:

- Respondents tend to *agree* more than they disagree with the statement provided—a.k.a., *yea-saying bias*. One explanation for acquiescence bias is the social

norm to be polite (Holbrook, 2008b³).

- This is explained as *social desirability bias*, the tendency for respondents to select among the options that they think are more socially acceptable or desirable, instead of a response that reflects their true feelings/opinion (Callegaro, 2008⁴).
- People also tend to want to minimize their psychological costs by selecting solutions that are satisfactory or acceptable enough, which is known as a *satisficing* principle (Simon, 1957, as cited in Krosnick & Alwin, 1987⁵).

4.3.2 Survey data impacted by psychological phenomena and environmental factors

The primacy effect, left-side selection bias, acquiescence bias, social desirability bias, and satisficing are psychological phenomena that survey respondents may exhibit when using response scales. As a survey designer, it's important to consider how your choice of ascending or descending order of the Likert options—an environmental factor—might interact with the psychological tendencies.

In an ascending order (when presented horizontally), positive response options are presented on the right side of the scale:



In a descending order (when presented horizontally), positive response options are presented on the left side of the scale:



Compared to the ascending order, the descending order is more likely influenced by a combination of the primacy effect, left-side selection bias, acquiescence bias (yea-saying bias), and social desirability bias. Respondents tend to perceive that selecting one of the positive responses (e.g., *Strongly agree* or *Somewhat agree*) is more socially desirable and they see the positive options first on the left side. Therefore, they tend to select a positive option, potentially leading to inflated positive responses.

4.3.3 Inflated survey data from survey items with descending-ordered response scales

In a study by Nicholls, Orr, Okubo, and Loftus (2006)⁶ involving 292 college students in Australia, participants were assigned to complete a survey with a 5-point Likert scale presented in one of two orders:

In ascending order:

○ Definitely disagree	○ Mostly disagree	○ Neither agree nor disagree	○ Mostly agree	○ Definitely agree
---------------------------------	-----------------------------	--	--------------------------	------------------------------

In descending order:

○ Definitely agree	○ Mostly agree	○ Neither agree nor disagree	○ Mostly disagree	○ Definitely disagree
------------------------------	--------------------------	--	-----------------------------	---------------------------------

The group that completed the survey with a descending-ordered scale produced significantly higher scores than the group with the ascending-ordered scale.

A similar pattern was found in a study by Hartley and Betts (2010)⁷. Participants were randomly assigned to one of four conditions created by using different directions of worded labels (Clear to Unclear, or Unclear to Clear) and different directions of numeric ratings (10 to 0, vs. 0 to 10):

	Label	Numeric Ratings	Label
	Clear	10 9 8 7 6 5 4 3 2 1 0	Unclear
	Clear	0 1 2 3 4 5 6 7 8 9 10	Unclear
	Unclear	10 9 8 7 6 5 4 3 2 1 0	Clear
	Unclear	0 1 2 3 4 5 6 7 8 9 10	Clear

The first format—the descending-ordered scale that began with a positively worded label (Clear) and a highest numerical value (10)—resulted in significantly higher ratings in comparison to the other three response options.

A simple solution to the issue of getting inflated data obtained from descending-ordered scales is to present response scales in ascending order.

Notes

1. Krosnick, J., & Alwin, D. (1987). [An evaluation of a cognitive theory of response-order effects in survey measurement](#). *The Public Opinion Quarterly*, 51(2), 201-219.
2. Holbrook, A. (2008a). [Acquiescence response bias](#). In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (Vol. 1, pp. 3-4). SAGE.
3. Holbrook, A. (2008b). [Response order effects](#). In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (Vol. 2, pp. 754-756). SAGE.
4. Callegaro, M. (2008). [Social desirability bias](#). In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (Vol. 2, pp. 825-826). SAGE.
5. Krosnick, J. A., & Alwin, D. F. (1987). [An evaluation of a cognitive theory of response-order effects in survey measurement](#). *Public Opinion Quarterly*, 51(2), 201-219.
6. Nicholls, M. E., Orr, C. A., Okubo, M., & Loftus, A. (2006). [Satisfaction guaranteed: The effect of spatial biases on responses to Likert scales](#). *Psychological Science*, 17(12), 1027-1028.
7. Hartley, J. & Betts, L. (2010). [Four layouts and a finding: The effects of changes in the order of the verbal labels and numerical values on Likert-type scales](#). *International Journal of Social Research Methodology*, 13(1), 17-27.

4.4 Chapter activity: How would you revise the current survey questionnaire?

You have recently started a new role as a training manager at an organization. You found out that your predecessor had drafted out a web-based survey questionnaire intended to evaluate the effectiveness of a mentoring training program provided to employees, as shown in Exhibit 4.

You are applying your knowledge about evidence-based survey design principles to the survey items to see what (if any) revisions should be made to create a final version of the survey questionnaire. You hope to use an average score of the 10 survey items to represent the quality of the program. Identify the parts that you want to change and describe the reasons why. Create a revised version.

Exhibit 4 An Existing Survey Questionnaire To Be Revised

Post-Training Evaluation

For each question, please select the option that best represents your opinion about the mentoring training program that you have just completed. This is an anonymous survey. Your input will be used to improve the quality of the program for future participants. Thank you!

1. The role-play helped me better understand the new mentoring techniques.

- Strongly disagree
- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

2. The video presentation was distracting.

- Strongly disagree
- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

3. The examples were irrelevant to my job.

- Strongly disagree
- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

4. The activities were appropriate for my current expertise on mentoring.

- Strongly agree
- Somewhat agree
- Neutral
- Somewhat disagree
- Strongly disagree

5. I will be able to interact with my employees as a more capable mentor.

- True
- Somewhat true
- Untrue
- I don't know

6. I am still not confident in using the new mentoring skills.

- Strongly agree
- Somewhat agree
- Somewhat disagree
- Strongly disagree

7. Are you satisfied with the quality of this program?

- Strongly agree
- Somewhat agree
- Somewhat disagree
- Strongly disagree

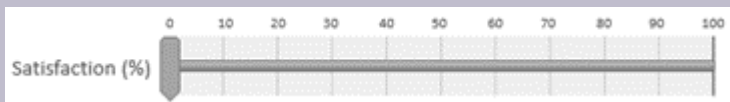
8. I would recommend this program to others.

- Not at all
- Maybe
- Definitely

9. How satisfied are you with the overall quality of the mentoring program?



10. How satisfied are you with the overall quality of the presenter?



5. DEVELOP SURVEYS FOR DIFFERENT PURPOSES

5.0 Introduction

As emphasized earlier in Section 2.4.5 *Project the type of data you need before designing survey items*, it is critical to determine the type of data you need before designing your survey questionnaire.

You conduct surveys to gather information you do not already know. In your L&D, HRD, OD, or PI practice, you may use surveys to solicit information at various stages of a project:

- **Before** implementing a program (e.g., during a performance analysis),
- **During** the development of a program (e.g., as part of an instructional design project), or
- **After** a program has been implemented (e.g., during an evaluation project).

Each of these purposes calls for different types of survey questions and corresponding response scales. As discussed in the previous chapter, factual information may require different response scales than questions that measure perceived quality or value.

Learning Objectives

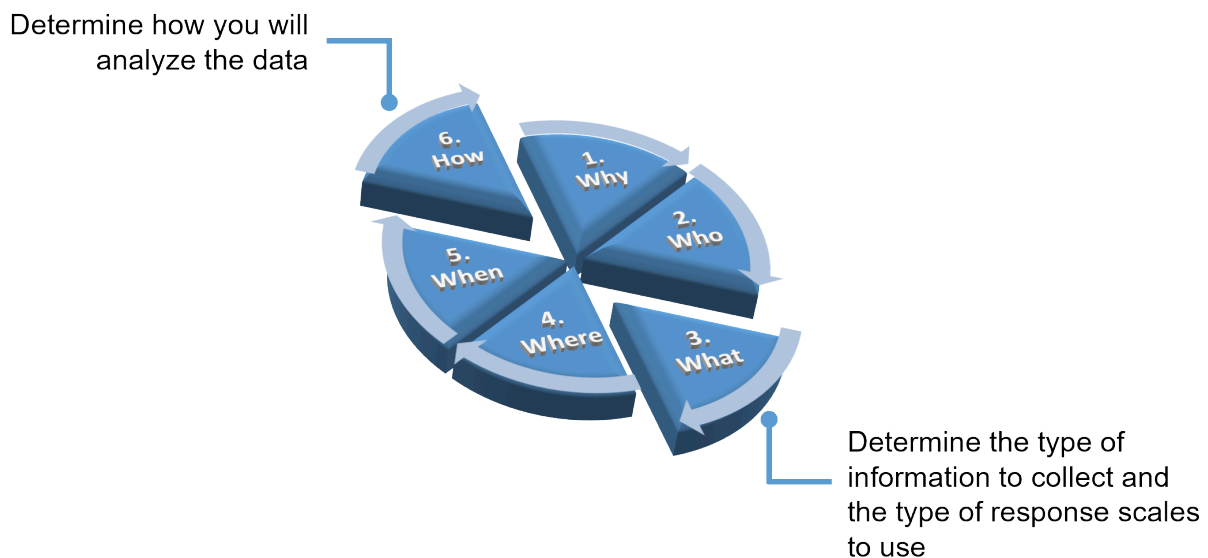
After completing this chapter, you will be able to:

- Explain the importance of identifying the purpose of a survey before designing it
- Recognize survey types used to collect data for evaluation projects
- Recognize survey types used to collect data for during performance analysis (needs assessment) projects
- Recognize survey types used to collect data for instructional design projects

5.1 Purpose of the survey and knowledge to be gained from the survey

Before designing your survey questionnaire, you must first clearly determine the purpose of the survey—this aligns with the first step in the survey design and data analysis process. Equally important is to keep the end in mind by thinking ahead to the final step of the process—how you will analyze your data and report it. This forward-thinking approach requires you to identify what kind of knowledge you want to gain from the survey and how you intend to present it. In other words, while you are working on the **What** step (designing survey questions and response scales), you should simultaneously consider the **How** step (data analysis and reporting), as illustrated in Figure 21.

Figure 21 *Project the How Step While Working on the What Step*



5.1.1 Developing the right type of survey item for the

intended report: Example 1

Consider the three scenarios presented in Table 5 where the purpose of conducting the survey is the same, but the way you plan to report the findings differs:

Table 5 *Designing Survey Items to Generate Specific Knowledge to Be Gained*

Purpose of the survey: To find out the effectiveness of the job aid

Design Factors	Scenario 1	Scenario 2	Scenario 3
Specific knowledge to be gained from the survey (how to write the report)	How frequently technicians have been using the job aid: e.g., N (X%) technicians use the job aid at least weekly.	If technicians find the job aid helpful or not: e.g., N (X%) technicians have found the job aid helpful.	How the job aid has helped the technicians: e.g., N (X%) technicians have found the job aid helpful in reducing errors.
Survey item to be used	Q1. How often did you use the job aid in the last month? <ul style="list-style-type: none"> ○ Once a day ○ Once a week ○ Once a month ○ Never 	Q2. Does the job aid assist you in your job performance? <ul style="list-style-type: none"> ○ Yes ○ No ○ I don't know because I never used it Please describe how the job aid helped or did not help:	Q3. How helpful is the job aid in assisting you in your job performance? Check all that apply. <ul style="list-style-type: none"> ○ Helps me complete my job tasks faster ○ Helps me complete my job tasks without asking for help ○ Helps me reduce errors ○ Other. Describe:
Data analysis	Frequency/percentage No average	Frequency/percentage No average	Frequency/percentage No average

In Scenario 1, the survey items is designed to capture factual, non-judgmental information. Reporting the frequency of job aid use does not directly indicate the quality of the job aid itself. For example, some technicians may not have used the job aid because they forgot about it or misplaced it.

In contrast, the survey items in Scenarios 2 and 3 aim to measure perceived quality

or value about the job aid. However, due to their different designs, these two survey items produce different types of knowledge about the quality/value of the job aid.

- If your primary goal is simply to determine *if* the job aid was helpful or not, “Q2 Does the job aid assist you in your job performance?” would be sufficient.
- If you want to learn *how* the job aid helped the technicians, however, Q2 will not provide enough detail. Instead, you should use a survey item such as “Q3 How helpful is the job aid in assisting you in your job performance?” which can reveal how technicians used the job aid and which aspects they found useful.

When you use Q3, you can still infer whether the job aid was helpful based on responses—so you likely would not need to include both Q2 and Q3 in your survey.

However, even Q3 may not give you a complete picture of the job aid effectiveness. It does not tell you *how much* the job aid has contributed to achieving specific outcomes. To gain this level of detail, you may use Q4 in Exhibit 5. Q4 is designed to generate the most comprehensive information about the effectiveness of the job aid usage, compared to Q1, Q2, and Q3.

Exhibit 5 A Sample Set of Survey Items That Measure How Much a Job Aid Helped

Q4. Please select an option that indicates how much the job aid has had an impact on how you do your job

Question	Very much negatively	Somewhat negatively	No change	Somewhat positively	Very much positively	Never used it
Complete job tasks faster	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Complete job tasks without asking for help	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reduce errors	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

When analyzing the data obtained from Q4, you may report the frequencies/percentages of individual opinions, for example:

Twenty and 25 technicians (38% and 47%) reported the job aid made a very positive or somewhat positive impact on completing their job tasks faster, respectively.

Three technicians (6%) indicated no impact on their job performance when using the job aid.

Three and two technicians (6% and 4%) stated that the job aid had a very negative or somewhat negative impact on completing their job tasks faster, respectively.

Alternatively, if you decide to treat this response scale as an interval scale, you might assign numerical values -2, -1, 0, +1, +2 (or 1, 2, 3, 4, 5) to the five response options and calculate an average score (excluding *Never used it*), as shown below:

- Very much negatively (-2)
- Somewhat negatively (-1)
- No change (0)
- Somewhat positively (1)
- Very much positively (2)
- Never used it (excluded)

This approach allows you to summarize the overall impact of the job aid with a single score.

For example, compare the two hypothetical datasets presented in Table 6, Table 7, and the bar graphs (Figure 22 and Figure 23).

- The average score of 0.6 presented in Table 6 means a small positive impact, falling between *no change* and *somewhat positively*.
- The average score of 1.5 presented in Table 7 means a substantial positive impact, falling between *somewhat positively* and *very much positively*.

This type of information would not be possible with Q2 or Q3 type survey items, which lack the precision necessary to evaluate the degree of impact.

Table 6 Hypothetical Data Results #1

Question	Very much negatively (-2)	Somewhat negatively (-1)	No change (0)	Somewhat positively (+1)	Very much positively (+2)	Never used it (excluded)
Frequency	2	3	3	5	7	2

Average score (while excluding “Never used it”) = $[(-2 \times 2) + (-1 \times 3) + (0 \times 3) + (1 \times 5) + (2 \times 7)] / (2 + 3 + 3 + 5 + 7) = 12 / 20 = 0.6$

Figure 22 A Bar Graph of Hypothetical Data Results #1

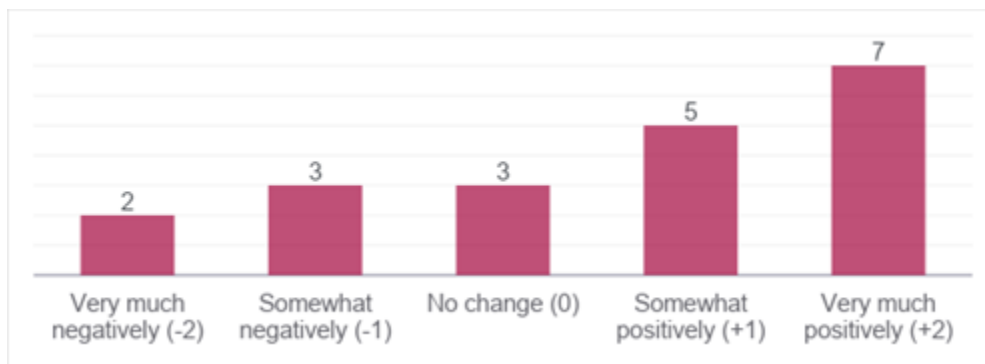
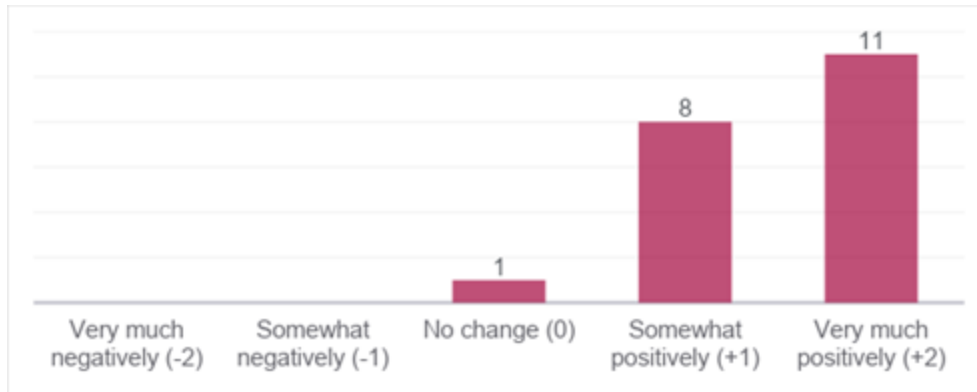


Table 7 Hypothetical Data Results #2

Question	Very much negatively (-2)	Somewhat negatively (-1)	No change (0)	Somewhat positively (+1)	Very much positively (+2)	Never used it (excluded)
Frequency	0	0	1	8	11	3

Average score (while excluding “Never used it”) = $[(-2 \times 0) + (-1 \times 0) + (0 \times 1) + (1 \times 8) + (2 \times 11)] / (0 + 0 + 1 + 8 + 11) = 30 / 20 = 1.5$

Figure 23 A Bar Graph of Hypothetical Data Results #2



Although it may be tempting to use survey items like Q3 or Q4 to collect more detailed information, it is not always feasible to know exactly how to ask respondents to rate the quality or value of the job aid—particularly in terms of exactly how it helped them. In such cases, consider the following strategies:

1. Interview the job aid developer to understand what the job aid is supposed to accomplish and use this insight to create survey items aligned with the developer's original intent.
2. Interview a small group of technicians (i.e., the actual users of the job aid) to gather feedback on the benefits of using the job aid and develop survey items based on the actual users' input.

Keep in mind, using a detailed survey item like Q4 can be overkill if your goal is simply to determine whether technicians found the job aid helpful or not. In that case, Q2 would be sufficient. As discussed earlier, if you are using a survey item like Q4, it is not necessary to include Q3.

To summarize, in your survey questionnaire, consider combining items in one of the following ways:

- Q1 (capturing factual usage) + Q2 (measuring perceived helpfulness)
- Q1 (capturing factual usage) + Q3 (measuring how the aid was used)
- Q1 (capturing factual usage) + Q4 (measuring degree of impact)

5.1.2 Developing the right type of survey item for the intended report: Example 2

Let's review another example. Suppose you are assessing new employees' job readiness after they have completed an onboarding program, and you've created a set of survey items using a 7-point scale. In addition, you also plan to report on whether new employees' job readiness levels correlate with their prior work experience, and you intend to use Pearson's r correlation coefficient, which requires interval or ratio-level data.

To conduct this analysis, you need to include an item that captures each respondent's prior work experience. Between the following two options, which one would you use?

Q1. How many years of prior work experience related to your job responsibilities do you have?

- 1-2 years of experience
- 3-5 years of experience
- 6-9 years of experience
- 10 or more years of experience

Q2. How many years of prior work experience related to your job responsibilities do you have?

Approximately _____ years

Between the two, Q2 is the better choice for your intended analysis. It captures continuous data, allowing you to calculate Pearson's r . By contrast, Q1 provides ordinal data which is not appropriate for Pearson's r . If you happen to use Q1, you may use alternative correlation methods, such as Spearman's rank-order correlation.

This example again reinforces the importance of keeping the end in mind—if you plan to perform specific statistical analyses, your survey items must be designed to yield data that meet the assumptions of those methods.

5.2 Sample surveys used during evaluation

When conducting program evaluations, you often need to demonstrate improvements in outcomes. These improvements can be presented as differences in the outcomes before and after the program. To achieve this, you may plan in advance to collect both pre- and post-program data, or you may choose to retrospectively assess pre-program data at the time of post-program data collection.

5.2.1 Pre-post surveys

Suppose you are evaluating a caregiver training program, designed for family members and friends who care for individuals with dementia and other conditions involving cognitive decline. The primary goal of the program is to help caregivers improve resilience. You intend to investigate short-term and long-term outcomes of the program by conducting surveys along with other data collection methods. The survey questionnaire could contain items such as those shown in Exhibit 6.

Exhibit 6 Sample Survey Items to Evaluate a Caregiver Training Program (Adapted from Wagnild & Young, 1993¹)

Please indicate your level of agreement with each of the following statements about the program by selecting one of the numbers from 1 to 7.

Question	Strongly disagree						Strongly agree
	1	2	3	4	5	6	7
I follow through with the plans I make.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am able to depend on myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have enough energy for daily activities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can find my way out of a difficult situation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can handle many things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can depend on myself more than anyone else.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am determined.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I keep interested in things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

You may plan to use repeated measures with the same survey questionnaire to indicate changes in caregiver resilience over time. You conduct a pre-survey at the beginning of the program and a post-survey at the end of the program. The changes between the pre- and post-survey data would indicate short-term outcomes. You may also follow up with program completers several weeks or months after the program ends and conduct a follow-up post survey. This may show long-term outcomes. In these types of repeated measures, you have to collect identifiable data (e.g., names, IDs, or other unique identifiers) with each survey administration, to match each participant’s pre-, post-, and follow-up post-data for analysis.

What	Pre-survey	Post-survey	Follow-up post survey
When	Before or in the beginning of the program	At the end of the program	Several weeks or months after the program

With the average scores collected from the repeated measures, you can present changes in caregiver resilience from the pre-survey to the post-survey, and the follow-up-post-survey using a line graph, as illustrated in Figure 24.

Figure 24 A Sample Line Graph Illustrating the Changed Levels among Repeated Measures



5.2.2 Retrospective surveys

Suppose you are evaluating the quality of a leadership development program that has been in operation for the last couple of years. The program targets first-level managers, aiming to enhance their leadership behaviors. You intend to investigate both short-term and long-term outcomes:

- Short-term outcomes: Whether the first-level managers are satisfied with the program and have acquired sufficient leadership knowledge and skills
- Long-term outcomes: How the managers' leadership styles have impacted their employees

Because the evaluation was initiated after the program had already begun, you are unable to collect pre-survey data. Therefore, you decide to use a retrospective survey.

To measure short-term outcomes, you've decided to investigate two dimensions:

1. Participants' satisfaction with the program
2. Participants' knowledge and skills in leadership

Thus, the survey questionnaire may contain survey items in two sections, as shown in Exhibit 7, and it can be administered with the participants (first-level managers) at the end of the program. Section 1 measures satisfaction, and Section 2 asks participants to rate their knowledge and skills before and after the program. Unlike pre-post surveys, which require you to collect identifiable data such as names or IDs to match pre-

and post-data, retrospective surveys collect both sets of data in one administration, allowing for anonymous responses, since no matching of pre- and post-responses is required. However, if you plan to conduct a follow-up post survey, you will still need to collect identifiable information.

To measure long-term outcomes of how managers' leadership styles have impacted their employees, several weeks or months after the program, you should survey the employees supervised by the program completers (first-level managers). For consistency, you may want to use survey items similar to those used in Section 2 of the managers' post-survey (Exhibit 7) that was administered with the program completers (first-level managers). The employee survey is shown in Exhibit 8.

Although you may use a retrospective survey asking employees to rate their manager's leadership skills before and after the manager completed the program, it can be difficult for employees to recall their managers' program participation timeframe accurately. In such cases, it may be better to ask employees to rate their manager's current leadership skills only.

Exhibit 7 *A Sample Survey Questionnaire Including Retrospective Survey Items to Measure Changed Outcomes*

Section 1. Please indicate your level of agreement with each of the following statements about the program by selecting one of the options.

Question	Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree	Don't know
The program provided information useful to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program content was interesting to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program was designed well.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program helped me become more knowledgeable about the impact of different leadership styles.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program was worth the time I spent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Section 2. Please indicate the levels of your leadership skills before and after you completed the program by selecting one of the numbers: 1 = Poor, 2 = Mediocre, 3 = Average, 4 = Good, 5 = Excellent.

Question	Before you completed the program					After you completed the program				
	1	2	3	4	5	1	2	3	4	5
How to:										
Delegate tasks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Motivate others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Provide feedback	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Resolve conflicts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be a good listener	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gain trust	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be fair	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Communicate clearly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be decisive when needed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be both confident and humble	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Exhibit 8 A Sample Survey Questionnaire to Measure Employees' Ratings on Their Managers' Leadership Skills Before and After Program

Please rate your manager's current leadership skills by selecting one of the numbers: 1 = Poor, 2 = Mediocre, 3 = Average, 4 = Good, 5 = Excellent.

How to:	Poor 1	Mediocre 2	Average 3	Good 4	Excellent 5
Delegate tasks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Motivate others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Provide feedback	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Resolve conflicts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be a good listener	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gain trust	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be fair	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Communicate clearly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be decisive when needed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Be both confident and humble	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

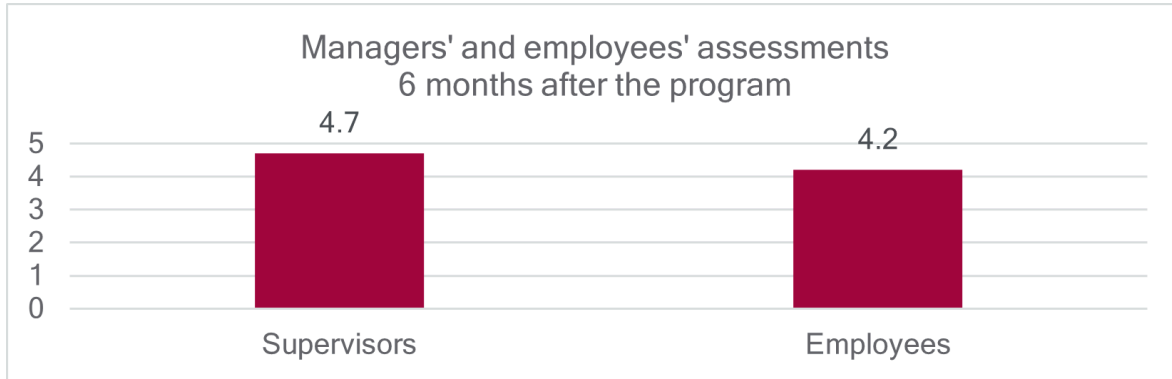
Then, the managers' pre-, post-, and follow-up-post-data can be compared with each other, as shown in Figure 25.

Figure 25 A Bar Graph Comparing Before, After, and 6 Months After Program



A similar follow-up-post survey can be administered to both program completers (first-level managers) and their employees. Then, their data can be compared to determine whether the managers' and employees' assessments are similar or different (Figure 26). Also see Appendix A for conducting a t-test to compare independent groups.

Figure 26 A Bar Graph Comparing Managers' and Employees' Assessments



5.2.3 Surveys to be used in four-level training evaluations

If you are conducting a training evaluation, you may apply the four-level training evaluation framework (Kirkpatrick, 2004²) to assess the following four outcomes:

1. Reaction—learners' reaction to the training program at the end of the program
2. Learning—the knowledge, skills, or attitudes gained as a result of the training program
3. Behavior—their changed behavior on the job after the training program
4. Results—the training program's contribution to overall organizational results

The four-level training evaluation framework is commonly associated with Donald Kirkpatrick as the developer. However, in his early publication in 1956, Kirkpatrick cited the four steps of conducting training evaluation listed by Raymond Katzell, an industrial-organizational psychologist (Kirkpatrick, 1956³), which indicates that the originator of the four steps of training evaluation was Raymond Katzell (Smith, 2008⁴; Thalheimer, 2018⁵). Therefore, in this book, we will recognize the four levels as the Katzell-based Kirkpatrick's four-level evaluation framework (Smith, 2008).

When conducting evaluations using this framework, you can employ various data collection methods (see Table 8). Surveys can be used in all four levels of training evaluation—however, it is important to note that surveys are not a direct measure for assessing learning, behavior, and results.

Let's explore how surveys can be used within each of the four levels of training evaluation.

Table 8 *Examples of Various Data Collection Methods for Four Levels of Training Evaluation (Table 38, pp, 122-123, Chyung, 2019⁶)*

Method	1. Reaction	2. Learning	3. Behavioral change	4. Results
Self-administered surveys	<p>Survey trainees' reaction (a.k.a. smiley sheet) at the end of a program</p> <p>Survey trainers about their experience with trainees</p>	<p>Survey trainees to self-assess their learning</p> <p>Survey trainers about their experience with trainees</p>	<p>Survey trainees to self-assess their behavioral change</p> <p>Survey trainees' co-workers or supervisors about their experience with trainees</p>	<p>Survey trainees and other stakeholders, and have them self-assess organizational results</p> <p>Survey customers about quality of organizational service provided to them</p>
Interviews and focus groups	<p>Interview trainees individually or as a group</p> <p>Interview trainers about their experience with trainees</p>	<p>Interview trainees to self-assess their learning</p> <p>Interview trainers, co-workers, or supervisors about their experience with trainees</p>	<p>Interview trainees to self-assess their behavioral change</p> <p>Interview trainees' co-workers or supervisors about their experience with trainees</p>	<p>Interview trainees and other stakeholders, and have them self-assess organizational results</p> <p>Interview customers about quality of organizational service provided to them</p>
Observations (including performance tests via observations)	<p>Observe trainees' reaction during the program (participant observation)</p>	<p>Observe trainees' new knowledge and skills during the learning process</p>	<p>Observe trainees' on-the-job behaviors (pre vs. post)</p> <p>Compare trainees' on-the-job behaviors to those of people who did not complete the training</p>	<p>Observe organizational results such as customers' reaction</p>

Method	1. Reaction	2. Learning	3. Behavioral change	4. Results
Extant data reviews	Review existing data that documented trainees' thoughts about the program quality	Review existing data that documented trainees' learning outcomes	Review existing data that documented trainees' job performance	Review existing data that documented organizational results Compare organizational results obtained from the group that participated in training and those obtained from a group that did not
Written tests	N/A	·Test trainees' new knowledge at the end of training, compared to pre-training knowledge Compare test results obtained from trainees to the test results obtained from people who did not complete the training	N/A	N/A

Level 1 Reaction

Surveys are most commonly used to measure trainees' satisfaction with a training program, capturing participants' immediate impressions and emotional responses to the training experience. Depending on the type of the training program and its context, the survey questionnaires would be designed to measure different aspects of the training program, such as:

- The training enrollment process
- The instructional content covered
- The instructional strategies used
- The usefulness of the learned information
- The trainer's knowledge
- The trainer's demeanor

- The training environment (location, classroom, duration, etc.)

Level 2 Learning

A direct way of measuring learning outcomes is through written or performance tests. Surveys are an indirect measure of learning outcomes. Surveys measure learning outcomes only in terms of:

- the learners' confidence in using the knowledge and skills learned
- their willingness or intention to apply the learned knowledge and skills in the workplace

This type of self-assessed information does not necessarily reflect the learners' actual competence or performance capacity. However, when actual knowledge testing is not practical or possible, surveys can still be used to estimate perceived learning outcomes. To improve the reliability of such data, consider triangulating it with other sources such as their supervisors' assessments or peer feedback.

For example, suppose you are evaluating an e-learning program on lab safety. You plan to measure level 1 reaction and level 2 learning (confidence in learned knowledge and skills). Exhibit 9 is a sample survey questionnaire to be used, in which Questions 1 through 6 measure level 1 reaction and Questions 7 through 12 measure level 2 confidence in learned knowledge and skills.

Exhibit 9 A Sample Survey Questionnaire That Measures Level 1 Reaction and Level 2 Confidence in Learned Knowledge and Skills

Please indicate your level of agreement with each of the following statements about the program and your learning by selecting one of the options.

About the program	Very untrue 1	2	3	4	Very true 5	Not sure
It was easy to log in and start the program.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to navigate through the program.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program covered the topic of lab safety sufficiently.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program kept me engaged.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program presented safety information in interesting ways.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The program used safety examples relevant to my work situation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

About your learning	Very untrue 1	2	3	4	Very true 5	Not sure
I am fully aware of safety policy and guidelines.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I know how to use safety gear when in the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I know how to use the safety checklist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I know how to operate the fume hood.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I know how to dispose of biohazardous waste.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I know how to complete proper hand-washing upon contact with chemicals and before leaving the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Level 3 Behavior

The most direct way to measure trainees' changed behaviors on the job is through

direction observation of performance or by reviewing concrete records of performance (e.g., computer log files, recorded customer service calls).

When direct observation is not feasible, an alternative method is to use surveys asking trainees to self-assess their behavioral changes. This survey data can then be triangulated with another source such as their supervisors' or customers' assessments.

Level 4 Results

Training programs intend to contribute to organizational results, but their actual contribution may vary. This includes improved customer retention, increased profits, increased membership, and improved customer satisfaction. Recall that you conduct surveys when the information cannot be directly observed and recorded. Some of the organizational results such as customer retention, increased profits, and increased membership can be recorded and found in documents. However, perceptions of impact—such as improved customer satisfaction may or improved team safety culture—may not be directly observable and must be measured via surveys.

Using the same lab safety e-learning program scenario earlier, you may survey the employees a few weeks or months after they completed training to measure how well they are following the safety guidelines and what impact they think the training has made on the overall safety-related outcomes. You may also conduct a similar survey with the supervisors of the employees.

Exhibit 10 is a sample survey questionnaire, in which Questions 1 through 6 measure level 3 behavior and Questions 7 through 10 measure level 4 results. Exhibit 11 is a sample survey questionnaire to be administered with the supervisors of the employees who completed the program. The supervisors may complete the survey for each employee or complete one survey to reflect the supervisor's perceptions of the group results. Then, you can compare the employees' self-assessments and their supervisors' assessments to see if their assessments are similar or different. If different, you can investigate the discrepancy further with other data collection methods such as interviews and/or observations.

When measuring level 3 and level 4 outcomes with surveys, you can also use retrospective surveys as shown in Exhibit 7 earlier.

Exhibit 10 A Sample Survey Questionnaire That Measures Self-Assessed Level 3 Behavior and Level 4 Results

Please indicate your level of lab safety compliance and how much you think your safety compliance has impacted the overall organizational results, by selecting one of the options.

About your work performance	Very untrue 1	2	3	4	Very true 5	Not sure
I am following safety guidelines.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I use safety gear when in the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I use the fume hood when handling chemicals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I use the safety checklist when in the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I appropriately dispose of biohazardous waste.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I complete proper hand-washing upon contact with chemicals and before leaving the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

About the overall results	Very untrue 1	2	3	4	Very true 5	Not sure
I think the level of lab safety is high.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel safe while in the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think the safety training helped eliminate safety-related accidents.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel a culture of safety consciousness among employees.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Exhibit 11 A Sample Survey Questionnaire That Measures Supervisors' Assessments on Level 3 Behavior

Please observe your employees' lab safety compliance for a few days, and indicate their level of lab safety compliance by selecting one of the options.

My employees:	Very untrue 1	2	3	4	Very true 5	Not sure
Are following safety guidelines.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use safety gear when in the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use the fume hood when handling chemicals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use the safety checklist when in the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Appropriately dispose of biohazardous waste.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Complete proper hand-washing upon contact with chemicals and before leaving the lab.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

These 4-level surveys may allow you to run correlation analyses to explore relationships between the levels and understand how different aspects of training outcomes are connected. For example, you may answer the following questions:

- Is there a correlation between level 1 and level 2 data? In other words, if trainees are more satisfied with the program, do they tend to feel more confident in their new knowledge and skills, and vice versa?
- Is there a correlation between level 2 and level 3 data? In other words, do employees with higher levels of confidence in their new knowledge and skills report higher levels of commitment to using their new knowledge and skills on the job?

Appendix A explains how to perform this type of correlation analyses with Excel.

Notes

1. Wagnild, G. M., & Young, H. M. (1993). Development and psychometric evaluation of the resilience scale. *Journal of Nursing Measurement*, 1(2), 165-178.
2. Kirkpatrick, D. (2004). How to start an objective evaluation of your training program. *T&D*, 58(5), 1-3.
3. Kirkpatrick, D. L. (1956). How to start an objective evaluation of your training program. *Journal of the American Society of Training Directors*, 10-11(May-June), 18-22.
4. Smith, S. (2008). Why follow levels when you can build bridges? *Training Development*, 62(9), 58-62.
5. Thalheimer, W. (2018). [Donald Kirkpatrick was not the originator of the four-level model of learning evaluation](#). *Work-Learning Research*.
6. Chyung, S. Y. (2019). *10-step evaluation for training and performance improvement*. SAGE.

5.3 Sample surveys used during performance analysis

The ultimate goal of training and performance improvement practice is to help organizations improve their outcomes. To achieve the goal, you will often begin by conducting a performance analysis, which involves a series of investigations aimed at uncovering the following:

- Whether there is a gap between the expected and current outcomes
- What the specific performance gap is
- Whether the performance gap is significant enough to warrant intervention
- What likely caused the identified performance gap
- If the probable causes are environmental factors, personal factors, or both
- What solutions could be implemented to reduce or eliminate the performance gap
- Which solutions are more cost-effective than others
- How changes are happening and being managed during the implementation of the selected solution

To investigate these issues, you will use various data collection methods, including surveys. When using surveys, it is important to consider different design approaches depending on the specific purpose of the survey.

5.3.1 Using a research approach common for program evaluations

First, compare this performance analysis situation to a program evaluation situation. When conducting a program evaluation, you may design your survey questionnaire with a set of multiple survey items using the same type of rating scales to determine the overall quality of the program.

For example, if you are investigating the level of job satisfaction (i.e., a construct) as a program outcome, you might develop a set of survey items grounded in Herzberg's

job satisfaction theory, as shown in Exhibit 12. You could administer this survey before and after the program and compare the average scores to determine changes in employee job satisfaction. You may also have groups of high- and low-performing employees complete a survey like the one shown in Exhibit 12 to compare differences between the groups.

This is a research approach, commonly used in program evaluations, and it can also be effectively applied in performance analysis to help identify and quantify performance gaps.

Exhibit 12 A Sample Survey Questionnaire Designed Based on Herzberg's Job Satisfaction Theory

Please rate your satisfaction level toward each of the following aspects about your job by selecting one of the numbers: 1 = No satisfaction, 2 = A little bit, 3 = Some, 4 = Quite a bit, 5 = High satisfaction.

Question	No satisfaction 1	A little bit 2	Some 3	Quite a bit 4	High satisfaction 5
I am praised when I perform well.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can be promoted based on my performance level.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find my job interesting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find my job challenging enough.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I see myself grow (develop) while doing my job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am allowed to make my own decisions while doing my job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

With these survey items, you can study if the program has helped increase employees' job satisfaction. In doing so, you may identify the overall job satisfaction level based on the average score of the data obtained from the multiple survey items. Perhaps, you may conduct this survey questionnaire before and after the program and

compare the differences in job satisfaction. This is a research approach, commonly used in program evaluations.

5.3.2 Using an assessment approach in performance analysis

Performance analysis—especially during cause analysis—more commonly requires an assessment approach to explore a broader range of environmental and personal factors that may be contributing to the performance issue.

For example, during your cause analysis, you might begin by interviewing a small, purposive sample of employees to assess probable causes to the performance issue. If those interviews reveal compelling evidence—such as low job satisfaction contributing to low performance—you might then conduct a survey with a larger employee group to validate the findings:

- Step 1: Interview with a small purposive sample
- Step 2: Identify probable causes
- Step 3: Conduct a survey with a large sample to validate the causes

In some cases, you may skip the initial interviews and instead use a well-designed survey questionnaire to assess the overall conditions influencing employee performance, including both environmental and personal factors. When doing so, you can apply established cause analysis frameworks such as Gilbert's (1978) Behavior Engineering Model¹ or Chevalier's (2003) updated model².

For example, suppose you are conducting a performance analysis at a manufacturing plant and have identified a significant performance gap related to safety violations among production workers. To investigate the causes of this performance gap, you might distribute a survey to both production workers and their supervisors. The survey could explore employees' understanding of safety-related policies and procedures as well as their actual safety-related behaviors.

In this context, Gilbert's Behavior Engineering Model (Table 9) can serve as a valuable framework for survey development.

Table 9 Thomas Gilbert's Behavior Engineering Model

Category	Information	Instrumentation	Motivation
E: Environmental Supports	<i>Data</i>	<i>Resources</i>	<i>Incentives</i>
P: Person's repertory of behavior	<i>Knowledge</i>	<i>Capacity</i>	<i>Motives</i>

The Behavior Engineering Model guides you to analyze six factors that contribute to performance outcomes: 1. data, 2. resources, 3. incentives, 4. knowledge, 5. capacity, and 6. motives. Gilbert viewed that these six factors are essential for engineering human behavior and for elevating typical performance to exemplary performance.

Using the six categories, you can develop a survey questionnaire like the one shown in Exhibit 13. If you plan to match the survey data with another set of data (e.g., their supervisor's data), the survey should be identifiable by including a field for the respondent's name or ID. Otherwise, you may administer the survey anonymously to encourage more candid responses.

If you decide to survey the supervisors of the production workers, the supervisor survey questionnaire should have slightly modified wording as shown in Exhibit 14.

Exhibit 13 A Sample Survey Questionnaire Designed Based on Gilbert's Behavior Engineering Model (Employee Version)

Please rate each of the following aspects about your job by selecting a number from the 5-point scale where 1 = Almost never and 5 = Almost always.

Environmental factor: Data	Almost never 1	2	3	4	Almost always 5
-----------------------------------	---------------------------	----------	----------	----------	----------------------------

- 1. I receive clear expectations about what tasks that I need to complete in my job.
- 2. I receive clear expectations about how well I need to complete my job tasks.
- 3. I receive sufficient guidance to help me perform my job tasks.
- 4. I receive sufficient feedback about my job performance.

Environmental factor: Instruments	Almost never 1	2	3	4	Almost always 5
--	---------------------------	----------	----------	----------	----------------------------

- 5. I have access to all the tools that I need to use to complete my job tasks.
- 6. I have access to all the information/materials that I need to use to complete my job tasks.
- 7. I am assigned to a job schedule that is optimal for me.
- 8. I find my workplace conducive to good performance.

Environmental factor: Incentives	Almost never 1	2	3	4	Almost always 5
---	---------------------------	----------	----------	----------	----------------------------

- 9. I face positive or negative consequences depending on the quality of my job performance.
- 10. I receive fair monetary compensation for my job performance.
- 11. I am provided with non-monetary incentives that are meaningful to me.
- 12. I am provided with career development opportunities.

Personal factor: Knowledge	Almost never 1	2	3	4	Almost always 5
-----------------------------------	---------------------------	----------	----------	----------	----------------------------

13. I have sufficient knowledge and skills to perform all my job tasks. 1 2 3 4 5

14. I receive adequate training or education when I need to gain more knowledge and skills. 1 2 3 4 5

15. I can develop my knowledge and skills to complete my job tasks better. 1 2 3 4 5

16. I have knowledge and skills valued by my job. 1 2 3 4 5

Personal factor: Capacity	Almost never 1	2	3	4	Almost always 5
----------------------------------	---------------------------	----------	----------	----------	----------------------------

17. I am physically fit to do all of my current job tasks. 1 2 3 4 5

18. I am mentally fit to do all of my current job tasks. 1 2 3 4 5

19. I have aptitudes to perform all of my current job tasks. 1 2 3 4 5

20. I am an appropriate person to do all of my current job tasks. 1 2 3 4 5

Personal factor: Motives	Almost never 1	2	3	4	Almost always 5
---------------------------------	---------------------------	----------	----------	----------	----------------------------

21. I am intrinsically motivated to do my job. 1 2 3 4 5

22. I am committed to produce good job performance. 1 2 3 4 5

23. I am willing to do my best when working on my job tasks. 1 2 3 4 5

24. I am satisfied with what I can do in my current job. 1 2 3 4 5

Let's say you conducted the survey with a group of plant production workers and the

corresponding survey with their supervisors. Table 10 presents partial hypothetical data from the BEM surveys (showing only the Data factor).

In this case, your primary interest is not just to assess the overall quality of the environmental and personal factors based on the overall average scores. Rather, you aim to identify which of the six categories are underperforming and in need of improvement.

To determine which categories may be lacking, you can calculate the average score of the multiple survey items within each of the six Behavior Engineering Model categories. For example, within the Data category:

- Employees' average score: $(4.05 + 3.60 + 2.55 + 2.15) / 4 = 3.09$
- Supervisors' average score: $(4.25 + 3.66 + 3.25 + 3.05) / 4 = 3.55$

The average scores of 3.09 for employees and 3.55 for supervisors suggest that the Data category may need improvement.

You can go further by investigating which specific items within each category require attention. For instance, in the Data category, survey items #3 and 4 may be of concern. In particular, survey item #4 (see Figure 27) shows a noticeable discrepancy between employees' and their supervisors' assessments, which signals a need for further investigation (e.g., through in-depth interviews).

- Employees: "Q4: I receive sufficient feedback on my job performance." (average = 2.15)
- Supervisors: "Q4: I provide my employee sufficient feedback on their job performance." (average = 3.05)

Exhibit 14 A Sample Survey Questionnaire Designed Based on Gilbert's Behavior Engineering Model (Supervisor Version)

Please rate each of the following aspects about each of the employees that you super-

use by selecting a number from the 5-point scale where 1 = Almost never and 5 = Almost always.

Employee's name that you are evaluating:

Environmental factor: Data	Almost never 1	2	3	4	Almost always 5
1. I provide my employee with clear expectations about what tasks they need to complete in their job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. I provide my employee with clear expectations about how well they need to complete their job tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. I provide my employee sufficient guidance to help them perform their job tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I provide my employee sufficient feedback on their job performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Environmental factor: Instruments	Almost never 1	2	3	4	Almost always 5
5. I give my employee access to all tools that they need to use to complete their job tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. I give my employee access to all information/materials that they need to use to complete their job tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. I assign my employee to a job schedule that is optimal for them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. I find my employee's workplace conducive to good performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Environmental factor: Incentives	Almost never 1	2	3	4	Almost always 5
---	---------------------------	----------	----------	----------	----------------------------

9. I give my employee positive or negative consequences depending on the quality of their job performance.

10. I give my employee fair monetary compensation for their job performance.

11. I provide my employee with non-monetary incentives that are meaningful to them.

12. I provide my employee with career development opportunities.

Personal factor: Knowledge	Almost never 1	2	3	4	Almost always 5
-----------------------------------	---------------------------	----------	----------	----------	----------------------------

13. I think my employee has sufficient knowledge and skills to perform all of their job tasks.

14. I provide my employee with adequate training and/or education when they need to gain more knowledge and skills.

15. I think my employee is able to develop their knowledge and skills to complete their job tasks better.

16. I value my employee's knowledge and skills for their job.

Personal factor: Capacity	Almost never 1	2	3	4	Almost always 5
----------------------------------	---------------------------	----------	----------	----------	----------------------------

17. My employee is physically fit to do all of their current job tasks.

18. My employee is mentally fit to do all of their current job tasks.

19. My employee has aptitudes to perform all of their current job tasks.

20. My employee is an appropriate person for doing all of their current job tasks.

Personal factor: Motives	Almost never 1	2	3	4	Almost always 5
21. My employee is intrinsically motivated to do their job.	○	○	○	○	○
22. My employee is committed to produce good job performance.	○	○	○	○	○
23. My employee is willing to do their best when working on their job tasks.	○	○	○	○	○
24. My employee is satisfied with what they can do in their current job.	○	○	○	○	○

Table 10 Hypothetical Survey Data Results Obtained from Employees and Supervisors

Employees	Almost never 1	2	3	4	Almost always 5	Mean
Environmental factor: Data						
1. I receive clear expectations about what tasks that I need to complete in my job.	1 (5%)	2 (10%)	2 (10%)	5 (25%)	10 (50%)	4.05
2. I receive clear expectations about how well I need to complete my job tasks.	2 (10%)	3 (15%)	3 (15%)	5 (25%)	7 (35%)	3.60
3. I receive sufficient guidance to help me perform my job tasks.	5 (25%)	6 (30%)	4 (20%)	3 (15%)	2 (10%)	2.55
4. I receive sufficient feedback on my job performance.	8 (40%)	5 (25%)	4 (20%)	2 (10%)	1 (5%)	2.15

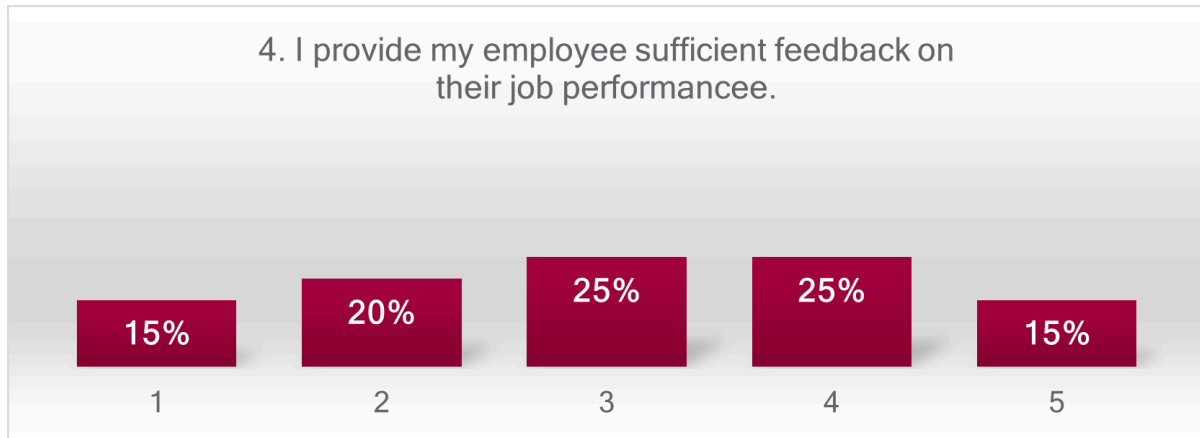
Supervisors	Almost never 1	2	3	4	Almost always 5	Mean
Environmental factor: Data						
1. I provide my employee with clear expectations about what tasks that they need to complete in their job.	0 (0%)	1 (5%)	3 (15%)	6 (30%)	10 (50%)	4.25
2. I provide my employee with clear expectations about how well they need to complete their job tasks.	1 (5%)	2 (10%)	5 (25%)	7 (35%)	5 (25%)	3.65
3. I provide my employee sufficient guidance to help them perform their job tasks.	3 (15%)	2 (10%)	5 (25%)	7 (35%)	3 (15%)	3.25
4. I provide my employee sufficient feedback on their job performance.	3 (15%)	4 (20%)	5 (25%)	5 (25%)	3 (15%)	3.05

Figure 27 Bar Graphs Showing Hypothetical Survey Data Results Obtained from Employees and Supervisors

Employees' assessments



Supervisors' assessments



Notes

1. Gilbert, T. F. (1978). *Human competence: Engineering worthy performance*. McGraw-Hill.
2. Chevalier, R. (2003). [Updating the behavior engineering model](#). *Performance Improvement*, 43(5), 8-14.

5.4 Sample surveys used during instructional design

While working on an instructional design project, you will have opportunities to collect data from various groups during the job task analysis, content analysis, learner analysis, and formative evaluation. For example, you may collect data from subject matter experts and learners (actual performers) to:

- *Determine specific tasks to be performed for the job*
- *Identify the desired level of knowledge, skills, and attitudes that learners should acquire*
- *Assess learners' current level of knowledge, skills, and attitudes*
- *Determine the necessary instructional content*
- *Solicit feedback on parts of the instruction that may need revision*

You may decide to conduct a survey in some of these cases.

For example, let's say you are offering an annual mandatory lab safety course to all lab technicians. You discovered that some lab technicians feel confident in their lab safety knowledge and do not perceive a need to complete this course. To improve learner motivation, you designed the course using Keller's ARCS model^{1,2}, addressing the following four elements in instruction:

- **Attention:** If learners find the presentation of instructional content to be clear, interesting, and engaging
- **Relevance:** If learners find the instructional content to be relevant to them and address their personal and professional needs
- **Confidence:** If learners feel confident in using the information they learned
- **Satisfaction:** If learners feel satisfied with the overall learning experience and outcomes

After completing your course design, you plan to conduct a pilot test (formative evaluation) with a small group of lab technicians to check how well the instruction works

and which parts of the instruction may need to be revised to maximize learning outcomes. After the pilot course, you decide to administer a survey to the pilot learners.

Since you designed the course with the ARCS model in mind, you wanted to evaluate whether the instruction effectively addresses each ARCS component. Therefore, you designed your survey items aligned with the four factors as shown in Exhibit 15.

Here, it is important to note that the ARCS components are interrelated and may not function as entirely separate constructs. For example, when instruction is easy to follow, interesting, and relevant, it not only holds learners' attention but also enhances their learning, confidence, and overall satisfaction. Thus, the overall construct that you are measuring with a survey questionnaire such as the one presented in Exhibit 15 is the course's motivational appeal to learners. The section headings—Attention, Relevance, Confidence, and Satisfaction—are presented only for organizational purposes in the survey design, not to imply separate statistical constructs.

Exhibit 15 A Sample Survey Questionnaire to Measure ARCS Factors in Instruction

Please rate each of the following aspects about the course based on your experience, by selecting a number from the 5-point scale where 1 = Not at all and 5 = Very much.

Attention	Not at all				Very much
	1	2	3	4	5
1. I was able to recognize the course objectives.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. I found the content to be easy to follow.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. I was able to pay attention to the content throughout the course.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I found the examples provided in the course interesting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Relevance	Not at all 1	2	3	4	Very much 5
------------------	-------------------------	----------	----------	----------	------------------------

5. I found the content to be relevant to lab safety in my job setting. 1 2 3 4 5

6. I found the content to be useful for improving lab safety in my lab. 1 2 3 4 5

7. I was able to follow and understand the content based on my previous knowledge of lab safety. 1 2 3 4 5

8. I think the multimedia used in the course helped me understand lab safety issues better. 1 2 3 4 5

Confidence	Not at all 1	2	3	4	Very much 5
-------------------	-------------------------	----------	----------	----------	------------------------

9. I found the exercises to be challenging enough. 1 2 3 4 5

10. I felt confident enough to continue the course. 1 2 3 4 5

11. I now fully understand what I need to do to ensure lab safety. 1 2 3 4 5

12. I will be able to apply the new knowledge to ensure lab safety. 1 2 3 4 5

Satisfaction	Not at all 1	2	3	4	Very much 5
---------------------	-------------------------	----------	----------	----------	------------------------

13. I am glad that I am knowledgeable about lab safety. 1 2 3 4 5

14. I feel it was worth spending time to complete this course. 1 2 3 4 5

15. I had a positive learning experience during the course. 1 2 3 4 5

16. I think other technicians should complete this course. 1 2 3 4 5

Notes

1. Keller, J. M. (1999). *Using the ARCS motivational process in computer-based instruction and distance education*. *New Directions for Teaching & Learning*, 1999(78), 37-47.
2. Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. New York: Springer.

5.5 Chapter activity: How would you design your own survey questionnaire?

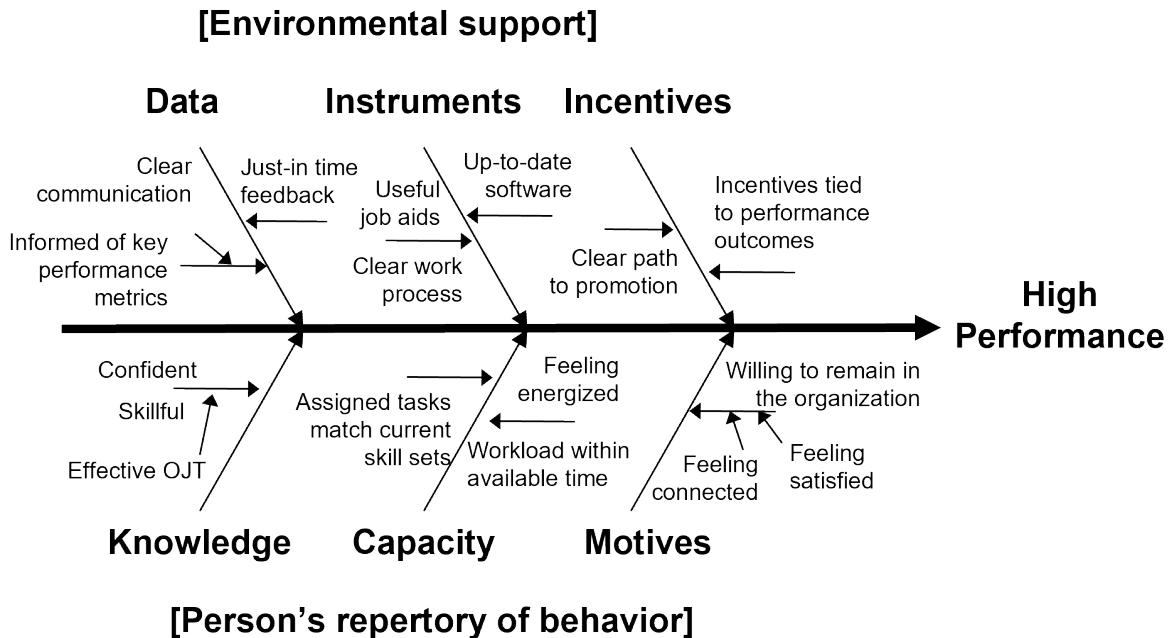
Imagine you are performing a performance analysis, designing instruction, or conducting an evaluation. Choose one of the following scenarios and design a survey questionnaire to be used in that context. Clearly describe what each survey item is intended to measure.

Performance analysis

During a performance analysis, you interviewed a small group of exemplary performers to identify possible contributors to their high performance. You analyzed the interview data using a combination of the Behavior Engineering Model (BEM) and Fishbone Analysis as frameworks. The results are shown in Figure 28. Now, you want to use this information to develop a survey questionnaire to collect data from all employees.

For your practice, design a total of 6-12 survey items based on this initial cause-and-effect analysis information shown in Figure 28 (choose the ones that you think are appropriate to include in a survey questionnaire).

Figure 28 A Cause-and-Effect Analysis with BEM Categories (Adapted from Chyung, 2008, p. 127¹)



[For accessibility, refer to the [long description](#) of the fishbone diagram.]

For example, based on this item “Useful job aids,” your survey item may look like one of the following, depending on whether you choose to use a statement format or a question format:

[Statement format]

The job aids are useful.

- Strongly disagree
- Somewhat disagree
- Neutral
- Somewhat agree
- Strongly agree

[Question format]

How useful are the job aids?

- Not at all
- A little bit
- Somewhat
- Quite a bit
- Very much

When selecting the format (including the response scale type), think ahead about how you plan to analyze and report the data. Apply the same format to all survey items for consistency.

Instructional design

You are designing an e-learning program on conflict resolution strategies, applying the following six e-learning design principles developed by Mayer and his colleagues (Clark, 2002², Mayer, 2003³):

1. The **multimedia** principle: Adding graphics to words can improve learning.
2. The **contiguity** principle: Placing text near graphics improves learning.
3. The **modality** principle: Explaining graphics with audio improves learning.
4. The **redundancy** principle: Explaining graphics with audio and redundant text can hurt learning.
5. The **coherence** principle: Using gratuitous visuals, text, and sounds can hurt learning.
6. The **personalization** principle: Use conversational tone and pedagogical agents to increase learning.

You plan to conduct a pilot test with a group of learners to assess whether your e-learning program aligns with these six principles. Design a survey questionnaire to gather feedback from learners. Your survey questionnaire should include 6-12 survey items, with at least one item representing each principle.

Evaluation

Your nonprofit organization recently delivered a workshop to parents, aimed at helping them improve parental resilience, defined as the process of managing “stress and functioning well when faced with challenges, adversity and trauma” (CSSP, 2018, p. 1⁴).

The main goals of the workshop were:

1. Increased ability to cope with stress and bounce back during parenting
2. Increased ability to solve problems associated with parenting
3. Increased ability to care for their own well-being, physically, mentally, and emotionally
4. Increased behavior of connecting with families and friends and seeking help when needed

5. Increased behavior of seeking assistance using community resources

It has been a couple of weeks after the workshop was delivered, and your supervisor has asked you to develop a follow-up survey questionnaire for the participants who completed the workshop.

Because you did not administer a pre-workshop survey before they completed the workshop, and your supervisor wants to see there were overall improvements in the intended outcomes, you've decided to use a retrospective survey method.

Design a survey questionnaire that allows you to measure improvements in the five main outcome areas. The survey should include 6-12 survey items, with at least one item for each outcome. You may include more than one survey item per outcome. While designing your items, consider how you plan to analyze and report the data, such as by calculating average scores or reporting percentages of perceived improvement.

Notes

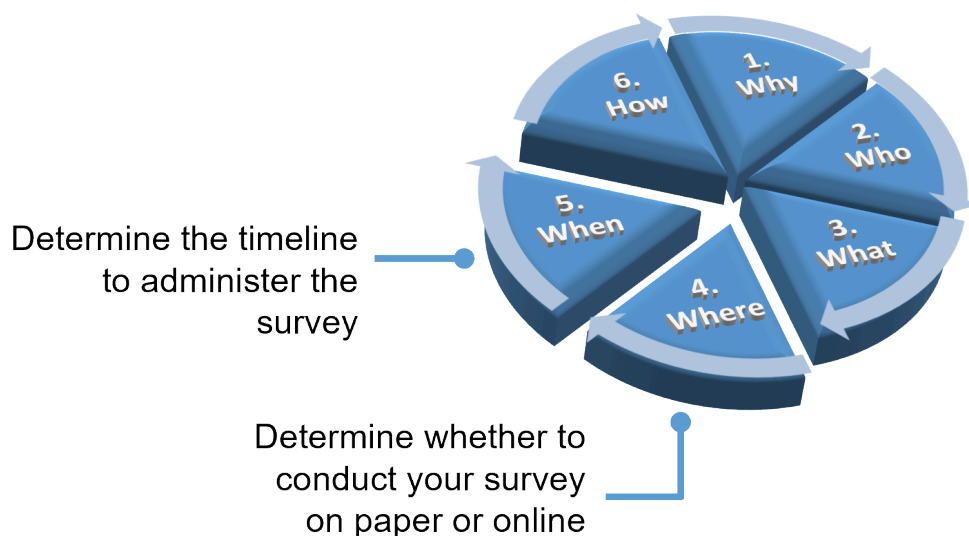
1. Chyung, S. Y. (2008). *Foundations of instructional and performance technology*. HRD Press.
2. Clark, R. (2002, September 10). *Six principles of effective e-learning: What works and why*. *The eLearning Guild's Learning Solutions Magazine*.
3. Mayer, R. E. (2003). *The promise of multimedia learning: Using the same instructional design methods across different media*. *Learning and Instruction*, 13(2), 125-139.
4. Center for the Study of Social Policy (CSSP) (2018). [Families and the protective factors framework](#).

6. ADMINISTER SURVEYS ON PAPER OR ONLINE

6.0 Introduction

The fourth and fifth steps in the survey design and data analysis process involve determining **Where** and **When** you will administer your survey (Figure 29).

Figure 29 The Where and When Steps in the Survey Design and Data Analysis Process



Where: Surveys can be administered on paper or online, depending on the context and accessibility of your participants:

- **Paper-based surveys** are ideal for settings where you have direct access to participants, such as at the end of in-person training programs or other on-site events. If you cannot distribute surveys in person, you may mail them to participants, often including a self-addressed, stamped envelope to encourage return.
- **Online surveys** are effective for reaching a large number of participants across various locations as long as they have access to the necessary technology (e.g., a computer or mobile device, Internet access, and email address to receive an invitation to the online survey). When designing online surveys, ensure they are optimized for multiple devices, including desktops, laptops, tablets, and smartphones.

When: When conducting an online survey, it is important to establish a **clear data collection timeframe**, typically lasting one or two weeks. You want to avoid scheduling your survey during holidays or particularly busy periods for participants—for example, at the end of a fiscal quarter or near tax deadlines in legal and financial organizations.

Learning Objectives

After completing this chapter, you will be able to:

- Identify effective design principles for print-based survey questionnaires
- Compare various web-based survey platforms and their key features
- Select an appropriate survey format and platform based on specific research or evaluation needs

1
Not at all
confident

2

3
Somewhat
confident

4

5
Very
confident

An alternative is to consider **a web-based survey system** to create your survey questionnaire. These systems often provide templates and menu-driven design features. They may also allow you to download the survey as a Word document, giving you the flexibility to adjust the formatting for paper distribution while maintaining clarity and structure.

There are many online survey tools available. In the following sections, we will explore three popular programs: Google forms, Qualtrics, and Survey Monkey.

6.2 Online surveys with Google Forms

After you create a [Google account](#), you gain free access to a suite of Google Apps, including **Google Drive**. Within Google Drive, you will find **Google Docs**, **Google Sheets**, **Google Slides**, and **Google Forms** (Figure 30). Among these, **Google Forms** is a simple yet effective tool for creating and administering online surveys.

Once you've created a Google Form, you can use the menu icons (Figure 31) to add a variety of elements to your survey, such as different types of questions, titles and descriptions, images, videos, and sections (similar to page breaks).

Figure 30 A Screen Shot of Google Drive Menu Showing Google Forms¹

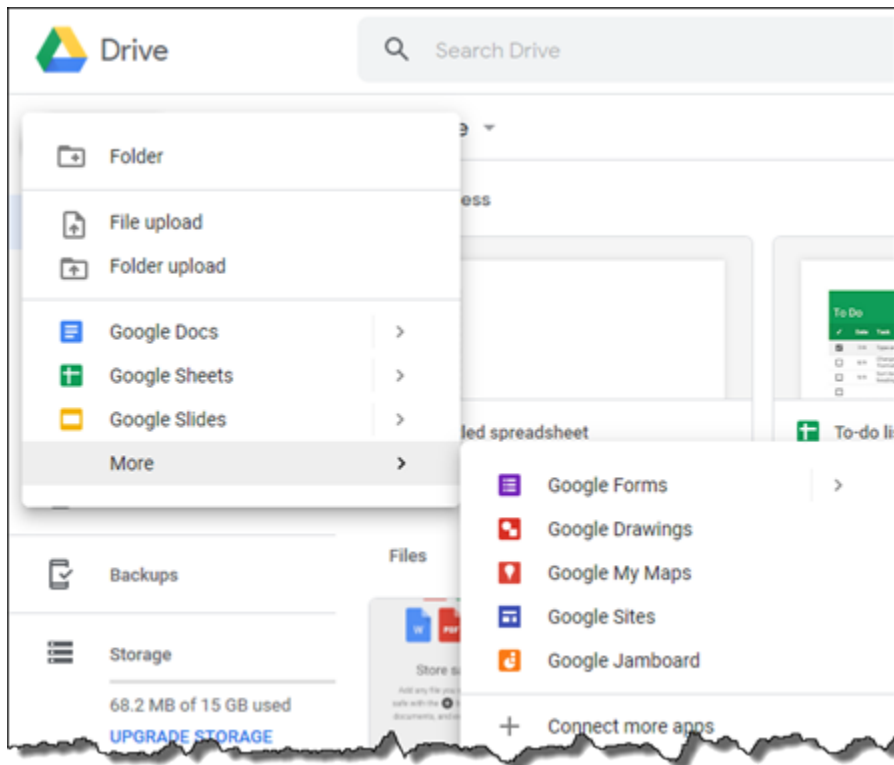
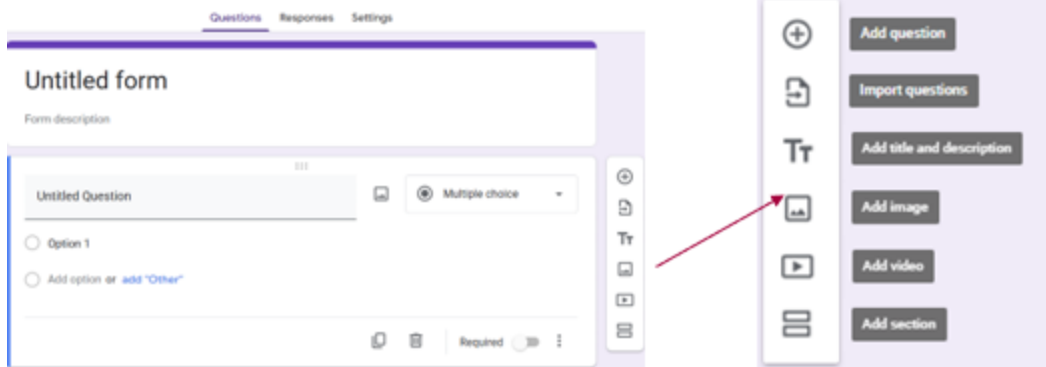


Figure 31 A Screen Shot of a New Google Form with Menu²



Google Forms allows you to use several rating scales including **Multiple choice**, **Checkboxes**, **Dropdown**, and **Linear scale** (Figure 32). The **Multiple choice grid** and **Checkbox grid** options present multiple choice and checkbox selections in a grid (table) format. Google Forms also provide **templates**, including those for course evaluations and event feedback (Figure 33). The data will be stored in a **linked Google Sheet**, making data management and analysis convenient.

Figure 32 Screen Shots of Scales Options Available in Google Forms³

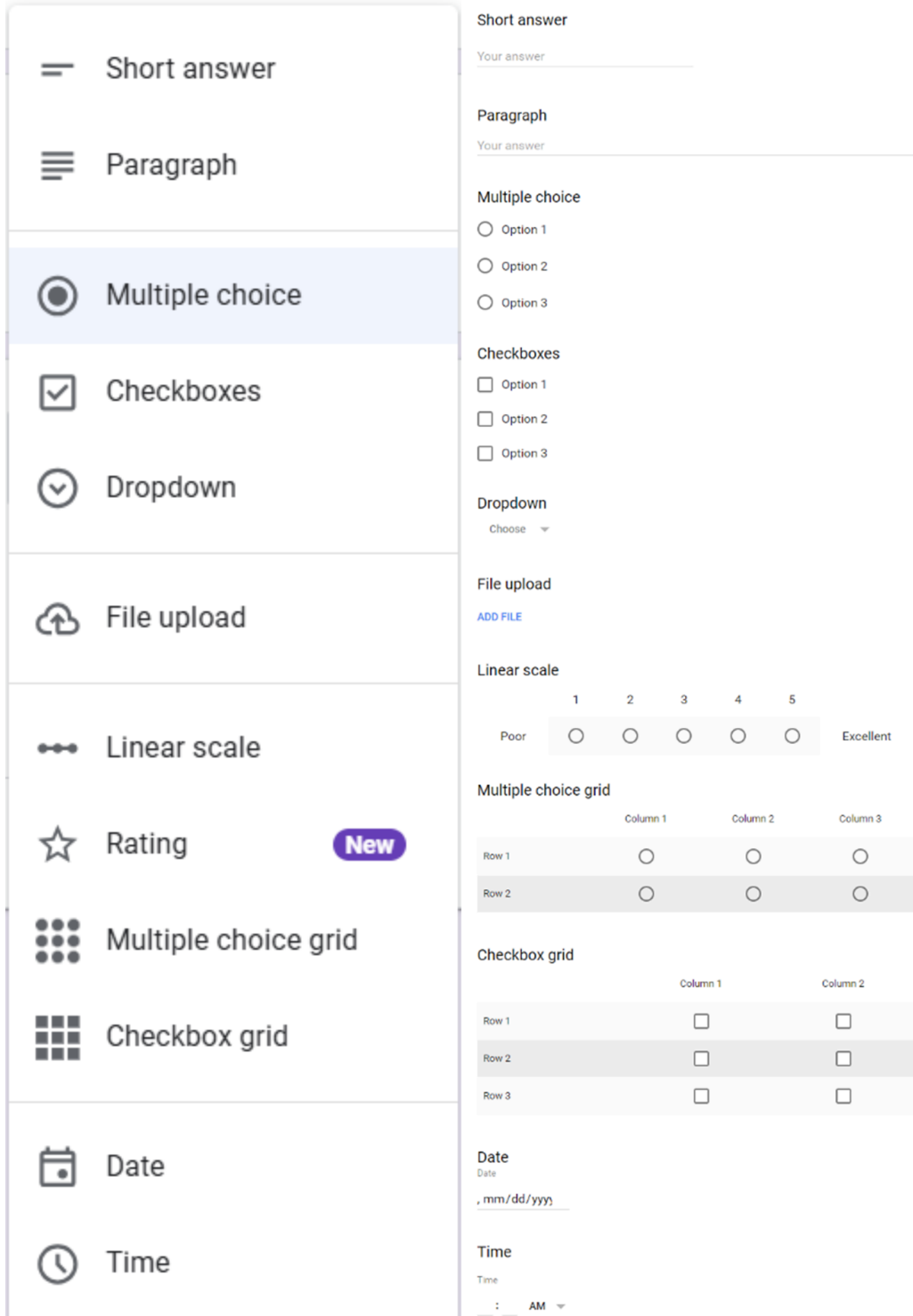


Figure 33 Screen Shots of Templates Available in Google Forms⁴

Course evaluation

Please submit feedback regarding the course you have just completed, including feedback on course structure, content, and instructor.

***Required**

Class name *

Your answer

Instructor *

Your answer

Level of effort

	Poor	Fair	Satisfactory	Very good	Excellent
Level of effort you put into the course	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Contribution to learning

	Poor	Fair	Satisfactory	Very good	Excellent
Level of skill/knowledge at start of course	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Level of skill/knowledge at end of course	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Level of skill/knowledge required to complete the course	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Contribution of course to your skill/knowledge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Skill and responsiveness of the instructor

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Instructor was an effective lecturer/demonstrator	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Presentations were clear and organized	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Instructor stimulated student interest	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Instructor effectively used time during class periods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Instructor was available and helpful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grading was prompt and had useful feedback	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Course content

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Learning objectives were clear	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Course content was organized and well planned	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Course workload was appropriate	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Course organized to allow all students to participate fully	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Event feedback

Thank you for participating in our event. We hope you had as much fun attending as we did organizing it.

We want to hear your feedback so we can keep improving our logistics and content. Please fill the quick survey and let us know your thoughts (your answers will be anonymous).

***Required**

How satisfied were you with the event? *

	1	2	3	4	5	
Not very	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very much

How relevant and helpful do you think it was for your job? *

	1	2	3	4	5	
Not very	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very much

What were your key take aways from this event?

Your answer

How satisfied were you with the logistics? *

1 = Very dissatisfied 5 = Very satisfied

	1	2	3	4	5	NA
Accommodation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Welcome kit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Communication emails	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Transportation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Welcome activity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Venue	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Closing ceremony	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Additional feedback on logistics *

Your answer

Which sessions did you find most relevant? *

	Not relevant	Relevant	Very relevant	Did not attend
Welcome activity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaker #1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Activity #1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaker #2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Activity #2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Closing activity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How satisfied were you with the session content?

Both presented and pre-read material

	1	2	3	4	5	
Poor	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Excellent

Notes

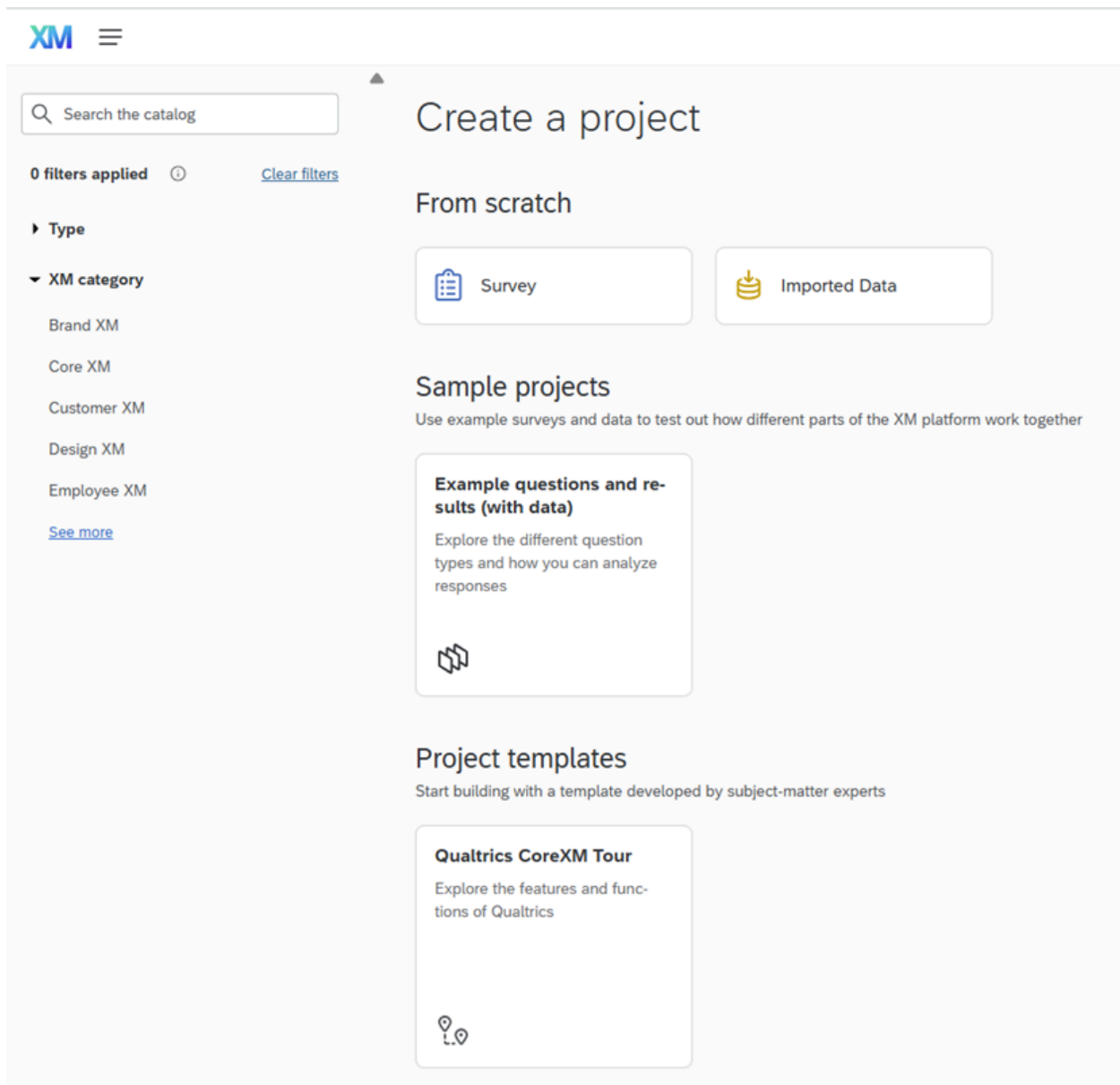
1. [Brand Resource Center](#). Google and Google Drive are trademarks of Google LLC and this book is not endorsed by or affiliated with Google in any way.
2. *Ibid.*
3. *Ibid.*

4. *Ibid.*

6.3 Online surveys with Qualtrics

[Qualtrics®](#) provides sophisticated survey design and administration solutions. It offers a wide range of templates and design options to suit various research and evaluation needs. When creating a new survey project in Qualtrics, you can create your project **from scratch** or select from a variety of **pre-designed templates** (Figure 34).

Figure 34 A Screen Shot of Qualtrics Menu for Creating a New Project¹



Even when designing a survey from scratch, the **intuitive graphical menu interface** makes it easy to navigate and select appropriate types of survey questions (Figure 35). Within a few clicks, you can add survey items using different **response scale formats**, including text-based or graphic-based scales displayed either horizontally or vertically. It suggests an extensive list of multiple-choice response options (Figure 36).

Figure 35 A Screen Shot of Question Types Available in Qualtrics

Question type

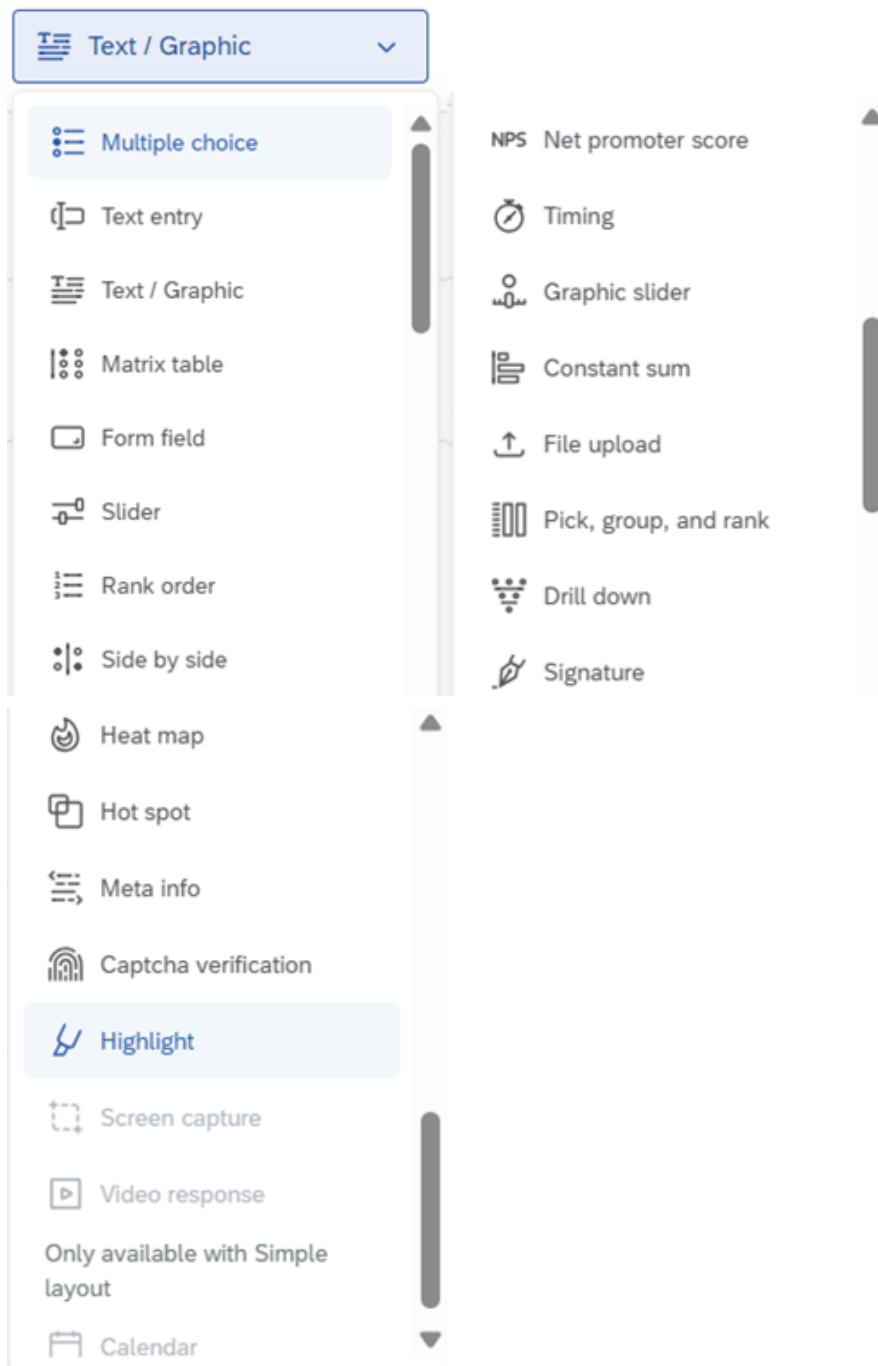
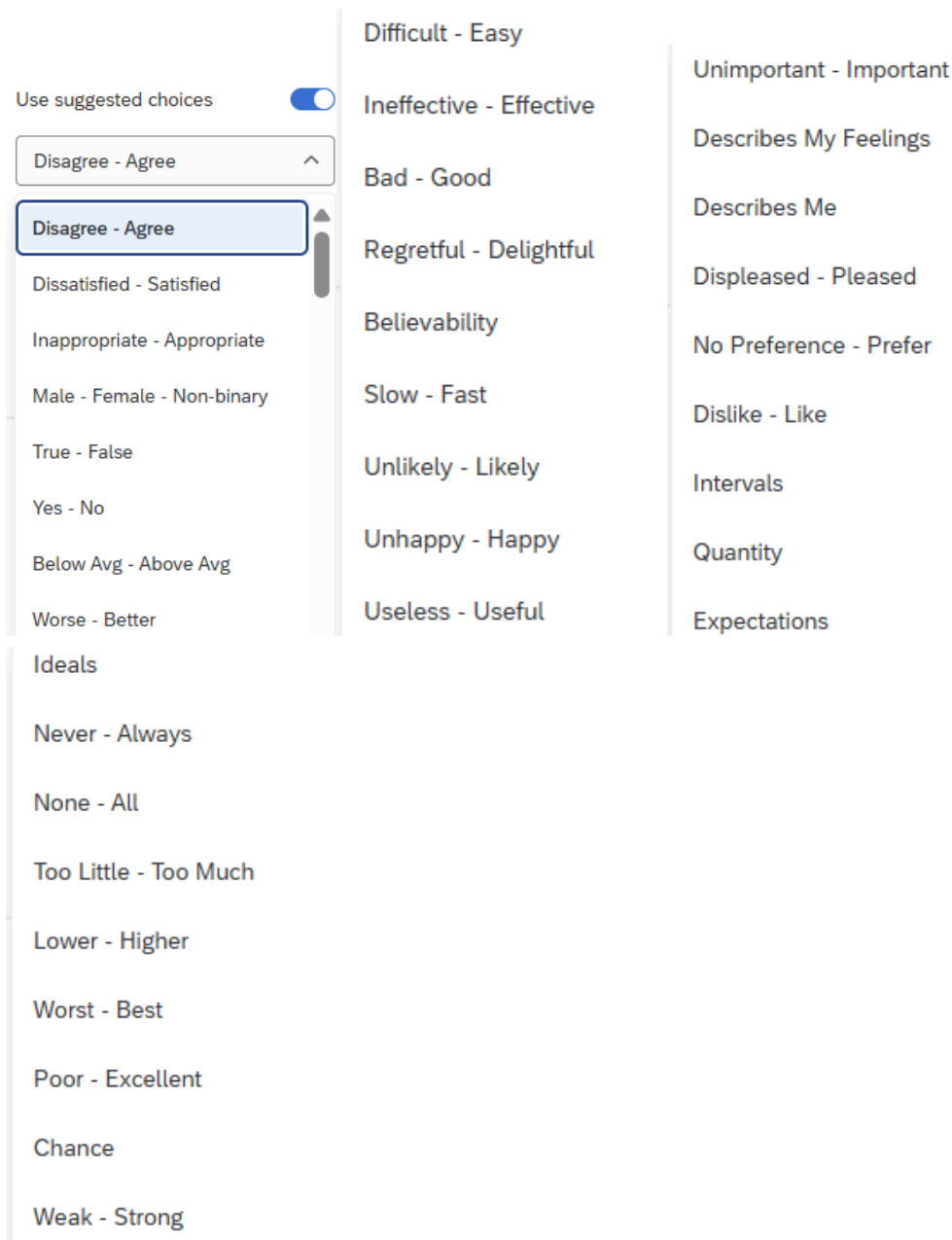


Figure 36 A Screen Shot of Various Multiple-Choice Options Available in Qualtrics



While building your survey, you can use the **preview** feature to test how the survey would appear on both **desktop and mobile devices**, helping you adjust the design as needed (Figure 37). This ensures your survey is mobile-friendly, an important consideration, as many respondents may complete the survey on their smartphone or tablet. You can **export** your survey as a Word document (Figure 38).

Figure 37 A Screen Shot of the Preview Feature Available in Qualtrics

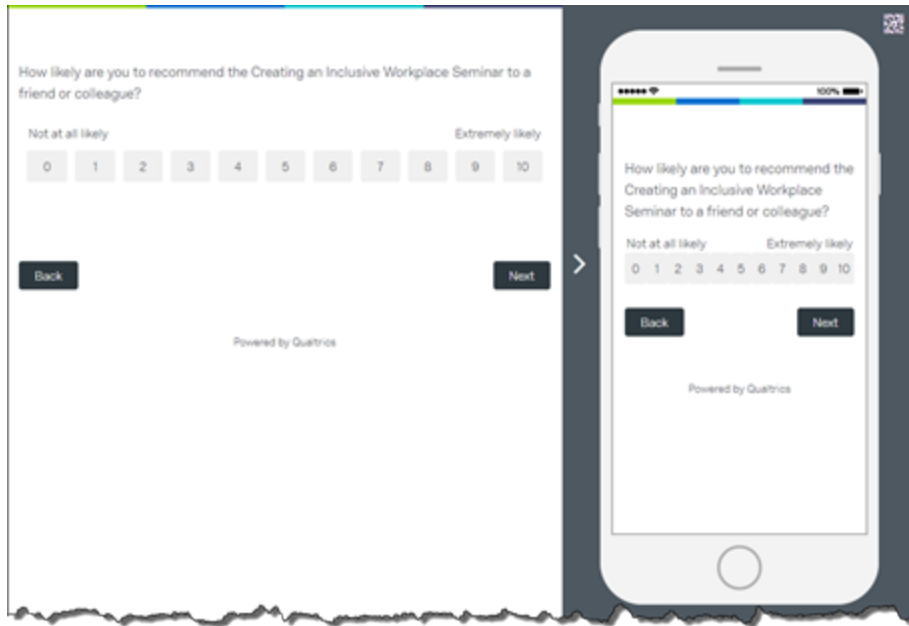
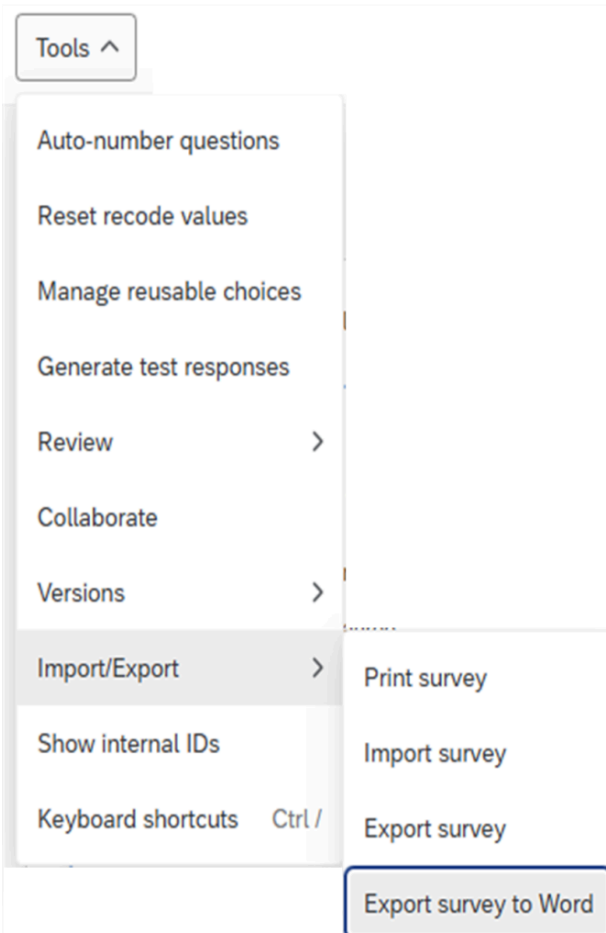


Figure 38 A Screen Shot of the Tools > Import/Export > Export Survey to Word Feature Available in Qualtrics



Notes

1. *Qualtrics is a trademark of Qualtrics International Inc., and this book is not endorsed by or affiliated with Qualtrics in any way*

6.4 Online surveys with SurveyMonkey

[SurveyMonkey®](#) is another widely used tool for developing, distributing, and managing online surveys. It offers a broad selection of **survey templates**, while also allowing you to create a survey **from scratch** (Figure 39).

Figure 39 Screen Shots of SurveyMonkey Options—“Choose a Starting Point” and “Pick a Popular Template”¹

Choose a starting point

The 'Choose a starting point' section contains seven cards:

- Start from scratch:** Begin with a blank survey or form. Then add your questions, test, and images.
- Copy an existing survey:** Choose a survey. Make a copy. Edit as needed.
- Build with AI:** Type a short prompt. AI will create a tailored survey or form just for you.
- Pick a popular template:** Ask the right questions and save time with a template built for your situation.
- Paste your content:** Add all your questions and answers at once. Or copy and paste one at a time.
- Build a target audience:** Choose audience demographics. Then build a survey when you're ready.
- Use proven research methods:** Launch studies with target audiences about product, marketing, and brand.

Explore templates

The 'Explore templates' section features a search bar and a 'Filters' sidebar on the left. The sidebar includes:

- Plan type:** Show free templates
- Categories:** Team templates (2), Most Popular (22), Forms (227), Events (23), Human Resources (302), Customer Feedback (342), General Business (232), Satisfaction (2), Marketing (24), Education (42), Benchmarkable (26), Just for Fun (27), Academic/Research (7), Quizzes (2)

The main area displays a grid of 15 template cards, each with a title, usage count, and a brief description:

- Customer Satisfaction Template:** Used 474622+ times. Keep your customers happy and turn them into advocates.
- Employee Engagement Template:** Used 282222+ times. Learn about people's experience with your organization.
- Market Research - Product Testing Template:** Used 24022+ times. Launching a new product isn't easy. Verify you have the right audience...
- Net Promoter® Score (NPS) Template:** Used 222222+ times. Measure customer loyalty and take action to reduce churn.
- Remote Work Employee Check-In Survey:** Used 20022+ times. Support employees who work from home (WFH) with weekly check-ins.
- Software and App Customer Feedback with NPS®:** Used 14022+ times. Would your customers recommend your software or app to others?...
- SurveyMonkey-Paradigm Belonging and Inclusion...:** Used 27022+ times. Developed with Paradigm, a leader in diversity and inclusion strategy...
- University Instructor Evaluation Template:** Used 16222+ times. Get student feedback on a professor's teaching style.
- Website Feedback Template:** Used 20222+ times. Gather feedback from website visitors to understand where and...
- General Event Feedback Template:** Used 20022+ times. Find out how people felt about your event to improve the next one.
- Scheduling Poll:** Used 222222+ times. Decide what time works best for your meeting or event.
- SurveyMonkey Question Type Test:** Used 22222+ times. How do you feel about the new question type?
- Volunteer Feedback Template:** Used 20022+ times. Ask volunteers how it was working with you.
- ACH Form:** ACH Form Template.
- Alumni Contact Information Form:** Used 20022+ times. Alumni Feedback Information Form.

Using the **preview** function, you can test how your survey will appear on desktop computers, tablets, and cell phones (Figure 40). When creating a survey item from scratch, you can choose from a wide range of question types (Figure 41).

Figure 40 Screen Shots of the Preview Feature Available in SurveyMonkey

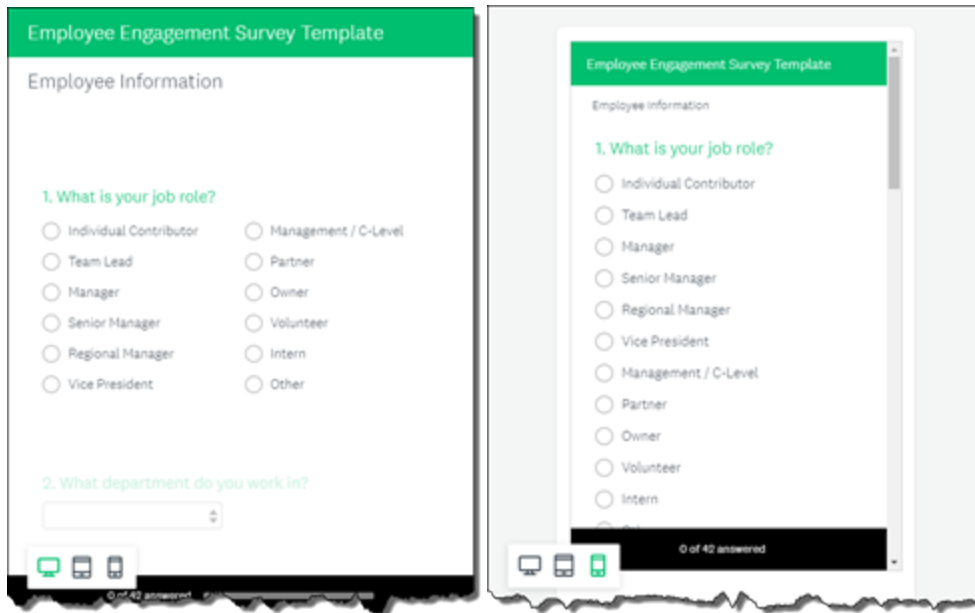
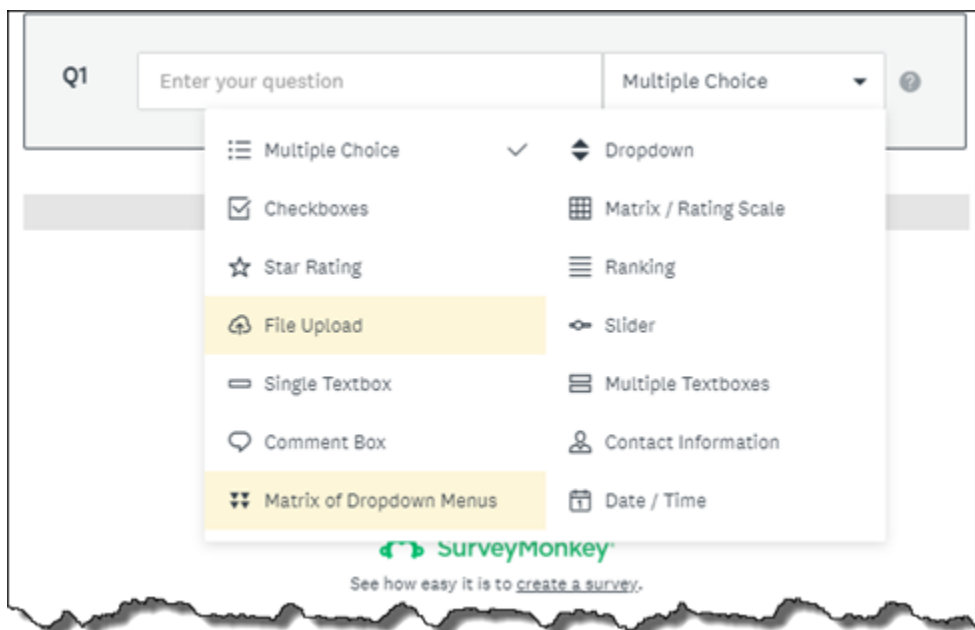
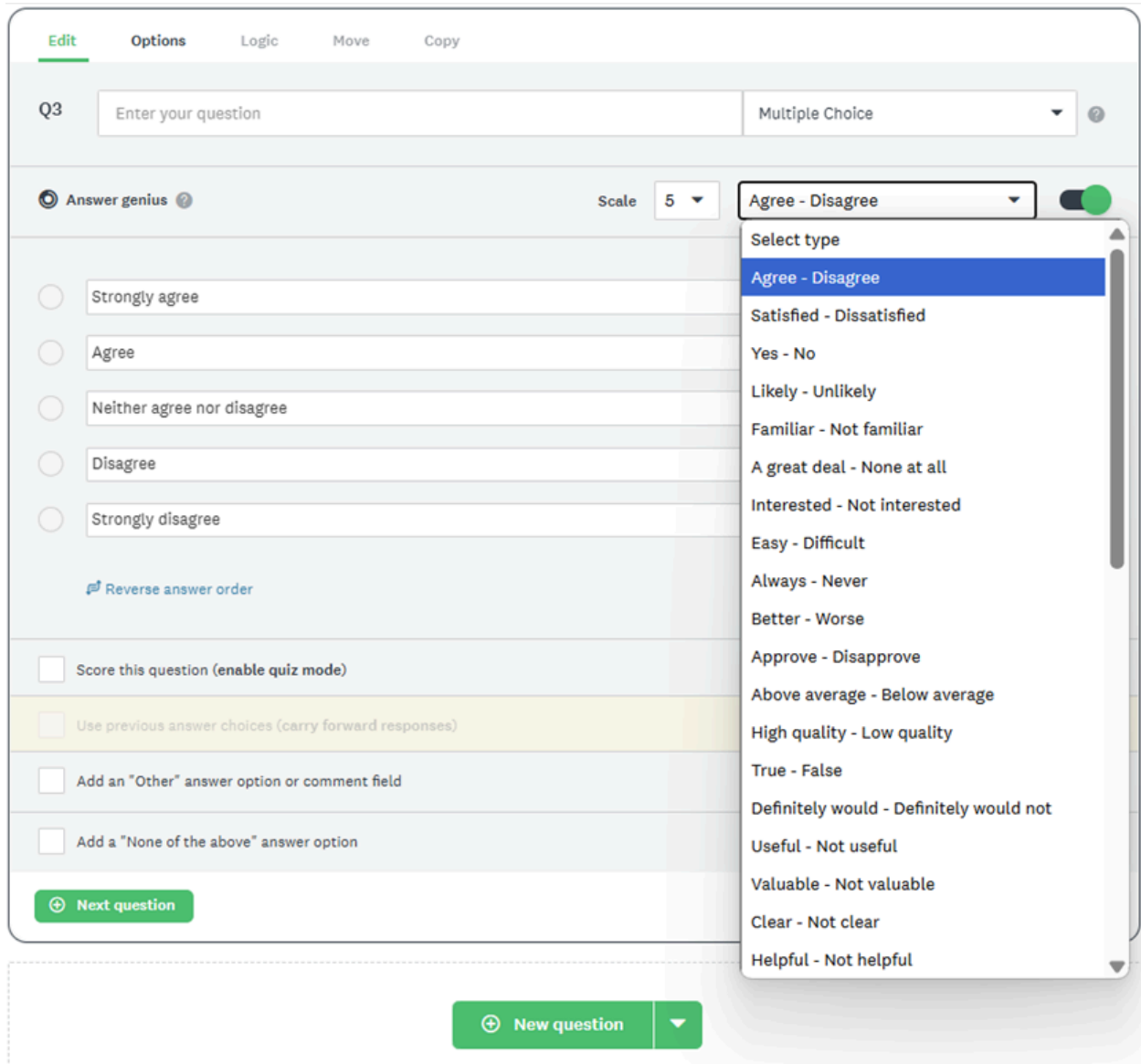


Figure 41 A Screen Shot of Scale Options Available in SurveyMonkey



For instance, to create a survey item with a Likert-type rating scale, you select **Multiple Choice**, choose one of the Likert-style options, and then customize the settings. You may adjust the number of scale points (e.g., 4-, 5-, or 7-point scales), reverse the order of response options (descending or ascending), and checkmark **add an “Other” answer option or comment field** (Figure 42).

Figure 42 A Screen Shot of Multiple-Choice Options Available in SurveyMonkey



Notes

1. SurveyMonkey is a trademark of SurveyMonkey Inc., and this book is not endorsed by or affiliated with SurveyMonkey in any way.

6.5 Chapter activity: How do different online survey systems work?

Google Forms is free to use with a Google account. **Qualtrics** and **SurveyMonkey** allow you to use a free trial version (or your institution may provide you with licensed access). For this activity, choose two of the three online survey platforms (or any other comparable survey software to which you have access). Using each platform, upload the survey questionnaire you designed in the Chapter 5 Activity. As you build your survey in each system, explore their available features. Then, **compare** and **contrast** the two systems you selected based on design tools, user interface, flexibility, and available support.

You may use the evaluation form provided in Exhibit 17 as a starter, but you are encouraged to add more survey items that you feel are relevant in evaluating the quality and utility of the systems. You may also modify the rating scales as needed. You should complete **one evaluation form for each system** (two total). Report your comparative evaluation results.

Exhibit 17 Survey System Evaluation Form

Product Name:

How difficult or easy is it to perform the following tasks?

1. start using the system (involving the sign-up, log-in process)	Very difficult 1	2	3	4	Very easy 5	N/A o
<ul style="list-style-type: none"> Describe pros and cons: 						
2. learn to use the features provided in the system	1	2	3	4	5	o
<ul style="list-style-type: none"> Describe pros and cons: 						
3. learn to create the first survey questionnaire from scratch (with at least five survey items)	1	2	3	4	5	o
<ul style="list-style-type: none"> Describe pros and cons: 						
4. create survey questionnaires by using the system's templates	1	2	3	4	5	o
<ul style="list-style-type: none"> Describe pros and cons: 						
5. find support documents (when needing assistance)	1	2	3	4	5	o
<ul style="list-style-type: none"> Describe pros and cons: 						
6. find answers to your questions from the support documents	1	2	3	4	5	o
<ul style="list-style-type: none"> Describe pros and cons: 						
7. Overall, how dissatisfied or satisfied with the system are you?	Very dissatisfied 1	2	3	4	Very satisfied 5	NA o
<ul style="list-style-type: none"> Describe the overall quality/value of this product and your recommendations: 						

7. PREPARE SURVEY DATA FOR ANALYSIS AND VISUALIZE DATA

7.0 Introduction

You may have administered survey questionnaires with closed-ended questions to collect data during projects involving evaluation, performance analysis, or instructional design. In the final **How** step of the survey design and data analysis process (Figure 43), the focus shifts to analyzing the quantitative data obtained from your surveys. One commonly used tool for this purpose is Excel, which offers practical features for analyzing quantitative survey data.

Figure 43 The How Step in the Survey Design and Data Analysis Process

Determine how you will analyze the data



Learning Objectives

After completing this chapter, you will be able to:

- Explain the purpose of data screening
- Identify common methods used in data screening
- Describe the benefits of visually presenting data

- Identify data visualization techniques using Excel

7.1 Data Screening

After you obtain data from your survey questionnaire, one of the first things you need to do is to screen the data to check if all data looks ready to be analyzed or what needs to be done if you see any abnormalities in your data. These initial data screening tasks include:

- Checking the survey return rate
- Handling missing or incomplete responses
- Identifying and addressing outliers
- Detecting and managing invalid or careless responses

7.1.1 Survey return rate

The first step in your data screening is to check the **survey return rate**. This begins by recalling how many individuals you invited to participate in your survey. You may have invited the entire target population or a sample from that population (refer to Chapter 1 for a review of population vs. sample).

For example, imagine you want to investigate employee engagement within your organization, which has a total of 2,000 employees. This total represents your **target population**. You could invite all 2,000 employees to complete your survey, or invite a **sample** of the population by using methods such as simple random sampling or stratified random sampling (see Chapter 1 for more on sampling methods).

Regardless of your sampling approach, participation in surveys is typically voluntary, so you cannot assume that everyone you invited to your survey will submit their data. Therefore, it is important to track how many people **you invited** versus how many **actually completed** the survey.

The survey return rate can be affected by several factors, including the type and length of your survey, organizational culture, the survey sponsorship, how relevant the survey topic is to respondents, timing and duration of the survey, incentives and follow-up strategies. Low return rates (10-30%) are common among busy professionals, especially when they experience **survey fatigue** (i.e., a state of being tired

of receiving frequent survey requests). However, with strong organization support, meaningful incentives, and a participatory culture, you may achieve return rates of 50% or higher.

In your survey report, you want to document the number of people you invited to the survey, number of people who completed the survey, and the calculated survey return rate. Below is an example of how to present this in your report:

A random sample of 460 employees from a total of 2,000 employees was invited to voluntarily participate in the survey. Of those invited, 200 completed the survey, resulting in a return rate of 43.5%.

7.1.2 Missing data (incomplete responses)

A survey return rate of 43.5% does not necessarily mean that all submitted surveys contain **complete responses**. When participation is voluntary, it's common that you do not set your survey items to be required (that is, the survey respondents may skip survey items if they want to or accidentally). This leads to **missing data**—visible as empty cells in your Excel data file.

It is important to distinguish missing data from nonresponse data. **Missing data** refers to incomplete responses from those who submitted the survey. In contrast, **nonresponses** refer to individuals (in the above example, 56.5%) who did not complete the survey at all.

For example, suppose your survey consists of five survey items, and 200 employees submitted their responses ($n = 200$). Some respondents skipped one or more items, resulting in incomplete data (each respondent is referred to as a case). Now, you need to decide how to handle these missing values. Below are several common situations and corresponding strategies. You may use a combination of multiple methods.

Situation 1: A small number of respondents (cases) have many missing values

You may exclude these cases from your analysis. For example, if two respondents did not complete 4 out of 5 items (80% missing), it is reasonable to remove the two cases. Then, your adjusted sample size is: $n = 200 - 2 = 198$.

Situation 2: Many respondents (cases) skipped the same survey item

It's possible that many respondents did not complete the same survey item—for example, Q5 in your survey. Maybe Q5 was irrelevant or unclear to them. Then, you may consider removing Q5 from your analysis. In other words, you analyze 200 cases on 4 survey items (excluding Q5).

Situation 3: A small number of respondents (cases) missed different survey items

Usually, missing data are spread out in your Excel data file. Let's say, 15 respondents did not complete one or two survey items (all different survey items). There are a few ways to handle the missing data:

Replace missing data with average scores:

This method is used when there are only a couple of missing data, and it is performed on each survey item. For example, let's say one respondent skipped Q2 (cell B3), and another respondent skipped Q5 (cell E4) (Figure 44). Calculate the average score of the complete data of Q2 (let's say, average = 3.5) and enter the average score (3.5) in cell B3 and calculate the average score of the complete data of Q5 (let's say, average = 3.0) and enter the average score (3.0) in cell E4.

Figure 44 *Replace with Average Scores*

Average = 3.5

	A	B	C	D	E
1	Q1	Q2	Q3	Q4	Q5
2	2	3	5	2	2
3	3		3	2	2
4	3	3	2	3	
5	4	4	3	3	4
6	4	4	5	4	4

Average = 3.0

Listwise deletion:

You delete the cases that contain any missing value (Figure 45); that is, when 15 out of 200 respondents have at least one missing data, your sample size $n = 200 - 15 = 185$.

Figure 45 Listwise Deletion

	A	B	C	D	E
1	Q1	Q2	Q3	Q4	Q5
2	2	3	5	2	2
3	3		3	2	2
4	3	3	2	3	
5	4	4	3	3	4
6	4	4	5	4	4

Delete

Pairwise deletion:

You analyze individual survey items with all available cases without deleting the cases that contain any missing value. You leave the missing data cell empty and calculate average scores of individual survey items (Figure 46).

Figure 46 Pairwise Deletion

	A	B	C	D	E
1	Q1	Q2	Q3	Q4	Q5
2	2	3	5	2	2
3	3		3	2	2
4	3	3	2	3	

199	3	4	3	4	4
200	4	4	3	4	4
201	3	4	5	4	2

In doing so, you may need to present average scores of individual survey items with different sample sizes, as shown below:

Survey Item	Mean
Q1	4.3 (n = 200)
Q2	3.5 (n = 199)
Q3	3.9 (n = 198)
Q4	3.4 (n = 200)
Q5	3.0 (n = 199)

7.1.3 Outliers

Outliers are extreme values, unusually high or low compared to the rest of your dataset. Because they can distort your results and conclusions, it is important to identify outliers and decide how to handle them. In some cases, you may exclude the case with an outlier; in others, you may investigate further before making a decision. Below are a couple of examples.

Time for completion:

Outliers can be found when checking whether respondents spent a reasonable amount of time completing the survey. For example, if your web-based survey platform records response times, and you expect the survey to take 2-3 minutes to complete, but find that a few respondents finished in just 24 seconds, these are likely outliers. If you estimate that it is almost impossible to complete the survey in such a short time (if they carefully read each survey question before they clicked on an option), then you may reasonably assume these responses reflect low engagement. In such cases, you might treat them as invalid data and exclude them from analysis.

	A
1	Duration (in seconds)
2	142
3	182
4	155
5	210
6	165
7	155
8	24
9	245
10	471
11	614
12	205

	A
1	Age
2	22
3	31
4	26
5	34
6	26
7	55
8	32
9	30

Unusual values (e.g., age):

Outliers can also emerge in specific survey responses due to data entry errors, misinterpretation, or other factors. For example, you asked respondents to enter their age and expected values between 20 and 40, but one response is “55” and another is “333”. In these cases, determine whether the outlier is a valid but unusual response (e.g., a genuinely older participant), or it is a typo or data entry error. Based on your evaluation, you can decide to keep the outlier, correct it (if possible), or treat it as missing data and exclude it from analysis.

7.1.4 Invalid data (careless responses)

When using 4-, 5-, or 7-point Likert-type response scales, you may observe unusual response patterns. For example, if most respondents selected 4s and 5s on a 5-point scale, but a couple of respondents selected 1s and 2s, those low scores may initially seem like outliers. However, these are not true statistical outliers—they may still be

plausible responses. It is possible that some respondents misinterpreted the direction of the response scale—for instance, selecting a 1 (“Poor”) when they actually intended to select a 5 (“Excellent”). This kind of mistake is considered a **careless response**. Below are several situations and ways to detect careless responses.

Contradictions between quantitative and qualitative data:

Careless responses may become evident when you compare quantitative data with corresponding open-ended (qualitative) responses. For example, if a respondent selects mostly 1s on the scale (indicating dissatisfaction), but writes very positive comments in the open-ended section, this inconsistency suggests the respondent may have misread the scale or responded carelessly (Figure 47). If you have **clear and compelling evidence** of a careless response, you may consider correcting the data—for example, by **reverse-coding** the quantitative responses to better align with the qualitative feedback.

Figure 47 Contradicting Quantitative and Qualitative Data

	A	B	C	D	E	F
1	Q1	Q2	Q3	Q4	Q5	Comments
2	1	2	1	1	1	Great program! Enjoyed it.

Where 1 = Poor and 5 = Excellent
Positive comments

Contradictory responses from paired survey items:

You may detect invalid data by using paired survey items that measure similar or opposite constructs. For example:

- I think the course is well organized.
Strongly disagree 1 2 3 4 5 Strongly agree
- The course organization was:
Very confusing 1 2 3 4 5 Very clear

You would expect responses to both items to be aligned—if someone rates the course as well organized, they should also rate the organization as clear.

You may also use positively and negatively worded item pairs, where the negatively worded item needs to be reverse-coded:

- I think the course is clearly organized.
Strongly disagree 1 2 3 4 5 Strongly agree
- I found the course organization to be confusing.
Strongly disagree 1 2 3 4 5 Strongly agree (← to be reverse-coded)

If responses to such paired items contradict each other (e.g., a high score on both the positive and negative item), this inconsistency may indicate carelessness or confusion. In such cases, it's helpful to examine other data from the same respondent for additional abnormalities.

Zero standard deviation:

Another method to detect potentially careless responses is to look for zero standard deviation across all survey items for a respondent. This occurs when a respondent selects the same rating for every question. This can be easily identified in Excel by calculating the standard deviation (SD) for each respondent's row. A standard deviation of 0 means there was no variation in their responses (Figure 48).

Figure 48 Zero Standard Deviation Detected

	A	B	C	D	E	F
1	Q1	Q2	Q3	Q4	Q5	SD
2	2	3	5	2	2	1.30
3	4	4	4	4	4	0.00
4	3	3	2	3	4	0.71
5	4	4	3	3	4	0.55
6	4	4	5	4	4	0.45

However, a zero standard deviation alone is not sufficient to conclude the data is invalid. It's possible that the uniform responses genuinely reflect the respondent's opinion. But if this is combined with other red flags—such as very short completion time—you may reasonably decide to treat the data as careless and exclude it.

To wrap up, in this example, you received 200 completed surveys from 460 employees, resulting in a 43.5% return rate. After screening the data for missing values, outliers, and careless responses, you may end up replacing or excluding certain cases. These decisions should be clearly documented in your report. Below is an example of a statement you may include:

During the data screening process, three missing data were replaced with item averages. Two cases were removed due to evidence of careless responses—each completed the survey in under 30 seconds and had zero standard deviation across items. The final sample size for analysis was: $n = 198$.

7.2 Chapter activity: How would you do your data screening?

Imagine you administered a survey questionnaire designed to measure employee engagement levels, consisting of 10 survey items. See the first tab “Ch.7.2 Data screening” in [Ch7 Data Screening and Charts.xlsx](#) (Figure 49), which displays the dataset you collected from 50 employees.

You are asked to perform your data screening based on the following criteria:

1. If any case has 7 or more missing values, delete the case.
2. If any case has 1, 2, or 3 missing values, replace the missing data with the case average scores.
3. Delete any cases with careless responses indicated by both:
 - Duration under 30 seconds, and
 - Standard deviation of zero

The “Ch.7.2 Data screening” tab already presents the Standard Deviation values in the “M” column and the average scores of individual survey items in the last (52nd) row.

Report your decisions by filling in the following:

1. Based on Criterion #1, I will delete case number(s) ____.
2. Based on Criterion #2, I will replace the case number ____’s cell ____ (e.g., C10) with an average score of ____.
3. Based on Criterion #3, I will delete case number(s) ____.

Figure 49 An Excel Data File for Data Screening

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Case number	Duration (in seconds)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Standard Deviation
2	1	235	4	4	4	4	4	4	4	4	4	4	0.000
3	2	212	5	4	5	5	5	4	5	5	5	5	0.422
4	3	243	5	4	5	5	4	3	4	4	4	4	0.632
5	4	132	5	5	5	4	4	4	5	5	5	5	0.483
6	5	152	5	5	5	5	5	3	5	5	4	5	0.675
7	6	92	3	1	3	2	3	1	1	1	1	1	0.949
8	7	123	4	4	4	4	4	4	4	4	4	4	0.000
9	8	365	5	5	5	5	5	4	5	5	5	5	0.316
10	9	16	5	5	5	5	5	5	5	5	5	5	0.000
11	10	154	4	5	5	4	4	4	5	5	4	4	0.516
12	11	228	5	5	5	5	5	4	4	5	5	4	0.483
13	12	158	5	4	5	5	3	4	4	4	5	5	0.699
14	13	211	4			5			5				0.577
15	14	259	5	5	5	5	4	4	5	5	5	5	0.422
16	15	1265	5	4	5	5	4	4	5	5	5	5	0.483
17	16	190	5	5	5	5	5	4	5	5	5	5	0.316
18	17	169	4	4	5	5	4	3	4	4	4	4	0.568
19	18	168	5	5	5	5	5	4	4	5	5	5	0.422
20	19	162	5	5	5	4	5	4	5	5	4	5	0.483
21	20	22	5	5	5	5	5	5	5	5	5	5	0.000
22	21	218	5	5	4	5	5	4	4	5	5	5	0.483
23	22	192	5	5	5	5	4	4	5	5	4	5	0.483
24	23	179	5	5	5	5	5	5	5	5	5	5	0.000
25	24	40	5	5	5	5	5	3	5	5	4	5	0.675
26	25	84	5	5	4	5	5	4	5	5	5	5	0.422
27	26	102	5	5	5	5	4	3	5	5	4	5	0.699
28	27	26	5	5	5	5	5	5	5	5	5	5	0.000
29	28	257	5	5	5	5	5	4	5	5	5	5	0.316
30	29	119	5	5	5	5	5	5	5	5	5	5	0.000
31	30	139	5	5	4	5	5	4	3	5	3	5	0.843
32	31	152	4	3	5	5	4	5	5	4	4	5	0.699
33	32	120	5	5	5	5	5	5	5	5	5	5	0.000
34	33	175	5	4	4	4	4	3	5	5	4	5	0.675
35	34	119	5	5	4	5	4	5	4	4	4	4	0.516
36	35	156	5	4	4	5	4	4	4	4	4	5	0.483
37	36	260	5	5	5	5	5	3	4	5	4	5	0.699
38	37	221	5	4	5	5	5	4	5	4	5	5	0.483
39	38	187	5	3	4	5	3	1	3	4	3		1.236
40	39	351	5	4	5	4	4	3	5	5	5	5	0.707
41	40	137	5	4	4	5	4	3	5	5	4	5	0.699
42	41	87	5	4	5	5	4	5	4	4	5	5	0.516
43	42	162	5	5	4	4	4	4	4	4	4	4	0.422
44	43	174	5	4	5	4	4	5	5	4	4	5	0.527
45	44	181	5	4	5	5	5	2	2	2	2	2	1.269
46	45	29	5	4	5								0.577
47	46	2197	5	4	5	5	4	4	5	5	5	5	0.483
48	47	298	5	5	5	5	5	5	5	5	5	5	0.000

7.3 Visualizing data and results

Your data analysis will typically include both **descriptive** and **inferential** statistics.

- **Descriptive statistics** are used to describe or summarize your dataset (such as mean scores, frequencies, or percentages).
- **Inferential statistics** are used to make generalizations or predictions about a population based on data from a sample.

Let's explore several ways to visually present your descriptive statistics.

In this section, we focus on how to visually present **descriptive statistics**, which often involve information about your survey participants—such as age, gender, race, marital status, job title, education level, employment status, years of experience, income, or geographic location. These types of data can be effectively communicated using various types of charts.

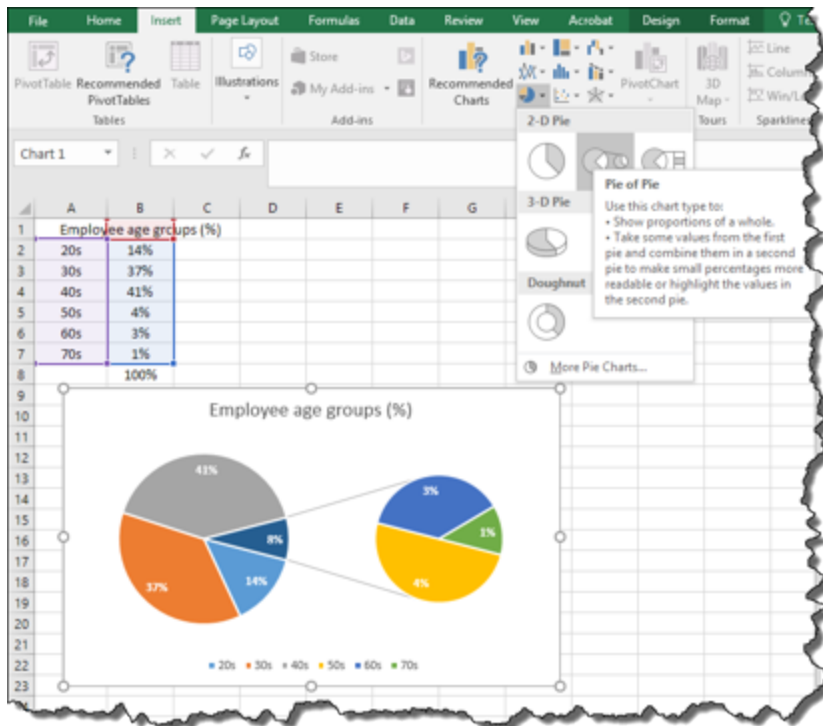
Pie charts

Pie charts are useful for illustrating the proportion (%) of categorical data, such as gender or job title. For example, if you invited a sample of employees (rather than the entire employee population) to complete your survey, you may want to check whether the characteristics of the respondents (e.g., age groups) reflect those of the broader population.

If your questionnaire included a question about age group, you can use a pie chart to show the proportion of respondents in each group.

- You can break out a slice of the pie to highlight a particular category.
- For categories with small percentages, you can use a “pie of pie” chart to improve readability (Figure 50).

Figure 50 A Sample Pie of Pie Chart



Bar graphs

Bar graphs are helpful when you need to:

- Display frequencies of categorical data, such as racial background or job title (Figure 51)
- Compare categorical data, such as gender or location, side by side (Figure 52)

Figure 51 A Sample Bar Graph Showing Employee Racial Background

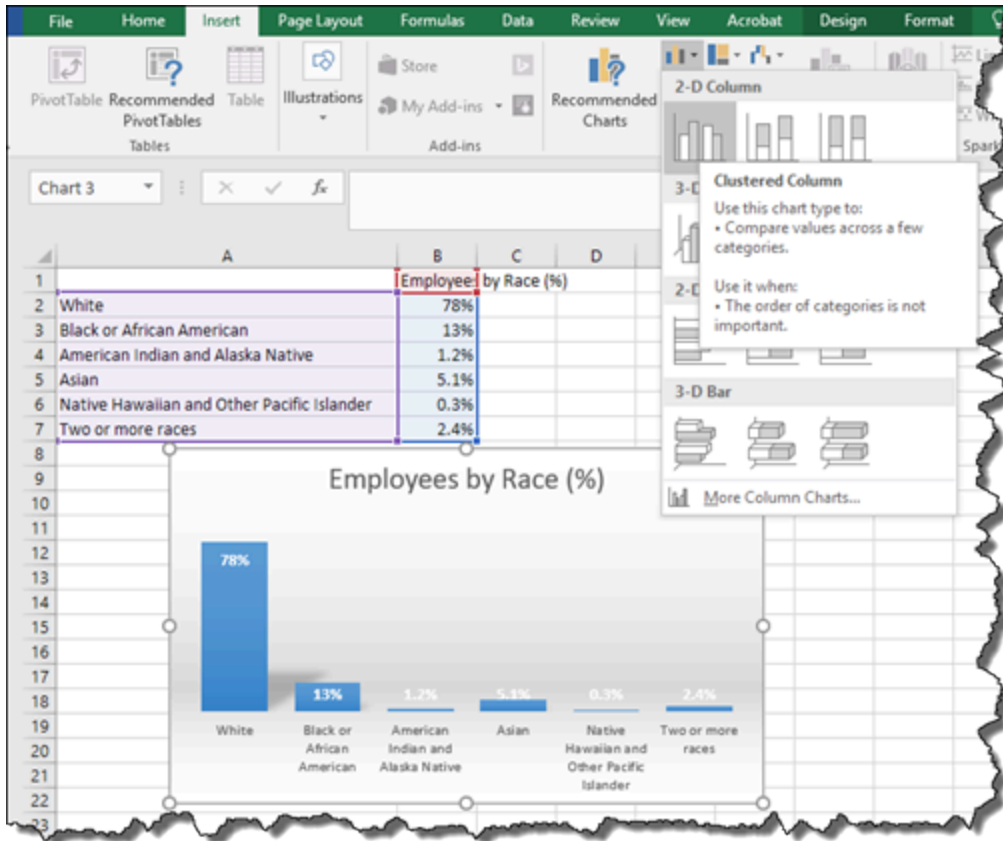
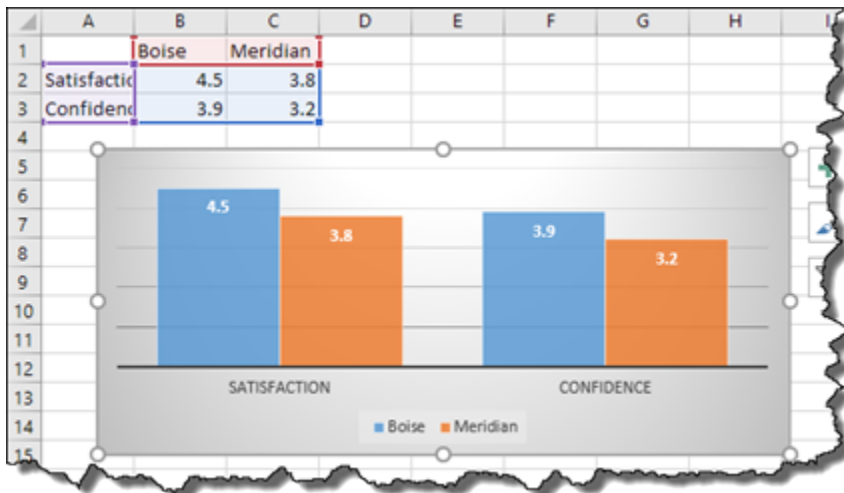


Figure 52 A Sample Bar Graph Comparing Employee Satisfaction and Confidence Levels from Two Locations



Line graphs

Line graphs are useful when you need to:

- Show trends over time (Figure 53)
- Illustrate interactive results such as changes over time by subgroups (e.g., before-and-after results by gender) (Figure 54)

Figure 53 A Sample Line Graph Showing Monthly Sales

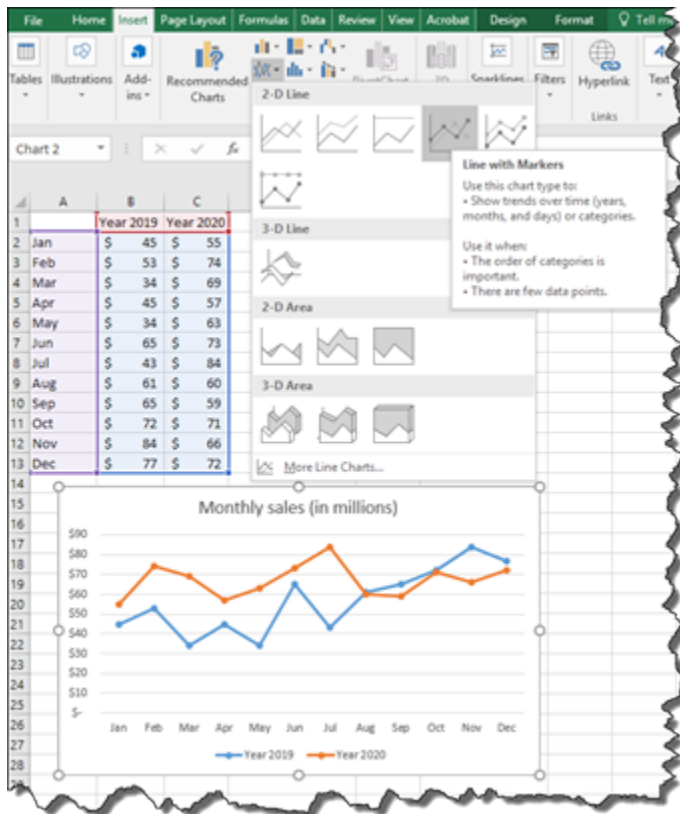
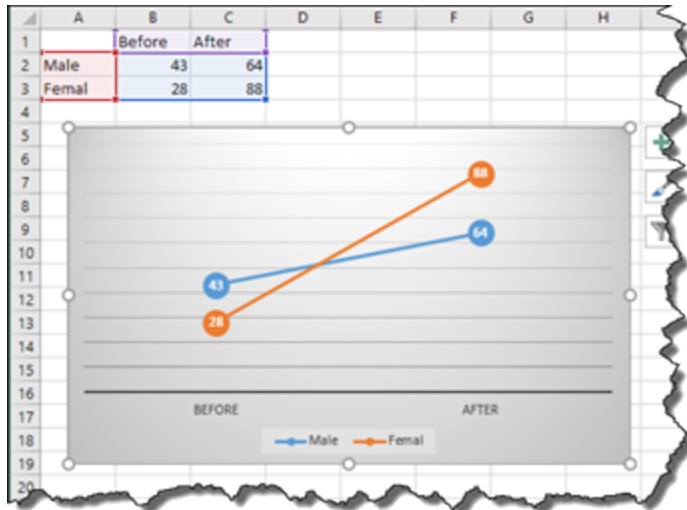


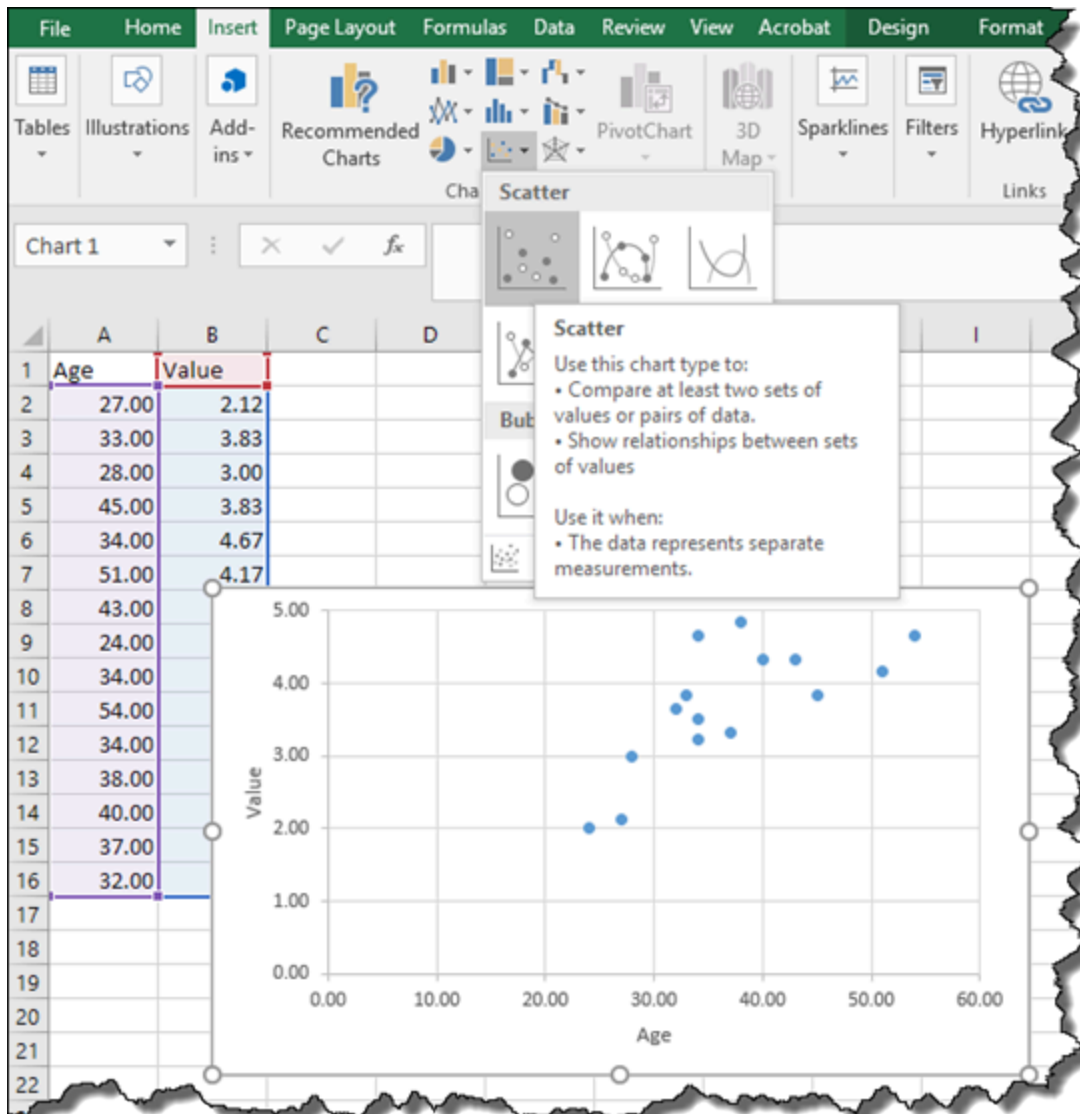
Figure 54 A Sample Line Graph Comparing Before and After Results by Male and Female



Scatter or bubble chart

A scatter or bubble chart is useful when presenting relationships between two sets of data. For example, you might visualize the relationship between employee age and their perceived value of an organizational strategy (Figure 55).

Figure 55 A Sample Scatter Chart Showing the Relationship Between Employee Age and Their Value Perception



7.4 Chapter activity: How would you visually present the data?

In this activity, you will practice creating visual representations of descriptive data using Excel charts.

Open the file [Ch7 Data Screening and Charts.xlsx](#), and use the remaining 6 tabs provided.

As illustrated in Section 7.3, create the following charts in the remaining 6 tabs:

- *Pie chart (age group)*
- *Bar graph (race)*
- *Bar graph (satisfaction)*
- *Line graph (trend)*
- *Line graph (before and after)*
- *Scatter plot (age x value)*

Appendix A

For more practice with data analysis, refer to [Appendix A. Correlation and Group Comparison](#).

About the Author

Seung Youn (Yonnie) Chyung, EdD, is a professor in the Organizational Performance and Workplace Learning (OPWL) program at Boise State University. She currently teaches a graduate course on Evaluation, and has designed and taught other graduate courses, including Survey Design and Data Analysis, Quantitative Research in Organizations, and Foundations of Instructional and Performance Technology.

Dr. Chyung's research focuses on synthesizing and developing evidence-based principles for survey design. She has published a series of practitioner-oriented articles on this topic in collaboration with her research teams composed of OPWL students and graduates.

She is the author of other books, [10-Step Evaluation for Training and Performance Improvement \(Sage\)](#) and [Foundations of Instructional and Performance Technology \(HRD Press\)](#).

In addition to her academic work, Dr. Chyung engages OPWL students and alumni in pro bono consulting projects that support organizations with program evaluations, survey design, and statistical data analysis.

More information about her scholarly activities can be found on her [GoogleScholar](#) page.

Feedback and Suggestions

Your feedback is greatly appreciated and will help us continue improving the book for future learners!

Please use the form below to let us know things that we should improve. If the form does not appear below, please use [this link](#).

Thank you!



One or more interactive elements has been excluded from this version of the text. You can view them online here: <https://boisestate.pressbooks.pub/surveydesign/?p=495>